An Annotated Syntax Reader

Linguistics: The Essential Readings

This series consists of comprehensive collections of classic and contemporary reprinted articles in a wide range of fields within linguistics. The primary works presented in each book are complemented by outstanding editorial material written by key figures in the field. Each volume stands as an excellent resource on its own, as well as an ideal companion to an introductory text.

An Annotated Syntax Reader

Lasting Insights and Questions

Edited by

Richard Kayne, Raffaella Zanuttini, and Thomas Leu

Contents

Acknowledgments

The work on this volume developed over a number of years during which we were able to count on the advice, encouragement, and patience of Tami Kaplan, Julia Kirk, and Danielle Descoteaux from Blackwell, to whom we are very grateful. We are also very grateful to Bob Frank, who readily let us tap his knowledge and appreciation of the theory and the literature. We would further like to express our gratitude to the students who helped us by providing insightful comments on our editorial material, in particular Simon Mauck from Georgetown University, Nathaniel Little and Phoebe Gaston from Yale University, and Gustavo Beritognolo from the Université du Québec à Montréal. We would also like to thank George Davidson for his very detailed and insightful copyediting of the manuscript. Finally, we would like to thank the authors of the articles included in this volume for their positive feedback on the project and their very helpful comments on the introductions.

Text Credits

The editors and publisher gratefully acknowledge the permission granted to reproduce the copyright material in this book:

Chapter 1
Paul Postal, On So-Called "Pronouns" in English. In Francis P. Dinneen (ed,), *Report of the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies* (Washington, DC: Georgetown University Press, 1966, pp. 177–206). [Reprinted in David A. Reibel and Sanford A. Schane (eds), *Modern Studies in English* (Englewood Cliffs, NJ: Prentice-Hall, 1969, pp. 201–224).]

Chapter 2
Joan W. Bresnan, On Complementizers: Toward a Syntactic Theory of Complement Types. In *Foundations of Language* 6 (1970): 297–321.

Chapter 13
Richard K. Larson, On the Double Object Construction. In *Linguistic Inquiry* 19 (1988): 335–391.

Chapter 14
Richard S. Kayne, Facets of Romance Past Participle Agreement. In Paola Benincà (ed.), *Dialect Variation and the Theory of Grammar* (Dordrecht: Foris, 1989, pp. 85–103). [Reprinted in Richard S. Kayne, *Parameters and Universals* (New York and Oxford: Oxford University Press, 2000, Ch. 3).]

Chapter 15
Jean-Yves Pollock, Verb Movement, Universal Grammar and the Structure of IP. In *Linguistic Inquiry* 20 (1989): 365–424.

Chapter 16
Lisa Travis, Parameters of Phrase Structure. In Mark R. Baltin and Anthony S. Kroch (eds), *Alternative Conceptions of Phrase Structure* (Chicago: University of Chicago Press, 1989, pp. 263–279).

Chapter 17
Liliane Haegeman and Raffaella Zanuttini, Negative Heads and the Neg Criterion. In *The Linguistic Review* 8 (1991): 233–251.

Chapter 18
Richard S. Kayne, Romance Clitics, Verb Movement and PRO. In *Linguistic Inquiry* 22 (1991): 647–686. [Reprinted in Richard S. Kayne, *Parameters and Universals* (New York and Oxford: Oxford University Press, 2000, Ch.5).]

Chapter 19
Hilda Koopman and Dominique Sportiche, On the Position of Subjects. In *Lingua* 85 (1991): 211–258.

Chapter 20
Kenneth Hale and Samuel Jay Keyser, On Argument Structure and the Lexical Expression of Syntactic Relations. In Kenneth Hale and Samuel Jay Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger* (Cambridge, MA: MIT Press, pp. 53–109).

Chapter 21
Giuseppe Longobardi, Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form. In *Linguistic Inquiry* 25 (1994): 609–665.

Chapter 22
Anna Szabolcsi, The Noun Phrase. In Ferenc Kiefer and Katalin É. Kiss (eds), *The Syntactic Structure of Hungarian*. Volume 27 of *Syntax and Semantics* (San Diego: Academic Press, 1994, pp. 179–274).

Chapter 23
Filippo Beghelli and Tom Stowell, Distributivity and Negation: The Syntax of *Each* and *Every*. In Anna Szabolcsi (ed.), *Ways of Scope Taking* (Dordrecht: Kluwer, 1997, pp. 71–107).

Chapter 24
Luigi Rizzi, *The Fine Structure of the Left Periphery*. In Liliane Haegeman (ed.), *Elements of Grammar: Handbook of Generative Syntax* (Dordrecht: Kluwer, 1997, pp. 281–337).

Chapter 25
Anna Cardinaletti and Michal Starke, The Typology of Structural Deficiency: A Case Study of the Three Classes of Pronouns. In Henk van Riemsdijk (ed.), *Clitics in the Languages of Europe*, volume 8 of *Empirical Approaches to Language Typology* (Berlin: Mouton de Gruyter, 1999, pp. 145–233).

Chapter 26
Lisa L. S. Cheng and Rint Sybesma, Bare and Not-So-Bare Nouns and the Structure of NP. In *Linguistic Inquiry* 30 (1999): 509–542.

Chapter 27
Anders Holmberg, Remarks on Holmberg's Generalization. In *Studia Linguistica* 53 (1999): 1–39.

Chapter 28
Norbert Hornstein, *Movement and Control*, 1999. In *Linguistic Inquiry* 30 (1999): 69–96.

Chapter 29
Diane Massam, VSO and VOS: Aspects of Niuean Word Order. In Andrew Carnie and Eithne Guilfoyle (eds), *The Syntax of Verb Initial Languages* (New York and Oxford: Oxford University Press, 2000, pp. 97–117).

Chapter 30
Noam Chomsky, Derivation by Phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language* (Cambridge, MA: MIT Press, 2001, pp. 1–52).

Chapter 31
K. A. Jayaseelan, IP-Internal Topic and Focus Phrases. In *Studia Linguistica* 55 (2001): 39–75.

Chapter 32
Julie Anne Legate, The Configurational Structure of a Nonconfigurational Language. In Pierre Pica (ed.), *Linguistic Variation Yearbook* 1 (Amsterdam and Philadelphia: John Benjamins, 2001, pp. 63–99).

Chapter 33
Richard S. Kayne, Antisymmetry and Japanese. *English Linguistics* 20 (2003): 1–40. [Reprinted in Richard S. Kayne, *Movement and Silence* (New York and Oxford: Oxford University Press, 2005, Ch. 9).]

Chapter 34
Guglielmo Cinque, "Restructuring" and Functional Structure. In Adriana Belletti (ed.), *Structures and Beyond*. Volume 3 of *The Cartography of Syntactic Structures* (New York and Oxford: Oxford University Press, 2004, pp. 132–191.

Chapter 35
Guglielmo Cinque, Deriving Greenberg's Universal 20 and Its Exceptions. In *Linguistic Inquiry* 36 (2005): 315–332.

Introduction

There are lasting insights to be found in each of the papers included here. At the same time, each of these papers gives rise, as in any scientific field, to further questions.

Each paper in this volume is, for reasons of space, presented in excerpted form, and is preceded by an introduction that provides some background and a small set of relevant references. Each paper is also followed by a set of questions intended to encourage the student to explore new lines of thought.

The number of syntactic papers (and books) published in the past half century that contain lasting insights and lasting results far exceeds the number that it was possible to include in a volume of this size. In choosing a set of 35 such papers, the editors were obliged to make arbitrary decisions, one of which was the decision not to include excerpts from books, as opposed to papers. Nor did we attempt to achieve exhaustivity across all subdomains of syntax, however characterized. We did not, either, try to maximize the number of syntacticians represented; as a result, some syntacticians, including one of the editors, are represented by more than one paper. (Had we been allotted twice or three times as many pages, we could readily have added another 35 or 70 papers of equal quality and importance to the 35 that we did include.) Finally, we decided to only include papers published before 2006.

In constructing the excerpts, that is, in shortening the papers to 10–12 pages (which allowed us to include many more papers than would otherwise have been possible), we faced a difficult task that inevitably led to compromises and imperfections. In some cases, the excerpt may not flow as smoothly as did the source paper (and in some cases we may have omitted parts of the paper particularly dear to the author). We did attempt, of course, to preserve readability to whatever extent possible. We eliminated most footnotes, but tried to supply updated references that were missing or not yet known at the time of the original publication.

An excerpt is by definition not exhaustive of the paper it is part of. Each excerpt can thus be taken as an invitation to the student to find and to read the entire paper. In a similar vein, while the introduction and questions that precede and follow each paper mention a certain number of references relevant to that paper, the set of references is not intended to be exhaustive, but rather to be an invitation to further reading, a door to the discovery of other papers on the chosen topic.

The questions that follow each excerpt are not limited to points touched on within the excerpt itself; that is to say, some questions themselves invite the student to turn to the entire paper from which the excerpt has been extracted. These questions number between 15 and 23 per paper, vary quite a bit in length and complexity, and are quite heterogeneous in type. Some of the questions can be answered in a short format and may be used as homework exercises. The majority of questions, however, are genuine research questions, which are meant to stimulate discussion and inspire avenues of research. (Related to this, there is no answer key.)

The introductions to the individual papers will be most useful to those who so far know little about syntax (the beginning student, or the scholar from another field) and the questions most useful to those who know more, or want to know more, or are beginning to do independent research in the field. The introductions typically contain: (i) a beginning paragraph that situates the issues raised in the paper in a broader context and says why they are important to the field; (ii) a description of the main (empirical and theoretical) contributions of the paper; (iii) a final paragraph that tries to highlight one idea that has had a lasting impact or has spurred interesting debate, while giving some references. For some of the older papers, the introduction also has the function of making them somewhat more accessible to today's reader, either by explaining some of the assumptions or by highlighting the core of the idea while leaving aside the details of the implementation, which are no longer current.

The questions are often, to one extent or another, English-oriented, in part because the author of the questions is a native speaker of English, in part because English is at present the single most widely discussed language in the syntactic literature, and in part because English is probably the language that is on average most familiar to students of syntax.

Students who are not native speakers of English can, if they know English very well, readily pursue the answers to such questions on their own, using their own knowledge of English (and of any other language explicitly mentioned) supplemented by recourse to descriptive grammars, to the syntactic literature, and to discussion with native speakers of English (and other languages). Students who are native speakers of English will have a slight initial advantage in this respect.

However, this initial advantage can be neutralized in at least two ways. First, students who are not native speakers of English can transpose English-oriented questions to their own native language (or languages). In some cases, the transposition, which will allow the student to proceed more quickly and more deeply, will be straightforward. In other cases, it may require substantial inventiveness. In every case, it will be enriching to the student to make the attempt.

Secondly, every one of the questions in this volume, whether English-oriented in its formulation or not, can and should lead the student to ask how other languages can be brought to bear on that question. As many other languages as possible. In the limit, all other languages. In practice, some.

Bringing in multiple other languages will inevitably lead to further questions of the comparative syntax sort. Is a given difference between two (or more) languages related to other differences between those languages? If it is, what theory of syntax is best adapted to capturing that linkage? What form of parameter must underlie that difference? What theory of parameters best characterizes the individual parameters discovered, best characterizes the limits on syntactic variation?

Answers require questions. The questions provided in this book are intended to lead to ever deeper questions, and to ever deeper answers.

In reading the excerpted papers included in this volume, the student will be reading some of the primary syntactic literature of recent decades, going back as far as the mid-sixties (Postal's 1966 paper) and as recently as the mid-2000s (two papers by Cinque). In reading the older among these papers, the student needs to learn to abstract away from those aspects of older theoretical frameworks that have been superseded. At the same time, the student needs to learn to perceive the commonalities between earlier and later frameworks.

Alongside advances in syntactic theory, it is essential to simultaneously see the advances in observational and descriptive adequacy (to use Chomsky's early terminology) that have taken place in the past half century (especially since the rhetoric of the field sometimes puts disproportionate emphasis on the explanatory frontiers at the expense of the observational and descriptive). In that regard, the reading of this volume can be complemented by the reading, for example, of the detailed and solid grammars of Italian, Spanish, Catalan, Basque, and English edited by Renzi, Salvi, and Cardinaletti (1988–1995), Bosque and Demonte (1999), Solà et al. (2002), Hualde and Ortiz de Urbina (2003), and Huddleston and Pullum (2002), respectively. (On the stability and clarity of syntactic data in the general case, from a psycholinguistic perspective, see Sprouse 2011). Broad overviews of syntactic theory can be found in Baltin and Collins (2001) and in Everaert et al.'s *The Blackwell Companion to Syntax* (2006). Broad overviews of comparative syntax, more specifically, can be found in Cinque and Kayne (2005) and Kayne (to appear).

Since the advent of generative syntax, our knowledge and understanding of human language syntax have become qualitatively better, at both descriptive and theoretical levels. On the theoretical side, advances have come in various forms. To take one example, bare phrase structure (external and internal merge) has superseded the less constrained "phrase structure rules + transformations" of the first 20 years of generative syntax. At the same time, it is essential to see that theoretical advances can, and typically do, coexist with theoretical stability (despite the mistaken impression among some that the field reinvents itself completely every few years). Core ideas can remain stable for decades or longer, and we have endeavored to choose papers that contain one or more ideas that have had that sort of long-lasting impact.

As just mentioned, bare phrase structure is newer than, superior to, and more constrained than the older phrase structure rules (plus transformations). That represents an advance. Yet the core notion of constituent structure itself (now expressed by Merge) has remained stable, going back to the beginning of generative syntax (and before). In addition, there is a wide consensus that constituents do not overlap, and that (cf. Kayne 1981) constituent structure is invariably binary branching. Binary branching in turn converges with Anderson and Chung's (1977, Ch. 5 of this volume) stable conclusion to the effect

that VSO languages have a VP from which the verb is extracted; binary branching also converges with Legate's (2003, Ch. 32 of this volume) by now stable conclusion that there are no (and can be no) nonconfigurational languages.

In the area of constituent structure, we can in addition note the long-term solidity of the notion "head of a phrase" that goes back to Chomsky (1970, Ch. 3 of this volume) and that has been extended, in work such as that of Rizzi (1997, Ch. 24 of this volume), to subareas of syntax not considered by him back then. Similarly, though a bit less immediate, is the continuity between the earlier idea of "cycle" (Chomsky 1965) and the newer idea of "phase" (Chomsky 2001, Ch. 30 of this volume), despite the innovative notion of "spellout" associated with the latter.

The existence of locality conditions on movement operations has been a perfectly stable aspect of syntax for a very long time. Moreover, despite evolving differences of detail, there has been a long-standing realization that some or many locality conditions are sensitive to a notion of "intervention," going back to Chomsky's (1964) A-over-A principle (with a containing category intervening to block extraction of a contained category of the same type) and then forward to Rizzi's (1990) relativized minimality and Chomsky's (1995) minimal link condition. As an example of a partially different locality condition, we can take Rizzi's (1980, Ch. 8 of this volume) use of subjacency in an elegant and lasting account of a variety of facts concerning extraction from *wh*-islands in Italian. (Rizzi's account seems likely to stand even as we search for a better understanding of discrepancies between certain other languages and Italian in this area of syntax.) In an even more obvious way, we can cite one subpart of Ross's (1967) coordinate structure constraint, namely the subpart that in English blocks sentences like *Who are you thinking of inviting John and?*, as being crosslinguistically stable, without exception.

The very notion of movement operation is itself common to syntactic theories going back to Chomsky's earliest work, even if the way in which syntactic movement is regulated has changed considerably, from the early use of highly specific transformations to the later probe-goal approach. If we abstract away from the different formal characterizations, we can see, more specifically, that many of the movement operations of very early generative syntax have remained highly stable, for example *wh*-movement, raising to subject position with *seem*, movement of the object to subject position in passives. Pronominal clitic movement of the sort prominent in Romance languages began to be discussed somewhat later, but has remained stable ever since. (Sportiche's 1995 recasting in terms of "pro"-movement leaves the core idea of clitic-type movement intact.) Verb movements of the sort discussed by den Besten (1983), Emonds (1978), and Pollock (1989, Ch. 15 of this volume) have an extremely stable core, in particular if we take them to be cut from largely the same cloth as the VP movements proposed in Massam (2000, Ch. 29 of this volume), Nilsen (2003), and Jayaseelan (2010) (even if in some cases V-movement and VP-movement are competitors).

Wh-movement in interrogatives is widespread and uncontroversial across many languages. Less widespread, at least at first glance (cf. Haegeman and Zanuttini 1991, Ch. 17 of this volume), is negative phrase movement of the sort readily seen in Scandinavian languages (cf. Engels 2012), and in French in the specific case of *rien* ('nothing') (cf. Kayne 1975, Ch. 1). Also less widespread than interrogative *wh*-movement is *wh*-movement in relative clauses of the sort readily seen in English and many other European languages.

On the other hand, some movement operations that were thought to exist in the early years of generative syntax are now seen not to, in the sense that the phenomena they were intended to account for can better be looked at in other terms. Examples are agent-postposing in passives (now to be rethought along the lines of Collins 2005), downward movement from subject position in sentences like *There has arrived a letter for you* (now to be rethought in terms of a lower original position for the superficial subject, as in Perlmutter 1978, 1989, and Burzio 1986), and rightward heavy-NP-shift (now to be interpreted as involving leftward movement instead, as in Larson (1988), Ch. 13 of this volume, and den Dikken 1995). (On sentential extraposition, see Kayne 1998.)

The abandoned movement operations of the previous paragraph had in common that they were rightward movements, which Kayne (1994) argued not to be available to the language faculty, as a consequence of antisymmetry, in his sense. Antisymmetry itself has come to be widely accepted (though there is not a consensus) as a restrictive characterization of the relation between linear and hierarchical order. Languages do not have access (for principled reasons – cf. Kayne 2011) to directionality parameters of the sort made use of especially in the 1980s (cf. Travis 1989, Ch. 16 of this volume). If antisymmetry is correct (and if Baker's 1988 UTAH principle or something like it is correct, and if Cinque's 1999 hierarchy is correct in constraining the crosslinguistic order of external merge), then word order differences across languages must invariably be due to differential applications of movement operations. An extension of the preceding would be that all crosslinguistic morpheme order differences can be traced back to differences in movement, which would fit in with Baker's (1988) solidly grounded view that morpheme order is tightly tied to syntax.

That languages differ in their (morpho)syntax is uncontroversial. We use the term "parameter" to refer to the properties of the language faculty that underlie these differences. An early proposal as to the form of (some) parameters was given by Bach (1971), in terms of the idea that a given language would choose a subset from a universally fixed finite set of transformations. Since then, the idea has come to the fore (and is almost certainly correct) that parameters are rather to be understood in terms of (relatively simple) properties of functional heads (Borer 1984; Chomsky 1995, as opposed to lexical heads – cf. Kayne 2006). (For a lucid discussion of parameters, see Rizzi 2009, 2011.) What is solidly established, in addition, is the fruitfulness of searching for correlations across syntactic differences as a means of establishing a new kind of window into the language faculty (cf. Kayne, to appear).

Returning to movement operations, we know that surface subject position is filled by internal merge/movement in a significant number of cases (passives, unaccusatives, *seem*) and it is virtually certain that surface subject position is in fact filled by internal merge/movement in all cases (cf. Koopman and Sportiche 1991, Ch. 19 of this volume). A related point is that we know with certainty that the number of arguments a predicate has cannot be read off the visible sentential structure in a superficial manner. In a sentence like *You seem to have made a mistake*, there might appear to be two arguments of *seem*, but there is really only one, from within which the subject *you* has been extracted. The same almost certainly holds for *People consider you to be intelligent*, with *you* raising from within the infinitive phrase, as extensively argued by Postal (1974). A third and more surprising example comes from sentences like *You have a little sister*, given Szabolcsi's (1983–1984 and 1994, Ch. 22 of this volume) analysis of possession sentences

(transposed to English in Kayne 1993), in which the possessor *you* clearly originates internal to the phrase *a little sister*, and does not originate as an argument of *have*.

That the appearance of a phrase in surface subject position is not a reliable indication that it originated there is a solid conclusion that has a partial counterpart in Koster's (1978, Ch. 7 of this volume) demonstration that finite subject sentences, as in *That you are right is obvious*, are not actually in subject position proper, but rather in a topic position. That *that you are right* cannot be in subject position suggests in turn that finite sentences cannot by themselves be arguments, but only subparts of arguments, as had been proposed by Rosenbaum (1967) in terms of a deleted *it*.

Rosenbaum's *it*-deletion proposal is to be seen as one instance of a much broader stable conclusion about syntax reached early in the history of generative syntax (and before), namely that the language faculty allows for various syntactically and semantically active elements to be deleted, or, alternatively put, not to be pronounced. These silent elements range from the traces/silent copies of movement operations to the silent pronominals "PRO" and "pro." The former, "PRO," is found as the subject of nonfinite sentences such as infinitives and gerunds. Its presence can be detected in various ways; one striking way involves comparing *We can't decide whether to go on the trip* and **We can't decide if to go on the trip*, and by examining the Romance counterparts of these, as in Kayne (1991, Ch. 18 of this volume). "Pro" is a partially distinct silent pronominal whose properties have been studied by many (see, for example, Rizzi 1982b, Ch. 11 of this volume; 1986).

Another solidly established subtype of silent element is found in VP-deletion sentences like *John hasn't solved it but Mary has*, in which the VP is silent. This type of sentence is available to some languages but not others (for example, French doesn't allow *. . . *mais Marie a*), and the same holds for the silent noun in *John has written three papers but Mary has written four* (whose word-for-word counterpart is likewise impossible in French). The silent VP and silent noun exemplified in this paragraph are silent elements with antecedents (that are found in the first half of their respective sentences).

Other sorts of silent elements, whose study goes back at least to Katz and Postal (1964), can be exemplified by *John has just turned twenty* (with a silent counterpart of *years* that would not be possible in French), as discussed in Kayne (2003), or by the baseball example *They won the game with two home runs in the seventh* (with a silent counterpart of *inning*). These silent elements don't have antecedents in the way that the silent elements of the preceding two paragraphs do, though it's been clear for decades that they are present (see, for example, early discussions of "the recoverability of deletion"). Their properties can be studied in a familiar fashion; putting things another way, their presence can be detected by slightly altering their environments, for example by noting that while **They won the game with two seventh home runs* is impossible, the sentence is much improved if extra material of a certain sort is added, as in ?*They won the game with two top of the seventh home runs*.

Addition of extra material also saves a (certain interpretation of a) sentence in *John criticized him* vs. *John criticized himself*; that is, addition of *self* allows for an anaphoric interpretation that would otherwise not be available. (A proposal that may permit linking these facts about anaphora to the preceding set of facts about silent *inning* is given in Kayne 2002, in a way that may not be compatible with Reuland 2011.) That *self* and the preceding pronoun have separate syntactic status had been proposed by Helke (1973)

and reinforced by Pica (1987), whose argument that the bimorphemic/bisyntactic status of *himself* has an effect on the locality conditions regulating it appears to have been correct. These locality conditions, and more generally put, the existence of locality conditions on reflexives, have been established beyond doubt for a long time (though how closely they should be tied to locality conditions on movement is not clear). The same holds for the existence of so-called Condition B effects in many languages (that is, the effects that prohibit *him* from taking (unstressed) *John* as antecedent in *John criticized him*). The same also holds for so-called Condition C effects, which prohibit (unstressed) *he* from taking *John* as antecedent in *He thinks that John is smart*, and which are found in language after language. (A separate question is whether Condition C is primitive, as in Lasnik 1976, or not, as in Kayne 2002.) As a final example of a stable, solidly established locality effect, we can mention the fact that if a verb agrees with a subject phrase, then that subject must be the verb's own subject and cannot be a higher subject (this would follow from Chomsky's Agree).

As a final example of a long-recognized property of the language faculty, of a rather different sort, we can think of the (virtually certain) fact that all languages have a verb–noun distinction (see Baker 2003; Kayne 2008; Aldridge 2009; Koch and Matthewson 2009).

With its goal of tying together a broad set of fundamental ideas and a series of explicit research questions, this volume should be of use to students and teachers of syntax at an intermediate or advanced undergraduate level (for example, in a second level or second semester course), as well as to students and teachers at the graduate level. It could be used as a source of inspiration for paper topics, or thesis or dissertation topics. This volume may also be of interest to scholars and scientists outside of linguistics proper, for example to philosophers, psychologists, and cognitive scientists who would like to gain a sense of what are some of the main insights and main issues in the field of syntax.

References

Aldridge, E. 2009. Minimalist questions for the nominalist analysis of Tagalog syntax. *Theoretical Linguistics* 35: 51–62.

Anderson, S. R., and S. Chung. 1977. On grammatical relations and clause structure in verb-initial languages. In P. Cole and J. M. Sadock (eds), *Grammatical Relations. Syntax and Semantics* 8, New York: Academic Press, pp. 1–25.

Bach, E. 1971. Questions. *Linguistic Inquiry* 2: 153–166.

Baker, M. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Baker, M. 2003. *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge: Cambridge University Press.

Baltin, M., and C. Collins. 2001. *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell.

den Besten, H. 1983. On the interaction of root transformations and lexical deletive rules. In W. Abraham (ed.), *On the Formal Syntax of the Westgermania. Linguistik Aktuell/Linguistics Today 3*, Amsterdam and Philadelphia: John Benjamins, pp. 47–131.

Borer, H. 1984. *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Dordrecht: Foris.

Bosque, I., and V. Demonte. 1999. *Gramática descriptiva de la lengua española (3 volumes)*. Madrid: Espasa.

Burzio, L. 1986. *Italian Syntax: A Government–Binding Approach*. Dordrecht: Reidel.

Chomsky, N. 1964. *Current Issues in Linguistic Theory*. The Hague: Mouton.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. 1970. Remarks on nominalization. In R. Jacobs and P. Rosenbaum (eds), *Readings in English Transformational Grammar*, Waltham, MA: Ginn, pp. 184–221.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, pp. 1–52.

Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York and Oxford: Oxford University Press.

Cinque, G., and R. S. Kayne. 2005. *The Oxford Handbook of Comparative Syntax*. New York and Oxford: Oxford University Press.

Collins, C. 2005. A smuggling approach to the passive in English. *Syntax* 8: 81–120.

den Dikken, M. 1995. *Particles: On the Syntax of Verb-Particle, Triadic, and Causative Constructions*. New York: Oxford University Press.

Di Sciullo, A. M., and C. Boeckx (eds). 2011. *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford Studies in Biolinguistics. Oxford and New York: Oxford University Press.

Emonds, J. E. 1978. The verbal complex V'-V in French. *Linguistic Inquiry* 9: 151–175.

Engels, E. 2012. Scandinavian negative indefinites and cyclic linearization. *Syntax* 15: 109–141.

Everaert, M., H. van Riemsdijk, R. Goedemans, and B. Hollebrandse (eds). 2006. *The Blackwell Companion to Syntax*. Oxford: Blackwell.

Haegeman, L., and R. Zanuttini. 1991. Negative heads and the NEG criterion. *The Linguistic Review* 8: 233–251.

Helke, M. 1973. On reflexives in English. *Linguistics* 106: 5–23.

Hualde, J. I., and J. Ortiz de Urbina. 2003. *A Grammar of Basque*. Berlin and New York: Mouton de Gruyter.

Huddleston, R., and G. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Jayaseelan, K. A. 2010. Stacking, stranding, and pied-piping: A proposal about word order. *Syntax* 13: 298–330.

Katz, J. J., and P. M. Postal. 1964. *An Integrated Theory of Linguistic Descriptions*. Cambridge, MA: MIT Press.

Kayne, R. S. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

Kayne, R. S. 1981. Unambiguous paths. In R. May and J. Koster (eds), *Levels of Syntactic Representation*, Dordrecht: Foris, pp. 143–183. [Reprinted in Kayne (1984, Ch. 7).]

Kayne, R. S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, R. S. 1991. Romance clitics, verb movement and PRO. *Linguistic Inquiry* 22: 647–686. [Reprinted in Kayne (2000, Ch. 5).]

Kayne, R. S. 1993. Toward a modular theory of auxiliary selection. *Studia Linguistica* 47: 3–31. [Reprinted in Kayne (2000, Ch. 7).]

Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. S. 1998. A note on prepositions and complementizers. Paper contributed to the Chomsky Virtual Celebration. [Printed in revised form in Kayne (2000, Ch. 15).]

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2002. Pronouns and their antecedents. In S. D. Epstein and T. D. Seely (eds), *Derivation and Explanation in the Minimalist Program*, Oxford: Blackwell, pp. 133–166. [Reprinted in Kayne (2005, Ch. 6).]

Kayne, R. S. 2003. Silent years, silent hours. In L.-O. Delsing, C. Falk, G. Josefsson, & H. Á. Sigurðsson (eds), *Grammar in Focus. Festschrift for Christer Platzack*, Lund: Wallin and Dalholm, vol. 2, pp. 209–226. [Reprinted in Kayne (2005, Ch. 10).]

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2006. On parameters and on principles of pronunciation. In H. Broekhuis, N. Corver, R. Huibregts, U. Kleinhenz, and J. Koster (eds), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, Berlin and New York: Mouton de Gruyter, pp. 289–299. [Reprinted in Kayne (2010, Ch. 4).]

Kayne, R. S. 2008. Antisymmetry and the lexicon. In J. Rooryck and J. van Craenenbroeck (eds), *Linguistic Variation Yearbook 8*, Amsterdam and Philadelphia: John Benjamins, pp. 1–31. [Reprinted in Kayne (2010, Ch. 9) and in Di Sciullo and Boeckx (2011, Ch. 14).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2011. Why are there no directionality parameters? In M. B. Washburn, K. McKinney-Bock, E. Varis, A. Sawyer, and B. Tomaszewicz (eds), *Proceedings of the 28th West Coast Conference on Formal Linguistics*, Somerville, MA: Cascadilla Proceedings Project, pp. 1–23.

Kayne, R. S. to appear. Comparative syntax. *Lingua*.

Koch, K., and L. Matthewson. 2009. The lexical category debate in Salish and its relevance for Tagalog. *Theoretical Linguistics* 35: 125–147.

Koopman, H., and D. Sportiche. 1991. On the position of subjects. *Lingua* 85: 211–258.

Koster, J. 1978. Why subject sentences don't exist. In S. J. Keyser (ed.), *Recent Transformational Studies in European Languages*, Cambridge, MA: MIT Press, pp. 53–64.

Larson, R. 1988. On the double object construction. *Linguistic Inquiry* 19: 335–391.

Lasnik, H. 1976. Remarks on coreference. *Linguistic Analysis* 2: 1–22.

Legate, J. A. 2003. Reconstructing nonconfigurationality. In A. M. Di Sciullo (ed.), *Asymmetry in Grammar: Syntax and Semantics. Linguistik Aktuell/Linguistics Today* 57, Amsterdam and Philadelphia: John Benjamins, pp. 99–116.

Massam, D. 2000. VSO and VOS. Aspects of Niuean word order. In A. Carnie and E. Guilfoyle (eds), *The Syntax of Verb Initial Languages*, New York and Oxford: Oxford University Press, pp. 97–117.

Nilsen, Ø. 2003. *Eliminating positions: the syntax and semantics of sentence modification.* Doctoral dissertation, Utrecht University.

Perlmutter, D. 1978. Impersonal passives and the unaccusativity hypothesis. In J. Jaeger (ed.), *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society, pp. 157–189.

Perlmutter, D. 1989. Multiattachment and the unaccusative hypothesis: The perfect auxiliary in Italian. *Probus* 1: 63–119.

Pica, P. 1987. On the nature of the reflexivization cycle. In J. McDonough and B. Plunkett (eds), *Proceedings of NELS 17*, Amherst: GLSA, University of Massachusetts, pp. 483–499.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Postal, P. 1966. On so-called 'pronouns' in English. In F. P. Dinneen (ed.), *Report on the 17th Annual Round Table Meeting on Linguistics and Language Studies*, Washington, DC: Georgetown University Press, pp. 177–184. [Reprinted as Postal (1969).]

Postal, P. 1969. On so-called 'pronouns' in English. In D. Reibel and S. Schane (eds), *Modern Studies in English: Readings in Transformational Grammar*, Englewood Cliffs, NJ: Prentice-Hall, pp. 201–224.

Postal, P. 1974. *On Raising*. Cambridge, MA: MIT Press.

Renzi, L., G. Salvi, and A. Cardinaletti. 1988–1995. *Grande grammatica italiana di consultazione* (3 volumes). Bologna: Il Mulino.

Reuland, E. J. 2011. *Anaphora and Language Design*. Cambridge, MA: MIT Press.

Rizzi, L. 1980. Violations of the Wh island constraint in Italian and the subjacency condition. *Journal of Italian Linguistics* 5: 157–195. [Reprinted in Rizzi (1982a, Ch. 2).]

Rizzi, L. 1982a. *Issues in Italian Syntax*. Dordrecht: Foris.

Rizzi, L. 1982b. Negation, Wh-movement, and the null subject parameter. In Rizzi (1982a, Ch. 4).

Rizzi, L. 1986. Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17: 501–557.

Rizzi, L. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 281–337.

Rizzi, L. 2009. Some elements of syntactic computations. In D. Bickerton and E. Szathmáry (eds), *Biological Foundations and Origin of Syntax*, Cambridge, MA: MIT Press, pp. 63–87.

Rizzi, L. 2011. On the elements of syntactic variation. In *Working Papers of Centro Interdipartimentale di Studi Cognitivi sul Linguaggio (CISCL)*, Università di Siena. Distributed by MIT Working Papers in Linguistics.

Rosenbaum, P. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: MIT Press.

Ross, J. R. 1967. *Constraints on variables in syntax*. Doctoral dissertation, MIT. [Published 1986 as *Infinite Syntax!* by ABLEX, Norwood, NJ.]

Solà, J., M. R. Lloret, J. Mascaró, and M. Pèrez Saldanya. 2002. *Gramàtica del català contemporani (3 volumes)*. Barcelona: Empúries.

Sportiche, D. 1995. Clitic constructions. In J. Rooryck and L. Zaring (eds), *Phrase Structure and the Lexicon*, Dordrecht: Kluwer, pp. 213–276.

Sprouse, J. 2011. A test of the cognitive assumptions of magnitude estimation: Commutativity does not hold for acceptability judgments. *Language* 87: 274–288.

Szabolcsi, A. 1983/84. The possessor that ran away from home. *The Linguistic Review* 3: 89–102.

Szabolcsi, A. 1994. The noun phrase. In F. Kiefer and K. É. Kiss (eds), *The Syntactic Structure of Hungarian. Syntax and Semantics 27*, San Diego: Academic Press, pp. 179–274.

Travis, L. 1989. Parameters of phrase structure. In M. Baltin and A. Kroch (eds), *Alternative Conceptions of Phrase Structure*, Chicago and London: University of Chicago Press, pp. 263–279.

1

On So-Called "Pronouns" in English

Paul Postal

1966

1.1   Introduction

How do personal pronouns (like *he* and *she*) and reflexives (like *himself* and *herself*) relate to noun phrases (like *the man* and *a woman*) in terms of phrase structure? If we think that they are all noun phrases, and that *the man* consists of a determiner followed by a noun, should *he* be viewed as a determiner (possibly followed by a null noun) or as a noun (possibly preceded by a null determiner)? Postal's article addresses this question and argues that pronouns like *he* are determiners, followed by a noun which is either deleted or, when followed by a restrictive relative clause, realized as *one*, whereas reflexives like *himself* consist of a determiner (*him*) plus the noun *self*.

Postal's proposal can be broken down into three main components. The first is that what we call pronouns are not in fact a subtype of noun. The second is that pronouns are a subtype of determiner, closely related to definite articles. The third is that pronouns, though not themselves nouns, are followed by a noun that is usually not pronounced.

Postal's view is that pronouns, like canonical determiners such as *the*, are added in the course of the syntactic derivation, and that their phonological realization depends on the features of the head noun. For example, if the head noun consists of a set of features that includes [+Pro, +Human, +Definite, +Masculine, +III, −II, −I, +Nominative], the determiner is *he*. But if the head noun consists of the features [+Pro, +Reflexive, +Human, +Definite, +Masculine, +III, −II, −I, +Genitive], the determiner is *him* and the resulting form *himself*.

One advantage of this proposal is that it straightforwardly accounts for the fact that pronouns behave like definite noun phrases in many ways. This is because a

pronoun is viewed as one member of the set of determiners that can lexicalize a definite feature. A second advantage is that it correctly predicts that pronouns should be able to co-occur with nouns, as other determiners do. This is in fact what we see in cases such as *we men, you guys, we honest policemen, you amusing comedians*, etc. In predicting forms like *he-one, we-ones, you-ones*, alongside the familiar *the one(s)* followed by a restrictive relative clause, this proposal also provides a natural treatment of the forms *we'uns, us'uns, you'uns* that are attested in certain varieties of English: they are simply viewed as resulting from the absence of a rule deleting the nonreflexive head noun *one*. Moreover, this approach to pronouns, which puts *they, them, their*, and *theirs* in the same class as *the, this, that, these*, and *those*, is consistent with the observation that this is the class of elements where we find a voiced interdental fricative /ð/. Finally, Postal's proposal makes the identity of pronouns and definite determiners found in Romance (e.g., Italian *la, lo*, Spanish *el*) seem entirely natural.

The development of the DP hypothesis in the 1980s (cf. Abney 1987 and, in this volume, Longobardi 1994, Ch. 21, and Szabolcsi 1994, Ch. 22) allowed an elegant integration of the main ideas of this article into contemporary syntactic theory. This, together with the strength of the empirical support for this proposal, contribute toward its being a continuing source of inspiration for much syntactic and semantic discussion of nominal pro-forms (e.g., Déchaine and Wiltschko 2002; Elbourne 2005; Roehrs 2009; among others).

1.2 From "ON SO-CALLED 'PRONOUNS' IN ENGLISH"

A Introduction

The following is an informal discussion of certain regularities in the syntactic behavior of forms traditionally called 'pronouns' in discussions of English syntax. By informal I mean that, although the analysis suggested involves a number of highly complex grammatical rules and a very special conception of the theory of grammar, no attempt has been made here to formulate or present any of the rules in their correct form. Nor is very much said about the theoretical assumptions these require. My aim is the much weaker one of trying to suggest that a class of facts requires that English grammar be formulated in such a way that it can contain such rules.

Our traditional lore about English grammar [Jespersen 1961, Part VII: 125–126; Curme 1977, Vol. II: 557; Long 1961: 338–356] recognizes a class of forms often called 'pronouns' or 'personal pronouns' which include *I, we, you, he, she, it, they*. At the start we may ignore for simplicity the various case forms *us, your, him*, etc., as well as reflexives, although these will become crucial later. Very often it was said that such forms 'stand for' or 'replace' or 'are substitutes for' previously mentioned or understood noun forms. Certain modern students of English such as Robert Allen have noted [in a paper read to the Linguistic Society of America], essentially correctly, that in many ways such forms actually 'replace' whole noun phrases (henceforth NP) rather than nouns, since they cannot occur with articles, relative phrases, and other elements which can occur in the same NP with ordinary nouns.

Compare:

(1)   the young girl said that she would go

where on one reading *she* can be said to 'stand for' the whole NP *the young girl* with:

(2)   the large girl can't stand the small one

where *one* can only be said to 'stand for' the noun *girl*. However, as I argue later, this contrast is a bit misleading since there is reason to assume that the form *one* or its variants is also relevant at one stage to the 'replacement' which occurs in sentences like (1).

Early transformational descriptions of English have shown that the vague and unclear traditional notion of 'stand for' can, in its <u>sentence internal</u> meaning, be precisely formalized by transformational derivation. Thus in a transformational grammar a structure like:

(3)   O'Hara is more intelligent than he seems to be

would be derived from a more abstract structure schematically like:

(4)   O'Hara is more intelligent than O'Hara seems to be

However, obviously not all pronouns can be so derived, which leads to a differentiation between transformationally introduced pronominal structures and those introduced in the underlying or basic forms, as in:

(5)   he is sick

The fact that pronouns have two different origins can then be suggested as the explanation for the ambiguity of reference of the pronoun in sequences like:

(6)   Schwartz claims he is sick

There is a great deal right in all this and no one who wishes to discuss English pronouns can afford to ignore the insights and observations which underlie the kinds of descriptions just mentioned. It is the thesis of this paper, however, that these analyses ignore some important facts and that there is concomitantly a good deal also wrong in them. Furthermore, what is wrong can be seen to arise from the almost inevitable tendency in grammatical research to assume wrongly that the surface or superficial syntactic forms of sentences provide direct insight into (or are even identical with) their deep syntactic forms.

B   The 'article' character of so-called pronouns

In a transformational grammar, each sentence and hence, derivatively, each part of each sentence has two distinct syntactic structures as part of its overall grammatical description; a highly abstract Deep structure relevant for semantic interpretation and a Surface structure relevant for phonetic interpretation. These two aspects of syntactic form are in general connected by a long and complex chain of transformational rules which, furthermore, derive a sequence of intermediate forms [cf. Chomsky 1965]. In such a grammar it makes no sense to ask such

traditional questions as: 'Is such and such occurrence of form F a noun?' It only makes sense to ask such questions <u>contextually</u> with respect to a specified structure. That is, one can ask whether such and such occurrence of a form F is a noun in the Deep structure, a noun in such and such intermediate structure, a noun in the Surface structure of the sentence, etc. The answer to some of these questions may be yes, to others no without contradiction. Furthermore and equally importantly, the fact that an element is present in the Surface form does not mean it was present in the Deep structure and, conversely, absence from the Surface form does not necessarily entail absence from the Deeper aspect of grammatical structure.

I mention all this only because it is fundamental to my basic claim which is that the so-called pronouns *I, our, they*, etc. are really <u>articles</u>, in fact types of <u>definite</u> article. However, article elements are only introduced as <u>segments</u> in intermediate syntactic structures. In the Deepest structures they are, I shall suggest, not present segmentally but are represented as syntactic features of nouns, features analogous to Animate, Human, Countable, etc. Rather deceptively, the articles which have traditionally been called pronouns are, as a result of certain transformational operations, in many cases assigned a derivative Noun status in Surface structures.

[. . .]

Moreover, further facts strongly suggest that, while it is right to assume that more abstract NP structures of Superficial pronoun-containing NP involve definite articles, it is wrong to assume either that the articles are *the* or that at the relevant stage the pronouns are nouns. Most important in this regard are the reflexive forms such as those in:

(19)   a.   Horace washed himself
        b.   the girl washed herself
        c.   I washed myself

As has been argued by Lees and Klima [1963], it is quite clear that reflexive elements must be derived transformationally from underlying NP which are identical to other preceding NP, this identity being subject to certain conditions. These have never been fully or exactly stated, but they concern occurrence of the two NP within the same <u>simple</u> sentence structure. This may be ignored here. Thus a sentence like (19)a must be derived from a more abstract, Deep structure of the sort schematically indicated: *Horace washed Horace*(subject of course to the remarks of footnote 4 [in the full article]). In previous transformational descriptions, reflexive words such as *myself*, *themselves*, etc. have been treated as compounds of pronouns and a special, transformationally introduced by the very rule which carries out the reflexivization operation as determined by NP identity within simple sentence structures.

This analysis of reflexive forms will not do, however. The identity and simple sentence constraints are fundamentally correct and unquestioned here although they involve some mysterious and far from fully solved problems. But the treatment of the element *self* as a grammatical formative is untenable. In fact *self* must be taken to be a noun stem as we see clearly in such phrases as *the expression of self in our*

society, selfish, selfless, etc. Compare *piggish, brutish, boyish* and *witless, spineless, timeless*, etc. Notice also the *self/selve* plural alternation parallel to that in such unquestioned noun stems as *wife/wive, life/live*, etc. If, however, the stem *self/selve* in reflexive words is a noun stem, what is the preceding element *my, our, him*, etc.? My answer is that they are, of course, articles, definite articles, in fact genitive type definite articles. I view the process of reflexivization as a complex of a number of partially independent operations, some of which are relevant for other grammatical developments such as nonreflexive pronominalization and, most crucially, determination of the Surface forms of so-called pronouns. The relevant rules include PRONOMINALIZATION, DEFINITIZATION, REFLEXIVIZATION, GENITIVIZATION, AND DEFINITE ARTICLE ATTACHMENT.

However, it will be impossible to understand these grammatical operations if it is not recognized that the terminal elements of Deep syntactic structures, i.e. the morphemes, are not unanalyzable atomic symbols. Rather, they are complexes of syntactic, phonological, and semantic features or properties. Phonology and semantics do not concern us here. But the fact that underlying noun stems have a syntactic feature analysis is crucial. The features involved for English must, apparently, include such as Animate, Human, Masculine, First Person (I), Second Person (II), Third Person (III), Definite, Demonstrative, Proper, Pronoun (Pro), Reflexive, Genitive, etc. The claim is then that, instead of nouns coocurring with article morphemes in Deep structures as in previous transformational and other treatments, Superficial structure article differences are represented at the most abstract level by differences in features of nouns, features like Definite, Demonstrative, and, as we see subsequently, also those involving person and gender properties.

[. . .]

Nothing in our analysis thus far accounts for the difference between the terminal two morpheme structure of reflexive words and the single formative character of nonreflexive pronominals. That is, what we have said would suggest that the output NP in Figure 1 [in the full article] should be *heone*. This is not the case here nor is the actual phonological form of the pronoun ever present in analogous forms in the standard language. We can only assume, therefore, the existence of a special rule to drop the nonreflexive pronoun stems in such cases. This is the rule called PRONOUN DELETION [. . .]. Although this seems a bit ad hoc, it in fact provides the basis for an interesting and important justification for the posited analysis which we shall give in the next section. I am definitely claiming, however, that were it not for this highly restricted and low level rule, our so-called pronouns would in fact have the terminal forms *Ione*, *usones*, *herone*, *itone* (or perhaps better *itthing* analogous to the indefinite [*something*]). This should make clear why I said earlier that the contrast pointed out by Allen between pronominals like *he, she, it*, etc., which replace whole NP, and pronouns like *one*, which replace individual nouns, is misleading in part. For in fact I claim that the pronoun which would be pronounced *one, thing*, etc. is also really present in the so-called pronominal cases as well. Further very strong evidence of this will be presented below.

[. . .]

C   Justification for the Analysis of the So-Called Pronouns as Articles

In the previous sections we have outlined an account of forms like *I, us, their*, etc. whereby they are treated as forms of definite article. In our terms this means that they are segments added to NP whose head nouns are [+Definite]. The contrasts among the various definite articles are due to other contrasting features of the head noun. The major motivation of this analysis thus far is the parallelism with respect to properties like Animate, Masculine, I, II, III, etc. between *he/him* and *himself*, *it* and *itself*, *I/me/my* and *myself*, etc. Once it is recognized that the reflexives consist of something plus a noun stem and that this something differs from the forms of pronouns only in case properties (Genitive and Nominative values), it is quite natural to assume that pronominalization and reflexivization involve specifying a noun as [+Pro, +Definite, −Demonstrative] and that these along with the inherent features of the noun then determine the form of the article. Hence by parallelism with *himself* we are led to regard *him* as an article whose underlying head noun (which would otherwise show up phonologically as *one*) has been deleted because it was [+Pro] either inherently or derivatively by identity. While perhaps not completely implausible, thus far we have certainly given little conclusive ground for accepting such an analysis. Basically it has been shown only that it is possible and that it provides a natural way of handling the definiteness of nonderivative pronouns like *I, him, you* and shape parallelisms between these and derivative pronoun forms of the reflexive and nonreflexive varieties. And furthermore the analysis is compatible with the hitherto ignored fact that *self/selve* is a noun stem. More serious evidence in favor of the article analysis is, however, available.

It should be emphasized that the analysis accounts for an otherwise unexplained gap in the NP system with respect to the concurrence of third person pronouns, definite articles, and restrictive relative phrases. One finds real pronouns actually occurring with the definite article *the* if there is a restrictive relative phrase or one of its reduced variants present in the NP:

(23)   a.   I met the one who Lucille divorced
       b.   I met the man who Lucille divorced
(24)   a.   I ate the one Schwartz gave me
       b.   I ate the apple Schwartz gave me
(25)   a.   I bred the small one
       b.   I bred the small lion

but without the restrictives, reduced or not, the pronoun form *one* cannot so occur:

(26)   a.   *I met the one
       b.   I met the man
(27)   a.   *I ate the one
       b.   I ate the apple

(28)   a.   *I bred the one
        b.   I bred the lion

Notice that the analogues with the indefinite article are alright regardless of whether the head noun is [+Pro] or not:

(29)   a.   I met someone
        b.   I met some man
(30)   a.   I ate something
        b.   I ate some apple
(31)   a.   I bred something
        b.   I bred some lion

My suggestion is that the gap left by the definite, nondemonstrative form with [+Pro] head absences in (26) – (28) is actually filled by the so-called pronoun forms, or, more precisely, by that subset which are third person. That is, the so-called third person pronouns, *it, he, her, them*, etc. are exactly the articles assigned to nouns containing the features [+Pro, +Definite, −Demonstrative, +III, −II, −I] in the absence of restrictive relative phrases in the relevant NP. This simultaneously explains the failure of the so-called third person pronouns to occur with restrictive relative phrases or their reductions.

[. . .]

However, in this discussion of underlying features for pronouns we have ignored the question of features like I and II. But these involve some of the most important problems and provide some of the most significant evidence for our analysis. One's initial impression is that, under the assumptions which have been made here, it will be necessary to restrict underlying feature specifications [+I] and [+II] in such a way that they occur only in nouns which are [+Pro] and only in nouns which do not have restrictive relatives. This will be necessary to prevent such impossible elements as *I boy*, *you person*, *you girl who Jack loves*, etc., allowing only abstract *Ione, youone, weones, youones*, which become actual Surface *I, you*, and [*we*]. However, although there are real restrictions here, the just given statement of them is certainly wrong, or rather too general. For it is fundamental to the present analysis that, in the plural, nonthird person elements can occur with both nonpronouns and/or restrictive relative phrases.

The first forms relevant to this claim are those such as we men, you guys, etc. which we take to be cases of [−Pro, +II . .]. Jespersen [1961, Part II: 85], who of course noticed such forms, implied in effect that they were derivatives from appositive relative clauses. In transformational terms this would naturally suggest derivations like, schematically: *we, who are men ⇒ we men*; *you who are children ⇒ you children*. If this solution could be maintained, it would obviate taking *we* and *you* to be articles in such phrases as is insisted here. But in fact this proposal of appositive derivation cannot be right since forms like *we men*, etc. occur in a variety of contexts where appositive relatives may not. Thus, for example, Smith [1964: 48–49] has noted that NP which are the objects in questions may not have appositive relatives:

(33) a. *did you see Bill, who is six feet tall
      b. *who wrote a novel, which was published by McGraw Hill

And as she also observed there are negative contexts which exclude appositive clauses:

(34) a. *he didn't eat the mango, which I bought for him yesterday
      b. *he didn't write a novel, which was banned as obscene

Similarly, other negative contexts exclude appositives:

(35) a. *no American, who was wise, remained in the country
      b. *none of the cars, which were Chevrolets, were any good
      c. *they never insulted the men, who were democrats

But the forms like *you guys* occur in all such appositive-excluding environments:

(36) a. did you see us guys
      b. who insulted you men
      c. he didn't like us Americans
      d. he did not insult you Communists
      e. none of you guys are any good
      f. neither of us professors is quitting
      g. they never agreed with us planners

Furthermore, there are other grounds for doubting the appositive analysis. Notice that the final relative phrase in such pre-article constructions as:

(37) a. that one of the men who is sick
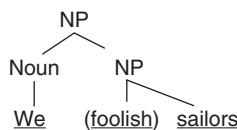
is really associated with the first noun *one*, as shown by the agreement with *sick*. There must therefore be a rule to shift it over the following structure to the end. In nonpronoun NP this following structure can include article, prenominal modifiers, and post-nominal modifiers:

(38) a. that one of the tall men who is sick
      b. that one of the men here who is sick
      c. that one of the men who I like who is sick

Observe, however, that the same relative shift rule must operate in pronoun-containing NP:

(39) a. that one of us who lives here
      b. that one of you guys who betrayed me
      c. that one of you foolish soldiers who deserted his post

Under the analysis suggested here, where *we, us, you*, etc. are articles, the structure over which the relative must shift in (39) is <u>exactly the same</u> as that in (38). But under the appositive analysis the structure would necessarily be radically different, complicating the shift rule, since the derived structure of elements like *we men, you foolish sailors*, etc. would have to be rather like [Figure 4]:

Finally, Jespersen to the contrary notwithstanding, the appositive derivation would assign the wrong interpretation since in fact such phrases do not have appositive meanings, at least not always. This is shown clearly by such examples as:

(40) a. you troops will embark but the other troops will remain
      b. lets us three men leave first

which are certainly not paraphrases of:

(41) a. you, who are troops, will embark but the other troops will remain
      b. *lets us, who are three men, leave first; lets us three, who are men, leave
          first

The fact that (41)b is in addition ungrammatical is further evidence of the inadequacy of an appositive derivation for such forms.

It seems clear then that the only conclusion is that such Surface NP as *we men*, etc. must be derived from underlying nouns which are [−Pro] and yet contain [+I] or [+II] specifications. Hence in such sequences we actually find the so-called pronouns *we/us* and *you* as <u>articles</u> in <u>Surface structures</u>. And this is among the strongest evidence for our overall claim that so-called pronouns have essentially the same type of derivation and status as traditionally recognized definite articles.

Having shown that in the plural first and second person forms can occur with ordinary nouns, we can turn to the question of their occurrence with restrictive relatives. And here also we find a contrast with the situation in the singular. For in fact such phrases as:

(42) a. you men who wish to escape
      b. we Americans who have been struggling here

seem perfectly natural. And this is even more true when the restrictives are reduced:

(43) a. you men here
      b. we honest policemen
      c. you amusing comedians
      d. you diligent Democrats shouldn't put up with lazy ones
      e. Johnas didn't criticize us intelligent workers, only the dumb ones

The occurrence of first and second person forms in the plural with restrictive relatives and their reductions leads to a significant justification for the claim that the so-called pronouns are articles and, in particular, for the claim that for standard English a more abstract set of forms *Ione, heone, weones, themones*, etc. underly [sic] the Surface elements *I, he, we, them*, etc. We illustrate a relevant derivation for the *we* case (on one analysis. I claim that *we* is in general ambiguous).

[. . .]

Most striking is the fact that the hypothetical pronoun stem *one* actually shows up in Surface structures in such forms as:

(44) a.   you great ones
      b.   us quieter ones
      c.   we religious ones

We take these to have structures exactly analogous to those of *you important men, we diligent Democrats*, etc.

[. . .]

Jespersen [1961, Part II: 261–262], who noticed examples like (44) had the following to say:

> '*Ones* may be used after a personal pronoun in the plural. This is not astonishing when an adjective intervenes (as in *you great ones* above . . . or . . . it is very annoying to *us quieter ones*); but it is more difficult to see why *ones* should have been added to a single *we* or *you*. This is found in Scotch dialect . . ., and it is evidently from Scotch that American has taken it. *We'uns* and *you'uns* are especially frequent in the vulgar speech of the Southern states . . .'

Jespersen obviously recognizes the problem which such forms as (44) cause for a view which treats *we, you*, etc. as pronouns. His remark that the occurrence of a following noun is not astonishing when an adjective intervenes is defensive. Why is it not astonishing? But even more, the view falls apart completely when faced with the dialect forms *we'uns, us'uns, you'uns*, etc. The latter provide one of the most crucial justifications for our analysis. For they illustrate a case where the hypothetical forms *weones, youones*, etc. actually are related to pronunciation without the ad hoc rule of nonreflexive pronoun stem deletion which must be posited for the standard language. In comparison to Jespersen's puzzlement, the analysis suggested in this paper provides a natural treatment of such forms. For such dialects as contain *us'uns*, etc. my claim would be that the underlying forms and most of the rules are identical to those suggested here for the standard language. But in these lower class systems the rule which drops nonreflexive pronoun stems after attached definite articles is, at least in first and second person cases, restricted to the <u>singular</u> and does not work for both singular and plural as in the standard language.

[. . .]

Given three features of two values, there are eight possible combinations. And in the plural, in fact, six of these occur:

$$\begin{bmatrix} +\text{III} \\ +\text{II} \\ +\text{I} \end{bmatrix} \quad \begin{bmatrix} +\text{III} \\ +\text{II} \\ -\text{I} \end{bmatrix} \quad \begin{bmatrix} +\text{III} \\ -\text{II} \\ -\text{I} \end{bmatrix} \quad \begin{bmatrix} +\text{III} \\ -\text{II} \\ +\text{I} \end{bmatrix} \quad \begin{bmatrix} -\text{III} \\ -\text{II} \\ -\text{I} \end{bmatrix} \quad \begin{bmatrix} -\text{III} \\ -\text{II} \\ +\text{I} \end{bmatrix} \quad \begin{bmatrix} -\text{III} \\ +\text{II} \\ +\text{I} \end{bmatrix} \quad \begin{bmatrix} -\text{III} \\ +\text{II} \\ -\text{I} \end{bmatrix}$$

$$\overset{}{(1)} \qquad \overset{}{(2)} \qquad \overset{}{(3)} \qquad \overset{}{(4)} \qquad \overset{*}{(5)} \qquad \overset{*}{(6)} \qquad \overset{}{(7)} \qquad \overset{}{(8)}$$

Only the combinations (5) and (6) are impossible in the plural. (1) is, for example, the analysis of the reflexive form in:

(48)   you and I and John can't perjure ourselves

(2) is the analysis of the reflexive in:

(49)   you and John shouldn't bother yourselves about it

(3) is of course the analysis of all so-called third person forms. (7) is the inclusive *we* and (4) is the exclusive *we*. Notice that only the former occurs in the environment after *let's*.

In the singular, on the other hand, only three of the eight possibilities are possible, namely, those in which one of the three features has a plus value and the other two minus values. But since more than four exist in the plural it is clear that two features will not suffice. It should be emphasized that in these analyses I agree very much with Long [1961: 338], who insists that *we* is not the plural of *I* in the same sense in which *boys* is the plural of *boy*. That is, in our terms none of the three possible combinations of features which yields the article *we* differ from the combination which yields the article *I* only in the value of the feature Singular. Features II and III necessarily have different values as well and the feature Pro may also differ since *we* can occur with nonpronouns while *I* cannot.

An important justification for the three feature analysis of person properties is that it provides an important part of the basis for giving a general characterization of the first person-second person interchange in questions and answers. Given feature analyses like those suggested above, the condition is simply that if the values of the features I and II do not agree in any noun form of the question, the 'corresponding' forms in the answer must have the opposite values for each. Thus *did you* (singular) *eat yet* where the underlying subject noun is [−I, +II] must be answered *yes I ate already* where the underlying noun is [+I, −II]. The question *did you* (plural) *leave* must be answered *yes we left* in which the underlying noun is [+I, −II], i.e. the *we* is understood as exclusive. But *we* can also answer questions which contain *we*. *do we have 10 dollars*; *yes we do*. This is possible because the question noun has the specification [+I, +II . . .], i.e. is inclusive, and does not meet the oppositeness condition requiring a switch in the answer. That *we* questions may also take *you* answers follows from the fact that some *we* are [+I, −II], i.e. exclusive. These facts of question-answer first and second person relations are thus good evidence of the ambiguity of English *we* NP. Obviously these question-answer facts discussed here are not really special to English but again involve universal features of language which must ultimately be built into any correct linguistic theory. That this

means features like I, II, III, Pro, etc. must be universals is simply a further confirmation since this seems clear on many other grounds.

There is one final minor argument in favor of the claim that the traditional personal pronouns are actually forms of definite article. Morphophonemically voicing is essentially predictable in dental, nonstrident continuants, i.e. there is no real [ ] – [ð] contrast in English. In particular, voicing may be predicted in such elements in articles, *the, this, that, these, those*, and in so-called pronouns, *they, them, their, theirs* (not too long ago one could of course have added *thee, thy, thine, thou*). But by assuming that pronouns are articles, these two environments are reduced to one. Analysis of generally so-called adverbial elements also suggests that forms like *then, there, thus* actually have the structure <u>definite article + certain types of pronoun</u> so that the same environment covers these as well.

Having mentioned phonology, I can conclude by observing that an analysis like that proposed here for English is to me even more obvious for languages like German and Spanish where, for example, the respective pronoun-definite article similarities between *er-der, sie-die* and *el-el, ella-la* are evidently no accidents. But I leave it for those who know these languages better than I to consider the possibility of such analyses.

1.3   Questions pertaining to Postal (1966)

1   In some languages, for example French, Postal's proposal is supported by the fact that some pronouns (the third person direct object clitic pronouns *le* (m.sg.), *la* (f.sg.), and *les* (pl.)) are identical in form to the definite article. Other third person pronouns are not, such as the subject clitics *il* (m.sg.), *elle* (f.sg.) and their plurals *ils* and *elles*. How might we reconcile this with Postal's proposal?

2   In English, no third person pronoun is identical to the definite article *the*. Yet *the* resembles other elements in English. What are they, and what conclusions could we draw from these resemblances?

3   (Extra credit) All languages seem to have pronouns, but many languages appear to lack definite articles. Give examples of such languages from five different families and discuss the significance of those languages for Postal's proposal.

4   In languages like French, the definite article and third person pronouns are either identical or similar. First and second person pronouns, on the other hand, don't look like the definite article at all. Assume that French is typical, in this way, of languages with definite articles. How should we interpret such a discrepancy between first and second person pronouns and third person pronouns?

5   In French and other Romance languages, as well as in German, Scandinavian languages, and Slavic languages, there is a strong parallelism in form and behavior between first and second person pronouns and some elements that we call reflexives. Pick three of these languages and illustrate this parallelism in as many ways as possible. (Hint: Contrast all of these pronouns with the third person ones.) (Second hint: Take a look at possessives.)

6   (Extra credit) Why should some reflexives be so parallel to first and second person pronouns (cf. question 5)? To what extent could first person vs. second person vs. reflexive be seen as parallel to *here* vs. *there* vs. *where*?

7   Although all languages have pronouns, not all languages have third person
     pronouns. Find an example of such a language. Why might third person pro-
     nouns be special in this way?

8   Are what we call pronouns single morphemes? Find evidence in English that
     they are not. (Hint: Concentrate on the third person pronouns. Bring in
     Corver and van Koppen (2011) and van Koppen (2012).)

9   Are reflexives of the English sort single morphemes? If not, why not? What is
     the significance of nonstandard forms like *hisself*? Bring in Helke (1973); Pica
     and Snyder (1997); Kayne (2002); and Reuland (2011).

10   If pronouns in at least some languages are not single morphemes, how
     does that bear on Postal's proposal that pronouns are akin to definite
     articles?

11   The *which* found in English relative clauses like *the book which is on the table* is
     often called a relative pronoun. From the perspective of Postal's proposal,
     what does this *which* have in common with *he, they*, etc. that could justify the
     use of the term relative pronoun? To what extent could we reasonably use the
     term interrogative pronoun for the *which* of *Which do you like best?*? Same
     question for the *which* of *Which book do you like best?*

12   From Postal's perspective, would it be reasonable to call the *these* of *I would
     prefer these* a demonstrative pronoun? How about in *I would prefer these
     books*?

13   (Extra credit) Discuss the difference that holds in English between *I would
     prefer these* and *I would prefer this*. Is *this* a pronoun in Postal's sense? Extend
     the discussion to three other languages of your choice.

14   Postal supports his idea that pronouns are definite articles by calling attention
     to the similarity between *the linguists* and *we linguists/us linguists*. Yet along-
     side *the linguist* there is no *I/me linguist*. Why would the plural act differently
     here from the singular? To what extent is this a problem for Postal? Bring in
     Delorme and Dougherty (1972).

15   Why does Postal link pronouns to definite articles and not to indefinite arti-
     cles? Are there languages where some pronouns have the same form as the
     indefinite article? (Big hint: Take a look at Romanian.) Are there languages
     where definite and indefinite articles (paradoxically) look alike? (Hint: Look
     around in Scandinavian.)

16   To what extent can we say that the *one* of *They have a blue one* is a pronoun?
     Similarly for *They have blue ones*. (Hint: Look at Spanish *unos*.)

References

Abney, S. 1987. *The English noun phrase in its sentential aspect*. Doctoral dissertation,
     MIT.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT
     Press.

Corver, N., and M. van Koppen. 2011. NP-ellipsis with adjectival remnants: a
     micro-comparative perspective. *Natural Language and Linguistic Theory* 29:
     371–421.

Curme, G. O. 1977 [2 vols, 1931, 1935]. *A Grammar of the English Language*. Essex, CT: Verbatim.

Déchaine, R.-M., and M. Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33: 409–442.

Delorme, E., and R. C. Dougherty. 1972. Appositive NP constructions: we, the men; we men; I, a man. *Foundations of Language* 8: 1–29.

Elbourne, P. 2005. *Situations and Individuals*. Cambridge, MA: MIT Press.

Helke, M. 1973. On reflexives in English. *Linguistics* 106: 5–23.

Jespersen, O. 1961 [1909–1949]. *A Modern English Grammar on Historical Principles (in 7 parts)*. London: Allen & Unwin / Copenhagen: Ejnar Munksgaard.

Kayne, R. S. 2002. Pronouns and their antecedents. In S. D. Epstein and T. D. Seely (eds), *Derivation and Explanation in the Minimalist Program*, Oxford: Blackwell, pp. 133–166. [Reprinted in Kayne (2005, Ch. 6).]

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

van Koppen, M. 2012. The distribution of phi-features in pronouns. *Natural Language and Linguistic Theory* 30: 135–177.

Lees, R. B., and E. S. Klima. 1963. Rules for English pronominalization. *Language* 39: 17–28.

Long, R. B. 1961. *The Sentence and Its Parts*. Chicago: University of Chicago Press.

Longobardi, G. 1994. Reference and proper names: A theory of N-Movement in syntax and logical form. *Linguistic Inquiry* 25: 609–665.

Pica, P., and W. Snyder. 1997. On the syntax and semantics of local anaphors in French and English. In A. M. Di Sciullo (ed.), *Projections and Interface Conditions*. Essays on Modularity, New York and Oxford: Oxford University Press, pp. 235–250.

Reuland, E. J. 2011. *Anaphora and Language Design*. Cambridge, MA: MIT Press.

Roehrs, D. 2009. *Demonstratives and Definite Articles as Nominal Auxiliaries*. Amsterdam and Philadelphia: John Benjamins.

Smith, C. S. 1964. Determiners and relative clauses in a generative grammar of English. *Language* 40: 37–52.

Szabolcsi, A. 1994. The noun phrase. In F. Kiefer and K. É. Kiss (eds), *The Syntactic Structure of Hungarian. Syntax and Semantics 27*, San Diego: Academic Press, pp. 179–274.

2

On Complementizers: Toward a Syntactic Theory of Complement Types

Joan W. Bresnan

1970

2.1 Introduction

In the late 1960s, the syntactic component of the grammar consisted of two parts: phrase structure rules, which determined the deep structure (DS) of a sentence, and the transformational component, which derived the surface structure (SS) from the DS, via a combination of deletion, insertion, and displacement operations. DS was taken to characterize the semantic interpretation of a sentence; from a single DS, transformations could derive distinct SSs representing the different surface forms, such as active and passive, that express the meaning represented by that DS. Since the role of DS was tightly linked to the expression of meaning, semantically vacuous elements were assumed to be absent at DS and inserted, by transformations, during the derivation of SS. Among such vacuous elements were thought to be complementizers (words like *that* and *for*). This meant that, for example, (1a) and (1b) would have the same DS, namely (1c):

(1) a. It may distress John for Mary to see his relatives.
     b. It may distress John that Mary sees his relatives.
     c. [[Mary sees his relatives] may [distress John]]

Bresnan (1970) takes issue with this view, arguing that complementizers must be present at DS.

Bresnan's article presents several pieces of evidence for the presence of complementizers at DS, two of which are as follows. First, if complementizers are inserted by transformational rules, one would expect them to be semantically empty – that

is, they should not affect the meaning of the sentence. Note, however, that (1a) and (1b) are not synonymous. (1b) presupposes that Mary sees John's relatives, while (1a) does not. Even more obvious is the semantic difference between the following sentences (contrasting the complementizers *if* and *that*):

(2)   a.   He doesn't care that she is a doctor.
      b.   He doesn't care if she is a doctor.

Secondly, complementizers seem to be involved in subcategorization, in which verbs select what types of arguments (subjects and objects) they can have. Subcategorization was assumed to be checked during the construction of DS, when lexical items were inserted into the phrase structure. Consider the following:

(3)   a.   That he eats cabbage means nothing.
      b.   For him to eat cabbage means nothing.
      c.   This means that he is of low birth.
      d.   *This means for him to be of low birth.

The verb *mean* can take sentential arguments as both subject and object. However, whereas in subject position they can be introduced by either *that* or *for*, in object position *for*-sentences are not possible. This would seem to be inexplicable under the view that complementizers are inserted, but it is easily explained under Bresnan's view that they are part of the DS of the sentence.

Bresnan's article proposes the presence of a COMP position that combines with a sentence to form an S′ constituent. It suggests that COMP hosts not only complementizers like *that* and *if*, but also *wh*-elements that are positioned in the front of a clause via a transformation, an idea that has been pursued in much other work (Chomsky 1973, 1977, 2004; Reinhart 1981; Szabolcsi 1994). As Bresnan notes, this proposal straightforwardly captures the complementary distribution of fronted *wh*-constituents and complementizers in English (cf. the "Doubly Filled COMP Filter" of Chomsky and Lasnik 1977).

Bresnan's idea that complementizers form an integral part of underlying clausal structure has been integrated into contemporary views. Later work has, however, questioned whether the position of the complementizer should be identified with that of fronted elements. For example, den Besten (1983) argues that the position of the verb in verb-second clauses in Dutch and German is identical to that of the complementizer, but is also preceded by a fronted constituent, suggesting that these elements occupy distinct structural positions. The integration of COMP into the X-bar template as the CP projection, suggested in Chomsky (1986), provides one way of reconciling this: the complementizer (and fronted verb) occurs in the head position, while fronted phrasal constituents occupy the specifier of CP. A line of work following Rizzi (1997) expands the articulation of the COMP position still further in distinct layers of projection associated with Topic, Focus, Force, and Finiteness, known as *the left periphery*.

2.2   From "ON COMPLEMENTIZERS: TOWARD A SYNTACTIC THEORY OF COMPLEMENT TYPES"
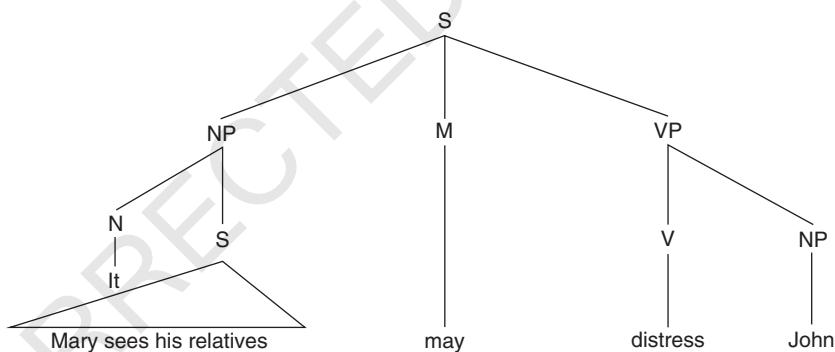
0   Introduction

In virtually all analyses of complementation within the framework of generative grammar the following sentences would differ only by optional transformations in derivations from a common deep structure:

(1)   It may distress John for Mary to see his relatives
(2)   It may distress John that Mary sees his relatives
(3)   Mary's seeing his relatives may distress John

The particles *that, for-to, 's-ing* – the so-called 'complementizers' – have been viewed as markers of syntactic subordination having neither semantic content nor significant syntactic function. Accordingly, examples (1)–(3) have usually been assigned a common underlying structure roughly similar to that represented in (4):

(4)

In the derivation of (1)–(3) from (4) one or more transformations have been held to insert a complementizer into the embedded sentence. This theory of complement types I will call the 'transformational hypothesis'. The transformational hypothesis entails that sentential complement types (*that-*, *for-to-*, and *'s-ing-*clauses) are not distinguishable in deep structure.

   Since it is clear that not all verbs taking complements can occur with every complementizer –
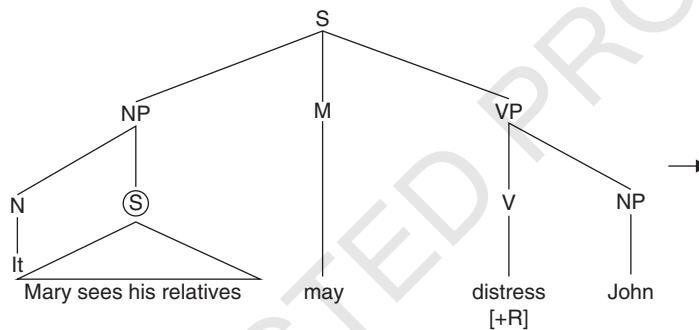
(5)   They decided that their children were happy
        *They managed that their children were happy
        *They decided for their children to be happy
         They managed for their children to be happy
        *They decided their children's being happy
        *They managed their children's being happy

– it is evident that some characteristic of the higher verb or predicate affects the choice of complementizers. Given the transformational hypothesis, one possible assumption is that a 'rule feature' (Lakoff, 1965) is associated with verbs and predicates; thus if the transformation introducing *for-to* is called *R*, then (judging from (1) and (5)) the verbs *distress, decide, manage* would be marked in the lexicon (either directly or by a redundancy rule) as [+R], [−R], [+R], respectively.
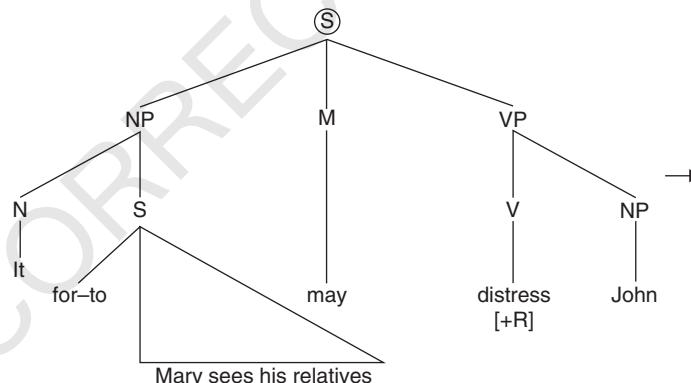
Because the complementizer-insertion transformations have to be sensitive to the rule feature on the higher verb, there is a peculiarity in their operation: they cannot insert complementizers into a sentence S during the transformational cycle on S, but only during the cycle on the next sentence dominating S.

To illustrate this feature of the transformational hypothesis, diagram (6) shows in a simplified way several steps in the derivation of (2); the S being cycled (see Chomsky (1965) on the notion of transformational cycle) is indicated by a circle:

(6)
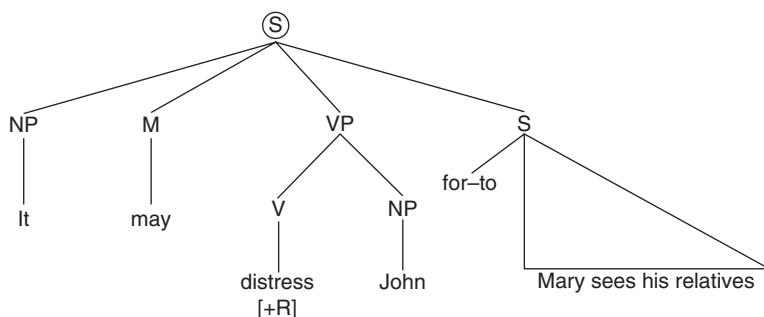
On the first cycle (6a) no transformations apply. If a transformation, say the passive, were to apply on this cycle, then ultimately the sentence *It may distress John for his parents to be seen by Mary* would result. On the second cycle (6b–c) the transformation inserting *for-to* (rule R) applies, because *distress* is marked [+R]. Finally, another transformation shifts the *for-to* complement to the end. The key point is that complementizer insertion could not occur on the first cycle because the transformation would not 'know' which complementizers are permitted by the verb *distress* until the next cycle. In other words, the structural description of any complementizer-insertion transformation cannot be limited to a complement clause, but must include the verb or adjective which that clause complements.

This peculiarity guarantees that nonembedded sentences will never appear with complementizers –

(7)   *That they imagined it

– but at the same time it violates an otherwise well-motivated universal stated by Chomsky (1965: 146), namely, that while transformations may remove material from embedded sentences, no transformation can insert morphological material into 'lower' sentences.
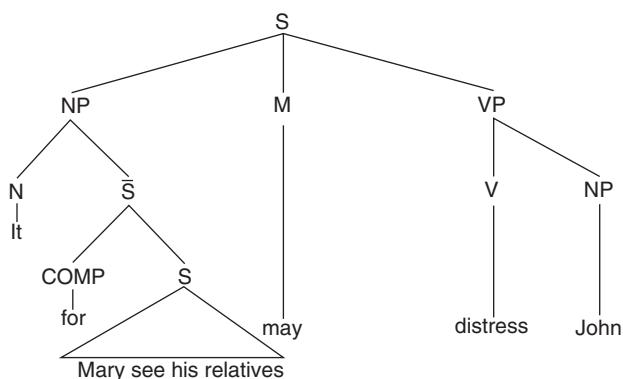
These facts and many others which are difficulties for the transformational hypothesis will be seen to follow naturally from an alternative theory – 'the phrase-structure hypothesis' – which I shall justify in this study. According to the phrase-structure hypothesis complementizers are specified in deep structure. Specifically, there exists a phrase-structure rule in English (and in every language having complementizers) which permits complement types to be distinguished in deep structure. The rule for English would look like

(8) $\bar{S} \rightarrow COMP\ S$

where the symbol COMP represents a deep structure node (or feature bundle (Chomsky, 1967)) which dominates (or is featurally specified for) complementizers. As I will show, the underlying sentential complementizers of English are *that*, *WH* (='Q'), and *for*. There is evidence from syntax, semantics, and universal grammar that complementizers are far from the semantically empty, syntactically trivial particles they have been assumed to be in most previous generative work.

As an illustration of the phrase structure hypothesis, note that (1)–(3) would have different underlying structures: (2), for example, would be represented as

(9)

Notice that the rule-feature [+R] is unnecessary: since complementizers are specified in deep structure, verbs may be *subcategorized* for the type of complement they take.

[…]

1 'COMP' as a node in deep structure

In this Section I give two basic arguments justifying the phrase structure rule $\bar{S}$ → COMP S, that is, establishing that COMP is a node in deep structure. First is the argument from subcategorization: complementizer selection is the *kind* of phenomenon characteristic of subcategorization, and in fact some aspects of complementizer-choice cannot even be described within the transformational hypothesis without extending it in undesirable ways. (The extensions necessary are undesirable because they mirror the subcategorization solution without restricting the theory of grammar correspondingly.) Second is an argument based on the interaction of a putative complementizer-insertion transformation with other transformations. Since it is known that verbs must be subcategorized for *WH* (or 'Q'),

(18) We $\left\{\begin{array}{c} \text{*believed} \\ \text{inquired} \end{array}\right\}$ whether he was there

(19) We $\left\{\begin{array}{c} \text{believed} \\ \text{*inquired} \end{array}\right\}$ that he was there

the first argument of this section will gain force when it is seen in the following section that *WH* is a complementizer; however, the arguments here are entirely independent of those in the next section.

To see that complementizers subcategorize verbs, consider such verbs as *mean, show, imply, reveal, entail, suggest, prove*, which take multiple sentential complements:

(20) That he eats cabbage means nothing
(21) This means that he is of low birth
(22) That he eats cabbage means that he is of low birth

The *that*-complement can occur as both subject and object, but the *for*-complement cannot:

(23) For him to eat cabbage means nothing
      = It means nothing for him to eat cabbage
(24) *This means for him to eat cabbage

If COMP were not in deep structure, these facts could not be described by known transformational means, including rule features; for if there were a complementizer-insertion transformation operant in (20)–(24), it would have to be sensitive not only to the rule feature in the verb, but to the subcategorization of the verb – *that* goes on subject or object, *for* on subject only.

There is another problem: sentences like the following are ungrammatical:

(25)   *For him to eat cabbage means that he will be sick

Apparently, then, there are interdependencies between complements which affect the choice of complementizer: when *that* occurs in the object, *for* cannot occur in the subject. Thus if COMP were not a node in deep structure, verbs could not be subcategorized for complementizers and the transformational machinery would have to be extended further in some *ad hoc* way. On the other hand, it is characteristic of subcategorization to display such interdependencies; for example, certain verbs prohibit noun phrase objects when they have sentential subjects:

(26)   That John eats cabbage implies that he likes cabbage
(27)   The first statement implies the second statement
(28)   The first statement implies that the second statement is true
(29)   *That the first statement is true implies the second statement

Further evidence that verbs are more naturally subcategorized for complementizers than marked with rule features for a complementizer–insertion transformation appears in the case of *tell*. For *tell* the choice of complementizer depends upon the presence of another object of the verb. Consider, for example,

(30)   Susie didn't tell $\begin{cases} \text{*that they had eaten} \\ \text{whether they had eaten} \end{cases}$

(31)   Susie didn't tell us $\begin{cases} \text{that they had eaten} \\ \text{whether they had eaten} \end{cases}$

These examples suggest that *that* is permitted with *tell* only if the indirect object is specified in deep structure. To be more explicit: there is evidence that *tell* always has objects in deep structure, for totally intransitive verbs may occur where *tell* cannot, as in *the sleeping man* vs. **the telling man*. This evidence presents at least two possibilities for explaining (30) and (31). The first possibility (Chomsky, 1964) is that the relevant object of *tell* can be deleted when it is a 'designated element' such as *someone*. In this case, the deletion rule might be made sensitive to the type of complement in a way that could describe (30) and (31) without recourse to subcategorization. But the presence of the additional morphological material in *tell someone* vs. *tell* has semantic consequences which argue against this alternative: note the difference between

(32)   I always like to tell stories – but not necessarily tò anyone. (Often
        I just spin a yarn for my own benefit.)

and

(33)   *I always like to tell stories to someone – but not necessarily tó anyone

– which is overtly contradictory.

The second possibility (Chomsky, 1964) is that the indirect object of *tell* may occasionally be a dummy element – that is, may occasionally fail to be filled by lexical insertion rules. This alternative, which seems to be semantically preferable, is the one which I am assuming in my discussion of (30) and (31). Since the option of having direct or indirect dummy objects is restricted to certain lexical items only (for example, *she ate*, but not *it frightened), it is a matter of subcategorization; it follows that the choice of *that* with *tell* depends upon subcategorization.

[…]

2   *WH* as a complementizer

The arguments of the first section support the phrase structure hypothesis, that there exists a phrase structure rule in English which permits complement types to be specified in deep structure. The arguments of this section, showing that *WH* is a complementizer, permit a more exact specification of English complement types.

(i)

Like recognized complementizers, *WH* can appear in multiple sentential complements to verbs:

(49)   How he acts early in the morning will show what he's really like

Both *how* and *what* in (49) reflect the presence of underlying *WH* complementizers. And like recognized complementizers, *WH* shows subcategorizational dependencies:

(50)   Whether or not his mouth turns black will show whether or not he's been nipping at the silver nitrate
(51)   *Whether or not his mouth turns black will show that he's been nipping at the silver nitrate
(52)   *That his mouth is turning black will show whether or not he's been nipping at the silver nitrate

Again, like recognized complementizers, *WH* appears in coordinate structures –

(53)   He didn't say whether Kíp flew to New York or whether Màry did

– which are optionally subject to Conjunction (or rather Disjunction) Reduction for some speakers:

(54)   Whether Kíp or whether Màry did it doesn't matter

These facts do not, of course, decide between subcategorizing verbs for complementizers (including *WH*) and permitting embedded 'Q' (as a non-complementizer)

to subcategorize verbs together with recognized complementizers. Though this decision indeed appears to be trivial, or merely notational, it nevertheless has significant consequences.

[…]

(iii)

In English complementizers maintain clause-initial position: this is true in all types of complement constructions and for all types of complementizers, so that if *WH* is a complementizer, the following paradigms are to be expected:

(65)   OBJECT COMPLEMENTATION
        I know that he is wise
        I prefer for you to speak English
        I am asking whether you will accompany me
(66)   SUBJECT COMPLEMENTATION
        That he was alone was obvious from the report
        For you to leave right now would be inconvenient
        Whether he eats cabbage or not simply doesn't matter
(67)   COMPLEX NP COMPLEMENTATION
        The idea that nobody will survive is appalling
        The command for all troops to move out was given Friday
        The question whether they'll strike remains unanswered
(68)   COPULAR COMPLEMENTATION
        Your problem is that you are arrogant
        The command was for all troops to move out
        The main question is whether they will support us
(69)   ADJOINT COMPLEMENTATION
        For his son to enjoy the army, he would have to try very hard
        Whether or not his son enjoys the army, he will try very hard
        That his son would not have to join the army, he joined himself

Further, these complementizers are often deletable: the important exceptions to note here are the cases of subject complementation:

(70)   *He was alone was obvious from the report
        *You to leave right now would be inconvenient
        *Does he eat cabbage or not doesn't matter
        *What does he eat doesn't matter

In 'highest' or nonembedded sentences, both *WH* and *that* are obligatorily deleted:

(71)   *That John is here
(72)   *Whether is John here

[…]

In summary, *WH* is like recognized English complementizers in being clause-initial, nondeletable in subject position, obligatorily deletable in non-embedded sentences, and in permitting preposed prepositions where it is deletable.

[…]

(v)

A final argument that *WH* is a complementizer (or rather, pleasant consequence of assuming that it is) is provided not by particular syntactic or semantic facts, but by the form of a linguistic universal: I will show that if *WH* is recognized as a complementizer, Baker's 'Q-universal' (Baker, 1970) can be strengthened in a way which immediately explains an unsolved problem noted by Baker – that of relating the universal behavior of relative clauses to that of questions. This result in turn leads to another possible universal and a means of relating underlying word order to the class of transformations possible in a language.

Baker's Q-universal is stated as follows (Baker, 1970):

> The first part of the hypothesis is that morphemes such as *if* and *whether*, and other words and particles in other languages in which such elements occur, are introduced into trees as lexical realizations of the Q morpheme … The second part of the hypothesis is that there is only one possible movement rule for questions, which differs in different languages only in the particular formatives mentioned in place of English *wh* [footnote 12 omitted]:

> Q X NP Y
> $1\ 2\ 3\ 4 \rightarrow 1, 3+2, \varnothing, 4$
> Condition: 3 [dominates] *wh*.

Note that this statement is restricted to questions (embedded or not) and that it requires the existence of a special Q morpheme which is also a node dominating lexical elements.

If my hypothesis that *WH* is a complementizer is correct then there is no such special node Q and the statement above must be reformulated in terms of the node COMP:

The Complementizer Substitution Universal
Only languages with clause-initial COMP permit a COMP-substitution transformation.

(Until a more thorough study is completed, the term *COMP-substitution transformation* may be understood informally to apply to any transformation moving a constituent over an essential variable into the position of the complementizer – for example, Relative-Clause Formation and Question Formation, or *WH-*movement. There may be more than one such transformation in a given language.) This statement is not restricted to questions, but applies to any clause containing a complementizer which 'attracts' elements of its clause over a variable.

One such complementizer is the *that* appearing in relatives: Klima (1964) and Emonds (1969) give evidence that this *that* is a complementizer, the same that

appears elsewhere; evidence that it is not a pronoun is provided by the observation that relative pronouns may be the objects of prepositions (*to whom, to which*), but *that* never may: **the man to that I spoke*. Moreover, it is well known that in many languages, relative clauses are introduced by a morpheme having the same shape as the declarative complementizer. Further, Emonds (1969) argues that some relative clauses are introduced by a *for* complementizer; thus the following sentences may be related by relativization rules:
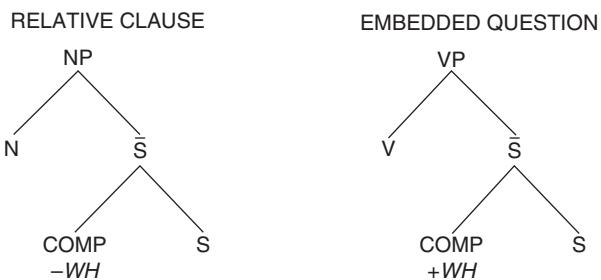
(94)   The weapons for us to practice with are on the table
(95)   The weapons with which to practice are on the table

A more detailed discussion of Complementizer-Substitution transformations, *WH-* and relative clauses, may be found in Bresnan (1972).

    If relative clauses are derived from complementized clauses, then the Complementizer-Substitution Universal immediately provides a solution to this problem noted by Baker:

> … there is one interesting respect in which future research may show that this hypothesis [quoted above], rather than being too general, is not general enough. This is that such phenomena as the moving of relativized constituents in relative-clause formation may eventually be shown to be governed by essentially the same restrictions. In particular, in the languages for which I have relevant information, the movement of a relativized constituent occurs only when the modified noun is to the left of the relative clause, and then only in a leftward direction. If this can be shown to hold true as a general rule, then it will clearly be insufficient to state two remarkably similar 'universal' transformations without providing some explanation for their similarity.

    The point to be emphasized here is that by recognizing Q to be a complementizer, two problems of universal grammar are solvable at once. It is totally obscure why relative clauses with heads to the left should be universally similar to clauses with question particles to the left (there is absolutely no evidence that relatives have the question morpheme: **the boy whether I saw is here*); yet it is clear how clauses having leftward positioned complementizers are structurally alike. The necessary assumption is that relative clauses with leftward heads are derived from clauses with leftward complementizers:

(I ignore the question of the exact labelling of nodes.) In the nonembedded case there is also a structural similarity:

DECLARATIVE   INTERROGATIVE

$\bar{S}$ $\bar{S}$

COMP   S   COMP   S
                          +*WH*

These structural similarities strongly suggest that it is the Phrase Structure rules which determine position of complementizers in a given language. Thus, it is no accident that English fails to have clause-initial *that* together with clause-final *whether*: **They knew he came whether*, or clause-final *how*: **They knew he did it how*.

To illustrate the Complementizer-Substitution Universal, consider an example from Japanese. (All facts about Japanese were related to me by Kinsuke Hasegawa.) This language has relative clause heads to the right and has no special rule for moving questioned constituents to initial position; thus in Japanese it is possible to ask questions unformulable in English:

(*anata wa*) *dare ni nani o ageta hito ni atta no desu ka*
　　(you)　　who　to what　gave man　　met
= 'You met a man who gave what to whom?'

(In English this sentence is possible only with 'echo' intonation, but in Japanese it is a typical and normal interrogative form.) The Complementizer-Substitution Universal predicts that Japanese, a language having rightward complementizing, should have no Question Formation or Relative Clause Formation *movement* rules analogous to those in English.

Above, I assumed that in languages having relative clause heads to the left these clauses are derived from clauses with leftward complementizers. It is natural to ask what justification there is for this assumption. If it is the Phrase Structure rules which determine complementizer position, then this assumption amounts to the claim that languages with leftward heads of relatives have *all* complementizers positioned to the left; that is, if a language contains

R_1 NP $\rightarrow$ N S,

then it contains

R_2 $\bar{S}$ $\rightarrow$ COMP S.

In particular, this claim implies that question particles will be to the left. On the basis of the data in Greenberg (1962: universals 3, 4, 9, 12, 24, and Appendix I), apparently it is generally true that languages which have R_1 are just those which have clause-initial question particles, or R_2. Here, then, is another possible universal:

The Expansion Universal
If R_1 belongs to the PS of the grammar of a language, then so does R_2.

There may well be a deeper explanation for this fact in terms of universal constraints on Phrase Structure expansion, and the above formulation is highly tentative.

Given the Expansion Universal, notice that the Complementizer-Substitution Universal may be restated as follows:

The Complementizer Substitution Universal
If R_2 belongs to the PS of the grammar of a language, then any member of T_i may belong to that grammar,

where T_i refers to the formally specifiable class of Complementizer-Substitution Transformations. Thus we have an obvious connection between underlying word order and the class of transformational rules possible in a language.

2.3   Questions pertaining to Bresnan (1970)

1   In some cases, English can omit complementizer *that*, e.g., in *Everybody knows (that) you're right*. But this is not always possible, for example in *That your friends are right is obvious*. Find as many other cases as possible in which *that* cannot be omitted.

2   The impossibility of **Your friends are right is obvious* has sometimes been thought to be related to the fact that *are* in this example could be misconstrued as the main verb of the sentence. Find an example in English (or in another language) that shows that that is not a sufficiently general solution.

3   English complementizers care about the form of the main verb in the embedded sentence that they introduce, in the sense that *that* is incompatible with nonfinite embedding (e.g., **We would like that (to) leave right away*) and *for* is incompatible with finite embedding (e.g., **We would like very much for John would leave right away*). To what extent is it satisfactory to analyze these restrictions as 'selectional'? Give your reasons.

4   Nonfinite relatives in English show the following restriction: *He needs a shelf on which to put his new books* vs. **He needs a shelf which to put his new books on*. If this restriction turned out to be the same as or similar to the one mentioned in the previous question concerning *that*, what might (or might not) that tell us about the status of *that*? Bring in Kayne (to appear).

5   English relative clauses with *that* sometimes allow *that* to be dropped, as in *the person (that) we were talking to*. Find three languages from different families in which relative clauses can be introduced by a complementizer and find out if that complementizer can be dropped in those languages, in the English fashion. Why might (or might not) the dropping of the complementizer be related to the fact that many languages (e.g., Chinese) arguably lack complementizers in relative clauses completely? Bring in the word order considerations discussed by Kayne (1994, section 9.3).

6   English complementizer *that* is identical in form to demonstrative *that*. Find another language closely related to English that has the same property.

7   *That* in English relative clauses is frequently called a complementizer. But some linguists would say that *that* is really a relative pronoun, in the same sense in which we say that *who* in *the man who we were talking to* is a relative pronoun. Find as many ways as possible in which relative *that* and relative *who* differ in their behavior.

8   Bresnan does not broach the question why English has complementizer *that*, but no complementizer *this*. To what extent is that a legitimate question? To what extent do you think Kayne's (to appear) proposal for distinguishing *that* from *this* is on the right track?

9   Should *to* in English be grouped with *that* and *for* as a complementizer? Give your reasons. (Bring in Saito 2010.)

10 In German embedded sentences in which the verb is in second position, there cannot be a complementizer present. Why might that be? (Hint: Look at Leu 2012.) In what way might English sentences such as *He insists that at no time has he told us anything but the truth* be relevant?

11 Using as many other languages as possible, test the hypothesis that the English vs. German contrast alluded to in the previous question is related to the VO vs. OV difference between English and German.

12 Find (at least) one consistent difference between the complementizers of English, German, and Dutch, on the one hand, and Romance languages, on the other.

13 There are languages in which the complementizer introducing embedded declarative clauses is identical or very similar in form to a verb. How seriously should we take that fact in our study of synchronic syntax? Bring into the discussion Koopman and Sportiche (1989).

14 Is it justified to ask why human languages have complementizers in the first place? Give your reasons.

15 Would it be justified to postulate the presence of unpronounced complementizers in a language which seemed not to have any? Why or why not?

16 To what extent is Bresnan's view of complementizers compatible with Rizzi's (1997)? To what extent is Rizzi (1997) important for Chomsky's (2008) view of the relation between C and T?

17 English gerunds differ from English infinitives in disallowing complementizer *for* (e.g., *John will start (*for) working hard tomorrow*). English gerunds also disallow an initial *wh*-phrase, as seen in **John doesn't remember where having put the book*, again as opposed to infinitives (e.g., *We have no idea where to put our books*). Why might (or might not) these two restrictions on gerunds be related? What is the relevance of English allowing *What book will Mary start reading tomorrow?*

18 English small clauses disallow complementizers, as seen in *They would like you (to be) here by 8 a.m.* vs. **They would like for you here by 8 a.m.* As with gerunds (see the previous question), small clauses are also incompatible with an initial *wh*-phrase (e.g., *You don't know what *(to be) satisfied with*). Are these two restrictions concerning small clauses related to each other? Are they related to the apparently similar restrictions on gerunds? Spell out your reasons.

19 Rosenbaum (1967) argued that in sentences such as *It's obvious that we should leave soon, it* and *that we…* start out as one constituent (that is split apart by extraposition). Discuss the possible importance of Rosenbaum's idea for Bresnan's analysis of complementizers.

20 Kayne (2003, section 4.6) proposed that for an IP to function as the argument of a higher predicate, it must be nominalized. How does this proposal bear on the status of complementizers, depending on whether or not Rosenbaum's proposal of the previous question is correct? How do English infinitives bear on Kayne's proposal?

References

Baker, C. L. 1970. Notes on the description of English questions: The role of an abstract question morpheme. *Foundations of Language* 6: 197–219.

den Besten, H. 1983. On the interaction of root transformations and lexical deletive rules. In W. Abraham (ed.), *On the Formal Syntax of the Westgermania. Linguistik Aktuell/Linguistics Today 3*, Amsterdam and Philadelphia: John Benjamins, pp. 47–131.

Bresnan, J. 1972. *The theory of complementation in English syntax*. Doctoral dissertation, MIT. [Published 1979 by Garland Publishers, New York.].

Chomsky, N. 1964. *Current Issues in Linguistic Theory*. The Hague: Mouton.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. 1970. Remarks on nominalization. In R. Jacobs and P. Rosenbaum (eds), *Readings in English Transformational Grammar*, Waltham, MA: Ginn, pp. 184–221.

Chomsky, N. 1973. Conditions on transformations. In S. Anderson and P. Kiparsky (eds), *A Festschrift for Morris Halle*, New York: Holt, Rinehart, and Winston, pp. 232–286.

Chomsky, N. 1977. On Wh-movement. In P. Culicover, T. Wasow, and A. Akmajian (eds), *Formal Syntax*, New York: Academic Press, pp. 71–132.

Chomsky, N. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 2004. Beyond explanatory adequacy. In A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures 3*, New York and Oxford: Oxford University Press, pp. 104–131.

Chomsky, N. 2008. On phases. In R. Freidin, C. P. Otero, and M. L. Zubizarreta (eds), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, Cambridge, MA: MIT Press, pp. 134–166.

Chomsky, N., and H. Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8: 425–504.

Emonds, J. E. 1970. *Root and structure-preserving transformations*. Doctoral dissertation, MIT.

Greenberg, J. H. 1962. Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (ed.), *Universals of Language*, Cambridge, MA: MIT Press, pp. 73–113. [2nd edition, 1966.]

Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. S. 2003. Antisymmetry and Japanese. *English Linguistics* 20: 1–40. [Reprinted in Kayne (2005, Ch. 9).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Kayne, R. S. to appear. Why isn't *this* a complementizer? In P. Svenonius (ed.), *Functional Structure from Top to Toe: A Festschrift for Tarald Taraldsen*, New York and Oxford: Oxford University Press. [Reprinted in Kayne (2010b, Ch. 10).]

Klima, E. 1964. Relatedness between grammatical systems. *Language* 40: 1–20.

Koopman, H., and D. Sportiche. 1989. Pronouns, logical variables, and logophoricity in Abe. *Linguistic Inquiry* 20: 555–588.

Lakoff, G. 1965. On the Nature of Syntactic Irregularity. Report NSF-16, The Computation Laboratory, Harvard University.

Leu, T. 2012. The German complementizers 'd-ass': Two Cs in one word. Ms., Université du Québec à Montréal.

Reinhart, T. 1981. A second COMP position. In A. Belletti, L. Brandi, and L. Rizzi (eds), *Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW conference*, Pisa: Scuola Normale Superiore, pp. 517–557.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 281–337.

Rosenbaum, P. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: MIT Press.

Saito, M. 2010. On the nature of the complementizer *To*. *Journal of Japanese Linguistics* 26 (Special Issue in Memory of S.-Y. Kuroda): 85–100.

Szabolcsi, A. 1994. The noun phrase. In F. Kiefer and K. É. Kiss (eds), *The Syntactic Structure of Hungarian. Syntax and Semantics 27*, San Diego: Academic Press, pp. 179–274.

3

Remarks on Nominalization

Noam Chomsky

1970

3.1   Introduction

A fundamental question in syntax concerns its role in word formation. Consider the triplet in (1), consisting of a sentence (1a) and two noun phrases, a gerundive nominal (1b) and a derived nominal (1c):

(1) a.   John has refused the offer.
     b.   John's refusing the offer
     c.   John's refusal of the offer

(1a,b,c) clearly have a number of things in common: they share core aspects of their meaning; the verb *refuse* and the nouns *refusing* and *refusal*, which have a common morphological root, function as predicates and impose the same selectional restrictions on their arguments (for example, *John* is an appropriate subject for all three, but an inanimate NP like *the rock* would be inappropriate for all three); they can all take either a noun phrase (e.g., *the offer*) or a sentence as their complement (e.g., *to leave his position*). The grammar, that is to say, the abstract model that we build to capture the native speaker's knowledge, should be able to express these similarities. In Lees's (1960) work on nominalizations, this is done by saying that they all originate from a single entry in the lexicon, in this case the verb *refuse*: both *refusing* and *refusal* are seen to arise from the application of a syntactic operation of nominalization, which takes a sentence such as (1a) as input and yields a corresponding nominal such as (1b) or (1c) as output.

However, there are also contrasts between sentences and their nominalization counterparts. In *Remarks*, Chomsky argues that an approach that derives (1a,b,c)

all from the same underlying Deep Structure representation cannot easily explain the differences that exist between them, in particular the contrast between (1c), on the one hand, and (1a,b) on the other. In the case of gerundive nominals, the correspondence with the associated clause is quite regular, both in the sense that it can apply to any verb and that the verb and the corresponding gerundive have the same meaning. This is exactly what is expected if gerundive nominals (1b) are derived from the same Deep Structure as the corresponding sentence (1a), as the result of syntactic transformations, since such operations are assumed to apply in a regular and invariant way.

In contrast, there is no systematic relationship between what are called derived nominals (1c) and associated verbal (or adjectival) forms: not every verb has a derived nominal, and not every derived nominal has a corresponding verb. Moreover, the semantic relation between the two can be quite idiosyncratic: for example, *ignore* and *ignorance*, while related, have very different meanings. This would be unexpected if derived nominals were derived from the same Deep Structure as their associated verbal (or adjectival) forms.

Hence, this paper argues that gerundive nominals are indeed derived by syntactic transformations from the same Deep Structure as the corresponding sentence, but derived nominals are not. Instead, there are two separate Deep Structures, one for the verbal form (*refuse*) and another for the nominal (*refusal*). Because both structures are derived from a single X-bar template, the roughly parallel structure of sentences and nominals can be captured. As for the commonalities between (1a,b) on the one hand, and (1c), on the other, Chomsky suggests that they can be captured with a category-neutral lexical entry that expresses their common selectional, subcategorization, and semantic features. To capture their category-specific idiosyncrasies, such a category-neutral lexical item includes conditions on lexical insertion that further determine the semantic and morphological properties of that item when used as a verb or as a noun. When this lexical item is inserted into a syntactic template, as determined by X-bar theory, the categorial features are determined and the idiosyncratic properties arise.

This approach is usually referred to as the lexicalist (as opposed to the transformationalist) hypothesis. It accounts straightforwardly for the potential lack of productivity and for the presence of idiosyncratic differences in meaning between sentences and derived nominals.

3.2   From "REMARKS ON NOMINALIZATION"

For the purposes of this paper, I will assume without question a certain framework of principles and will explore some of the problems that arise when they are applied in the study of a central area of the syntax of English, and, presumably, any human language.

A person who has learned a language has acquired a system of rules that relate sound and meaning in a certain specific way. He has, in other words, acquired a certain competence that he puts to use in producing and understanding speech.

The central task of descriptive linguistics is to construct grammars of specific languages, each of which seeks to characterize in a precise way the competence that has been acquired by a speaker of this language. The theory of grammar attempts to discover the formal conditions that must be satisfied by a system of rules that qualifies as the grammar of a human language, the principles that govern the empirical interpretation of such a system, and the factors that determine the selection of a system of the appropriate form on the basis of the data available to the language learner. Such a "universal grammar" (to modify slightly a traditional usage) prescribes a schema that defines implicitly the infinite class of "attainable grammars"; it formulates principles that determine how each such system relates sound and meaning; it provides a procedure of evaluation for grammars of the appropriate form. Abstractly, and under a radical but quite useful idealization, we may then think of language-learning as the process of selecting a grammar of the appropriate form that relates sound and meaning in a way consistent with the available data and that is valued as highly, in terms of the evaluation measure, as any grammar meeting these empirical conditions.

I will assume that a grammar contains a base consisting of a categorial component (which I will assume to be a context-free grammar) and a lexicon. The lexicon consists of lexical entries, each of which is a system of specified features. The nonterminal vocabulary of the context-free grammar is drawn from a universal and rather limited vocabulary, some aspects of which will be considered below. The context-free grammar generates phrase-markers, with a dummy symbol as one of the terminal elements. A general principle of lexical insertion permits lexical entries to replace the dummy symbol in ways determined by their feature content. The formal object constructed in this way is a *deep structure*. The grammar contains a system of transformations, each of which maps phrase-markers into phrase-markers. Application of a sequence of transformations to a deep structure, in accordance with certain universal conditions and certain particular constraints of the grammar in question, determines ultimately a phrase-marker which we call a *surface structure*. The base and the transformational rules constitute the syntax. The grammar contains phonological rules that assign to each surface structure a phonetic representation in a universal phonetic alphabet. Furthermore, it contains semantic rules that assign to each paired deep and surface structure generated by the syntax a semantic interpretation, presumably, in a universal semantics, concerning which little is known in any detail. I will assume, furthermore, that grammatical relations are defined in a general way in terms of configurations within phrase-markers and that semantic interpretation involves only those grammatical relations specified in deep structures (although it may also involve certain properties of surface structures). I will be concerned here with problems of syntax exclusively; it is clear, however, that phonetic and semantic considerations provide empirical conditions of adequacy that must be met by the syntactic rules.

As anyone who has studied grammatical structures in detail is well aware, a grammar is a tightly organized system; a modification of one part generally involves widespread modifications of other facets. I will make various tacit

assumptions about the grammar of English, holding certain parts constant and dealing with questions that arise with regard to properties of other parts of the grammar.

In general, it is to be expected that enrichment of one component of the grammar will permit simplification in other parts. Thus certain descriptive problems can be handled by enriching the lexicon and simplifying the categorial component of the base, or conversely; or by simplifying the base at the cost of greater complexity of transformations, or conversely. The proper balance between various components of the grammar is entirely an empirical issue. We have no a priori insight into the "trading relation" between the various parts. There are no general considerations that settle this matter. In particular, it is senseless to look to the evaluation procedure for the correct answer. Rather, the evaluation procedure must itself be selected on empirical grounds so as to provide whatever answer it is that is correct. It would be pure dogmatism to maintain, without empirical evidence, that the categorial component, or the lexicon, or the transformational component must be narrowly constrained by universal conditions, the variety and complexity of language being attributed to the other components.

Crucial evidence is not easy to obtain, but there can be no doubt as to the empirical nature of the issue. Furthermore, it is often possible to obtain evidence that is relevant to the correct choice of an evaluation measure and hence, indirectly, to the correct decision as to the variety and complexity that universal grammar permits in the several components of the grammar.

To illustrate the problem in an artificially isolated case, consider such words as *feel*, which, in surface structure, take predicate phrases as complements. Thus we have such sentences as:

(1) John felt angry (sad, weak, courageous, above such things, inclined to agree to their request, sorry for what he did, etc.).

We might introduce such expressions into English grammar in various ways. We might extend the categorial component of the base, permitting structures of the form noun phrase–verb–predicate, and specifying *feel* in the lexicon as an item that can appear in prepredicate position in deep structures. Alternatively, we might exclude such structures from the base, and take the deep structures to be of the form noun phrase–verb–sentence, where the underlying structure *John felt* [$_S$*John be sad*]$_S$ is converted to *John felt sad* by a series of transformations. Restricting ourselves to these alternatives for the sake of the illustrative example, we see that one approach extends the base, treating *John felt angry* as a NP–V–Pred expression roughly analogous to *his hair turned gray* or *John felt anger* (NP–V–NP), while the second approach extends the transformational component, treating *John felt angry* as a NP–V–S expression roughly analogous to *John believed that he would win* or *John felt that he was angry*. A priori considerations give us no insight into which of these approaches is correct. There is, in particular, no a priori concept of "evaluation" that informs us whether it is "simpler," in an absolute sense, to complicate the base or the transformation.

There is, however, relevant empirical evidence, namely, regarding the semantic interpretation of these sentences. To feel angry is not necessarily to feel that one is angry or to feel oneself to be angry; the same is true of most of the other predicate expressions that appear in such sentences as (1). If we are correct in assuming that it is the grammatical relations of the deep structure that determine the semantic interpretation, it follows that the deep structure of (1) must not be of the NP–V–S form, and that, in fact, the correct solution is to extend the base. Some supporting evidence from syntax is that many sentences of the form (1) appear with the progressive aspect (*John is feeling angry*, like *John is feeling anger*, etc.), but the corresponding sentences of the form NP–V–S do not (* *John is feeling that he is angry*). This small amount of syntactic and semantic evidence therefore suggests that the evaluation procedure must be selected in such a way as to prefer an elaboration of the base to an elaboration of the transformational component in such a case as this. Of course this empirical hypothesis is extremely strong; the evaluation procedure is a part of universal grammar, and when made precise, the proposal of the preceding sentence will have large-scale effects in the grammars of all languages, effects which must be tested against the empirical evidence exactly as in the single case just cited.

This paper will be devoted to another example of the same general sort, one that is much more crucial for the study of English structure and of linguistic theory as a whole.

Among the various types of nominal expressions in English there are two of particular importance, each roughly of propositional form. Thus corresponding to the sentences of (2) we have the gerundive nominals of (3) and the derived nominals of (4):

(2) a.   John is eager to please.
     b.   John has refused the offer.
     c.   John criticized the book.
(3) a.   John's being eager to please
     b.   John's refusing the offer
     c.   John's criticizing the book
(4) a.   John's eagerness to please
     b.   John's refusal of the offer
     c.   John's criticism of the book

Many differences have been noted between these two types of nominalization. The most striking differences have to do with the productivity of the process in question, the generality of the relation between the nominal and the associated proposition, and the internal structure of the nominal phrase.

Gerundive nominals can be formed fairly freely from propositions of subject–predicate form, and the relation of meaning between the nominal and the proposition is quite regular. Furthermore, the nominal does not have the internal structure of a noun phrase; thus we cannot replace *John's* by any determiner (e.g., *that, the*) in (3), nor can we insert adjectives into the gerundive nominal. These are precisely the consequences that follow, without elaboration or qualifications, from the

assumption that gerundive nominalization involves a grammatical transformation from an underlying sentencelike structure. We might assume that one of the forms of NP introduced by rules of the categorial component of the base is (5), and that general rules of affix placement give the freely generated surface forms of the gerundive nominal:

(5) $[_S \text{NP } nom \text{ (Aspect) VP}]_S$

The semantic interpretation of a gerundive nominalization is straightforward in terms of the grammatical relations of the underlying proposition in the deep structure.

Derived nominals such as (4) are very different in all of these respects. Productivity is much more restricted, the semantic relations between the associated proposition and the derived nominal are quite varied and idiosyncratic, and the nominal has the internal structure of a noun phrase. I will comment on these matters directly. They raise the question of whether the derived nominals are, in fact, transformationally related to the associated propositions. The question, then, is analogous to that raised earlier concerning the status of verbs such as *feel*. We might extend the base rules to accommodate the derived nominal directly (I will refer to this as the "lexicalist position"), thus simplifying the transformational component; or, alternatively, we might simplify the base structures, excluding these forms, and derive them by some extension of the transformational apparatus (the "transformationalist position"). As in the illustrative example discussed earlier, there is no a priori insight into universal grammar – specifically, into the nature of an evaluation measure – that bears on this question, which is a purely empirical one. The problem is to find empirical evidence that supports one or the other of the alternatives. It is, furthermore, quite possible to imagine a compromise solution that adopts the lexicalist position for certain items and the transformationalist position for others. Again, this is entirely an empirical issue. We must fix the principles of universal grammar – in particular, the character of the evaluation measure – so that it provides the description that is factually correct, noting as before that any such hypothesis about universal grammar must also be tested against the evidence from other parts of English grammar and the grammars of other languages.

In the earliest work on transformational grammar [cf. Lees (1960)], the correctness of the transformationalist position was taken for granted; and, in fact, there was really no alternative as the theory of grammar was formulated at that time. However, the extension of grammatical theory to incorporate syntactic features [as in Chomsky (1965, Chapter 2)] permits a formulation of the lexicalist position, and therefore raises the issue of choice between the alternatives. My purpose here is to investigate the lexicalist position and to explore some of the consequences that it suggests for the theory of syntax more generally.

Consider first the matter of productivity. As noted above, the transformation that gives gerundive nominals applies quite freely. There are, however, many restrictions on the formation of derived nominals. The structures underlying (6), for example, are transformed to the gerundive nominals of (7) but not to the derived nominals of (8):

(6) a.   John is easy (difficult) to please.
     b.   John is certain (likely) to win the prize.
     c.   John amused (interested) the children with his stories.
(7) a. *John's being easy (difficult) to please
     b.   John's being certain (likely) to win the prize
     c.   John's amusing (interesting) the children with his stories
(8) a. *John's easiness (difficulty) to please
     b. *John's certainty (likelihood) to win the prize
     c. *John's amusement (interest) of the children with his stories

There are, of course, derived nominals that superficially resemble those of (8), for example, those of (9), which pair with the gerundive nominals of (10):

(9)   a.   John's eagerness to please [(2a), (4a)]
      b.   John's certainty that Bill will win the prize
      c.   John's amusement at (interest in) the children's antics
(10) a.   John's being eager to please [(2a), (3a)]
      b.   John's being certain that Bill will win the prize
      c.   John's being amused at (interested in) the children's antics

These discrepancies between gerundive and derived nominals call for explanation. Specifically, we must determine why the examples of (8) are ruled out although those of (9) are permitted.

The idiosyncratic character of the relation between the derived nominal and the associated verb has been so often remarked that discussion is superfluous. Consider, for example, such nominals as *laughter, marriage, construction, actions, activities, revolution, belief, doubt, conversion, permutation, trial, residence, qualifications, specifications*, and so on, with their individual ranges of meaning and varied semantic relations to the base forms. There are a few subregularities that have frequently been noted, but the range of variation and its rather accidental character are typical of lexical structure. To accommodate these facts within the transformational approach (assuming, as above, that it is the grammatical relations in the deep structure that determine meaning), it is necessary to resort to the artifice of assigning a range of meanings to the base form, stipulating that with certain semantic features the form must nominalize and with others it cannot. Furthermore, the appeal to this highly unsatisfactory device, which reduces the hypothesis that transformations do not have semantic content to near vacuity, would have to be quite extensive.

The third major difference noted above between gerundive and derived nominals is that only the latter have the internal structure of noun phrases. Thus we can have such expressions as *the proof of the theorem* (* *the proving the theorem*, with a gerundive nominal), *John's unmotivated criticism of the book* (*John's unmotivated criticizing the book*), and so on. Correspondingly, the derived nominals cannot contain aspect; there is no derived nominal analogous to *John's having criticized the book*. Furthermore, many derived nominals pluralize and occur with the full range of determiners (*John's three proofs of the theorem*,

several of John's proofs of the theorem, etc.). And derived nominals, in fact, can appear freely in the full range of noun phrase structures. For example, the sentence *John gave Bill advice* is just like any other indirect object structure in that it has the double passive [*advice was given* (*to*) *Bill, Bill was given advice*]. It is difficult to see how a transformational approach to derived nominals can account for the fact that the structures in which they appear as well as their internal structure and, often, morphological properties, are those of ordinary noun phrases. None of these problems arises, as noted earlier, in the case of gerundive nominals.

These properties of derived nominals are quite consistent with a lexicalist approach and, in part, can even be explained from this point of view. Before going into this matter, let us elaborate the lexicalist position in slightly greater detail.

I noted earlier that the lexicalist position was not formulable within the framework of syntactic theory available at the time of Lees's work on nominalizations. The problem was that the obvious generalizations concerning the distributional properties of the base and derived forms were expressible, in that framework, only in terms of grammatical transformations. There was no other way to express the fact that the contexts in which *refuse* appears as a verb and *refusal* as a noun are closely related. However, when the lexicon is separated from the categorial component of the base and its entries are analyzed in terms of contextual features, this difficulty disappears. We can enter *refuse* in the lexicon as an item with certain fixed selectional and strict subcategorization features, which is free with respect to the categorial features [noun] and [verb]. Fairly idiosyncratic morphological rules will determine the phonological form of *refuse*, *destroy*, etc., when these items appear in the noun position. The fact that *refuse* takes a noun phrase complement or a reduced sentential complement and *destroy* only a noun phrase complement, either as a noun or as a verb, is expressed by the feature structure of the "neutral" lexical entry, as are selectional properties. Details aside, it is clear that syntactic features provide a great deal of flexibility for the expression of generalizations regarding distributional similarities. Hence what was a decisive objection to the lexicalist position no longer has any force.

Let us propose, then, as a tentative hypothesis, that a great many items appear in the lexicon with fixed selectional and strict subcategorization features, but with a choice as to the features associated with the lexical categories noun, verb, adjective. The lexical entry may specify that semantic features are in part dependent on the choice of one or another of these categorial features. This is, of course, the typical situation within the lexicon; in general, lexical entries involve certain Boolean conditions on features, expressing conditional dependencies of various sorts. Insofar as there are regularities [. . .], these can be expressed by redundancy rules in the lexicon.

[. . .]

Evidence in favor of the lexicalist position appears to be fairly substantial. It is important, therefore, to look into the further consequences of this position, and the difficulties that stand in the way of incorporating it into the theory of syntax.

Suppose that such phrases as *eagerness* (*for John*) *to please, refusal of the offer, belief in a supreme being*, etc., are base noun phrases. Clearly, if this approach is to be pursued, then the rules of the categorial component of the base must introduce an extensive range of complements within the noun phrase, as they do within the verb phrase and the adjective phrase. As a first approximation, to be revised later on, we might propose that the rules of the categorial component include the following:

(20)   a.   NP → N Comp
       b.   VP → V Comp
       c.   AP → A Comp
(21)   Comp → NP, S, NP S, NP Prep–P, Prep–P Prep–P, etc.

Is there any independent support, apart from the phenomena of derived nominalization, for such rules? An investigation of noun phrases shows that there is a good deal of support for a system such as this.

Consider such phrases as the following:

(22)   a.   the *weather* in England
       b.   the *weather* in 1965
       c.   the *story* of Bill's exploits
       d.   the *bottom* of the barrel
       e.   the *back* of the room
       f.   the *message* from Bill to Tom about the meeting
       g.   a *war* of aggression against France
       h.   *atrocities* against civilians
       i.   the *author* of the book
       j.   John's *attitude* of defiance towards Bill
       k.   his *advantage* over his rivals
       l.   his *anguish* over his crimes
       m.   his *mercy* toward the victims
       n.   a *man* to do the job
       o.   a *house* in the woods
       p.   his *habit* of interrupting
       q.   the *reason* for his refusal
       r.   the *question* whether John should leave
       s.   the *prospects* for peace
       t.   the *algebra* of revolution
       u.   *prolegomena* to any future metaphysics
       v.   my *candidate* for a trip to the moon
       w.   a *nation* of shopkeepers

In each of these, and many similar forms, it seems to me to make very good sense – in some cases, to be quite necessary – to regard the italicized form as the noun of a determiner–noun–complement construction which constitutes a simple base noun phrase. The only alternative would be to regard the whole expression as a

transform with the italicized element being a nominalized verb or adjective, or to take the complement to be a reduced relative clause. In such cases as those of (22), neither alternative seems to be at all motivated, although each has been proposed for certain of these examples. Space prevents a detailed analysis of each case, but a few remarks may be useful.

The analysis of the head noun as a nominalized verb requires that we establish abstract verbs that are automatically subject to nominalization. This requires devices of great descriptive power which should, correspondingly, be very "costly" in terms of a reasonable evaluation measure. Nevertheless, it is an interesting possibility. Perhaps the strongest case for such an approach is the class of examples of which (22i) is an example. It has been argued, quite plausibly, that such phrases as *the owner of the house* derive from underlying structures such as *the one who owns the house*; correspondingly (22i) might be derived from the structure *the one who *auths the book*, **auth* being postulated as a verb that is lexically marked as obligatorily subject to nominalization. However, the plausibility of this approach diminishes when one recognizes that there is no more reason to give this analysis for (22i) than there is for *the general secretary of the party, the assistant vice-chancellor of the university*, and similarly for every function that can be characterized by a nominal phrase. Another fact sometimes put forth in support of the analysis of these phrases as nominalizations is the ambiguity of such expressions as *good dentist* (*dentist who is a good man, man who is good as a dentist*). But this argument is also quite weak. The ambiguity, being characteristic of all expressions that refer to humans by virtue of some function that they fulfill, can be handled by a general principle of semantic interpretation; furthermore, it is hardly plausible that the ambiguity of *good assistant vice-chancellor* should be explained in this way.

For some of the cases of (22), an analysis in terms of reduced relatives is plausible; for example, (22o). But even for such cases there are difficulties in this approach. Notice that there are narrow restrictions on the head noun in (22o). Thus we have the phrase *John's house in the woods* meaning *the house of John's which is in the woods;* but we cannot form *John's book* (*dog, brother, . . .*) *in the woods* (*on the table, . . .*). If John and I each have a house in the woods, I can refer to his, with contrastive stress on *John's*, as *JOHN'S house in the woods*; if we each have a book on the table, I cannot, analogously, refer to his as *JOHN'S book on the table*. Such observations suggest that the surface structure of *John's house in the woods* is *John's – house in the woods*, with *house in the woods* being some sort of nominal expression. On the other hand, in a true reduced relative such as *that book on the table*, there is, presumably, no main constituent break before *book*.

The analysis as a reduced relative is also possible in the case of (22r) and (22s). Thus we have such sentences as (23), with the associated noun phrases of (24):

(23) a.   The question is whether John should leave.
      b.   The prospects are for peace.
      c.   The plan is for John to leave.
      d.   The excuse was that John had left.

(24)   a.   the question whether John should leave
       b.   the prospects for peace
       c.   the plan for John to leave
       d.   the excuse that John had left

Despite the unnaturalness of relative clauses formed in the usual way with (23) as the embedded proposition, one might argue that these are the sources of (24), as reduced relatives. Alternatively, one might argue that the sentences of (23) are derived from structures incorporating (24). The latter assumption is far more plausible however. Thus there are no such sentences as (25):

(25)   a.   *The question whether John should leave is why Bill stayed.
       b.   *The prospects for peace are for a long delay.
       c.   *The plan for John to leave is that Bill should stay.
       d.   *The excuse that John had left was that Bill should stay.

Under the reduced relative assumption, there is no reason why (25) should be ruled out. This would be explained, however, if we assumed that such sentences as (23) are derived from structures incorporating the base noun phrases (24); for example, it might be proposed that (23) derives from (26) by replacement of the unspecified predicate Δ by the complement of the subject noun:

(26) $[_{NP} \text{Det N Comp }]_{NP}$ be $[_{Pred} \Delta]_{Pred}$

Under this analysis, the copula serves as a kind of existential operator. Structures such as (26) are motivated by other data as well; for example, as the matrix structure for such sentences as *what John did was hurt himself*, which might be derived from $[_{NP} \textit{it that John hurt John}]_{NP}$ be $[_{Pred} \Delta]_{Pred}$, through a series of operations to which we return below. In any event, there is an argument for taking the forms of (24) to underlie (23), rather than conversely.

The structures (22), and others like them, raise many problems; they do, however, suggest quite strongly that there are base noun phrases of the form determiner–noun–complement, quite apart from nominalizations. In fact, the range of noun complements seems almost as great as the range of verb complements, and the two sets are remarkably similar. There is also a wide range of adjective complements [*eager* (*for Bill*) *to leave, proud of John*, etc.]. Therefore, it is quite natural to suppose that the categorial component of the base contains rules with the effect of (20), (21), a conclusion which lends further support to the lexicalist assumption.

    [. . .]

Continuing to explore consequences of the lexicalist hypothesis, let us return to the rules (21) which expand NP, VP, and AP into expressions containing optional complements. The phrase category "complement" seems to play no role in transformations. We can easily abolish this category if we replace the rules (21) by a single schema, with a variable standing for the lexical categories N, A, V. To introduce a more uniform notation, let us use the symbol $\overline{X}$ for a phrase containing X as its head. Then the base rules introducing N, A, and V will be replaced by a

schema (48), where in place of . . . there appears the full range of structures that serve as complements and X can be any one of N, A, or V:

(48) $\quad \overline{X} \rightarrow X \ldots$

Continuing with the same notation, the phrases immediately dominating $\overline{N}, \overline{A}$ and $\overline{V}$ will be designated $\overline{\overline{N}}, \overline{\overline{A}}, \overline{\overline{V}}$ respectively. To introduce further terminological uniformity, let us refer to the phrase associated with $\overline{N}, \overline{A}, \overline{V}$ in the base structure as the "specifier" of these elements. Then the elements $\overline{N}, \overline{A}, \overline{V}$ might themselves be introduced in the base component by the schema (49):
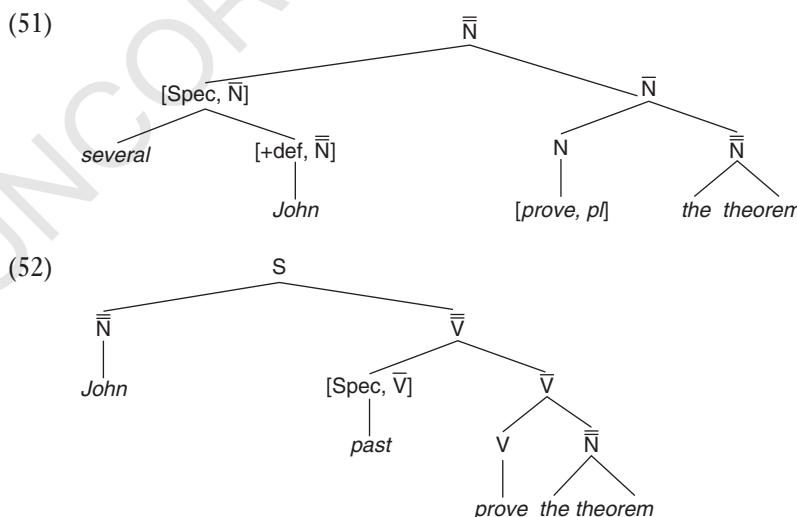
(49) $\quad \overline{\overline{X}} \rightarrow [\text{Spec}, \overline{X}]\, \overline{X}$

where $[\text{Spec}, \overline{N}]$ will be analyzed as the determiner, $[\text{Spec}, \overline{V}]$ as the auxiliary (perhaps with time adverbials associated), and $[\text{Spec}, \overline{A}]$ perhaps as the system of qualifying elements associated with adjective phrases (comparative structures, *very*, etc.). The initial rule of the base grammar would then be (50) (with possible optional elements added):

(50) $\quad S \rightarrow \overline{\overline{N}}\,\overline{\overline{V}}.$

Thus a skeletal form of the base is induced by the "primitive" categories N, A, V (which, as noted earlier, may themselves be the reflection of an underlying feature structure).

In other respects, the primitive categories might differ, for example, if V is analyzed into a copula–predicate construction. Furthermore, it can be expected that the base rules for any language will contain language-specific modifications of the general pattern. If this line of thought is correct, the structure of derived nominals would be something like (51), and the structure of a related sentence, like (52) (omitting much detail):

(51)

(52)

The internal structure of the nominal (51) mirrors that of the sentence (52). The strict subcategorization features of the lexical item *prove* take account of the phrases $\overline{V}$ and $\overline{N}$ dominating the category to which it is assigned in (51), (52), respectively. Its selectional features refer to the heads of the associated phrases, which are the same in both cases. The category $\overline{\overline{N}}$, like S, is a recursive element of the base. Correspondingly, it would be natural to suppose that in the cyclic application of transformations, the phrases of the form $\overline{\overline{N}}$ play the same role as the phrases of the form S in specifying the domain of transformations.

A structure of the sort just outlined is reminiscent of the system of phrase structure analysis developed by Harris in the 1940's. In Harris's system, statements applying to categories represented in the form X^n (*n* a numeral) applied also to categories represented in the form X^m ($m<n$). One might seek analogous properties of the system just analyzed.

[. . .]

3.3   Questions pertaining to Chomsky (1970)

1   The list of noun complements given in (22) in Chomsky's article contains only English examples. Choose another language and see to what extent it allows the same range of possibilities.

2   If English were to turn out to be atypically generous in allowing the wide range of cases given in (22), would that weaken Chomsky's hypothesis concerning the parallelism between nouns and verbs? Give your reasons.

3   Chomsky argues against deriving nominal phrases like *the destruction of the city* from full sentences/propositions. To what extent, if any, has his argument been weakened by the subsequent introduction of the notion "small clause" (cf. Williams 1975 and the papers in Cardinaletti and Guasti 1995).

4   Same question relative to Fodor's (1970) argument against deriving *kill* from *cause to die*. As usual, give your reasons.

5   Chomsky argues for a lexical (nontransformational) approach to English sentences like *That book is readable*. How might he try to integrate (colloquial) sentences like *That book is readable by almost anybody*? Same question for (somewhat marginal) sentences like ?*That chair isn't/ain't sittable in*, with preposition-stranding.

6   How might Chomsky try to cope with the fact that, although English has the same "agentive" preposition *by* in *The bridge was destroyed by the enemy* and *its destruction by the enemy*, Italian makes a distinction, using *da* in sentential passives, yet *da parte di* (a bit like 'on the part of't) in passive-like derived nominals? Do facts like these favor or disfavor (or neither) Collins's (2005) analysis of passives, as compared with Chomsky's?

7   Derived nominals apparently display complements in cases like *the murder of John, the offer of a job*. On the other hand, there are nouns identical in form to verbs (as are *murder* and *offer*) that don't allow such complements, e.g., **the/*their kick of the table* (vs. *They gave the table a kick*), **the/*their punch of the robber* (vs. *They gave the robber a punch*). Find as many other such nouns in English as you can. (Extra credit: Make a proposal as to why *kick, punch*, etc.

have this property and then evaluate it relative to Kayne's (2008) claim that nouns and verbs are not parallel in the way that Chomsky has it.)

8   As Chomsky shows, various complex constructions in which the NP following the verb is not a true direct object of the verb don't allow derived nominals, for example *John considers Mary a genius* vs. **John's/*the consideration of Mary a genius*. The same holds for the corresponding passives, e.g., *Mary is considered a genius by almost everybody* vs. **her consideration a genius (by almost everybody)*. Elsewhere in the paper, Chomsky discusses the absence of derived nominals corresponding to raising constructions, e.g., *That book is certain to be on the reading list* vs. **its certainty to be on the reading list*. How might one try to unify the *consideration examples and the *certainty examples. (Hint: Take a look at Kayne 1981.) Or do you think it would be a mistake to try to do so? If so, give your reasons.

9   One of the examples given in (22) is *the reason for his refusal*. Give arguments both for and against the idea that such nominals involve (partially hidden) relative clauses like *the reason why he refused, the reason for which he refused*.

10   To what extent is it significant that some derived nominals (and adjectivals) seem to be based on phrases rather than on words, e.g., *It must be a quarter after five-ish*, *?the off-the-wall-ishness of his answer?*

11   In languages with more grammatical gender than English, derived nominals necessarily have some gender; for example, in French *destruction* is feminine in gender. To what extent can this be reconciled with Chomsky's idea that the root (here *destruct*) is "free with respect to the categorial features [noun] and [verb]"? (Hint: Take a look at Williams 1981.)

12   This paper contains extensive discussion of various idiosyncrasies involving derived nominals (and forms with *-able/-ible*). Another type of apparent idiosyncrasy in English is *That book is likely/certain/*probable/*possible to be on the reading list*. To what extent can this kind of idiosyncrasy be treated in the same way as the ones discussed in the paper? (Extra credit: Make a proposal that (to some degree) explains why *likely* and *certain* differ in this way from *probable* and *possible*.)

13   Do you think that the case just given is to be grouped with the apparent idiosyncrasy found in *This question is impossible/*?possible to answer*? Why, or why not?

14   In his last footnote, Chomsky rejects deriving *Tomatoes grow* from "NP grows tomatoes" on the grounds that it would imply deriving *Children grow* from the impossible "NP grows children." To what extent do you think the use of "imply" here is justified? Would a distinction between "agent" and "causer" be of help? If so, how exactly? To what extent is your view similar to that of Marantz (1997)?

15   (Extra credit) Causatives like *John is making his tomatoes grow (by . . .)* are called "agentive," as opposed to causatives like *The cold weather is making the tomatoes grow*. Find as many differences as possible between these two types of causative sentences.

16   In his discussion of (29) in the full article, Chomsky argues that the inalienable interpretation of *John's leg* does not involve a relative clause source. If so,

then it might be a case of a base-generated (possessive) specifier of the noun leg. To what extent do you think that the presence of –'*s* is important here?

17 Chomsky provides quite a number of cases of what certainly looks like an idiosyncratic interpretation of derived nominals (and of *-able/-ible* forms). How important is it that idiosyncratic interpretations are also found in sentential constructions like *John blew his top, Put your money where your mouth is*, etc.? (Hint: Take a look at Nunberg et al. 1994).

18 As Chomsky discusses, English gerunds share some properties with derived nominals. On the other hand, English infinitives seem not to. Illustrate this fact in as much detail as possible. Why might it hold?

19 Choose three languages other than English and discuss whether their infinitives are more like English gerunds or more like English infinitives.

20 In (55) in the full article, Chomsky mentions examples like *John's actions are self-destructive*, but he does not explicitly address the contrast with **John's actions are himself-destructive*. What other facts of English (or other languages) might the impossibility of *himself* here be related to? Would one good candidate be the contrast between *They are long-standing Nixon-haters* and **They are long-standing him-haters*? Give your reasons.

21 To what extent might the unpredictable interpretation of the derived nominal "ignorance" (relative to the verb "ignore") be akin to the (partially) unpredictable interpretation of what we call "idioms." Bring in particular examples of idioms from your native language.

References

Cardinaletti, A., and M. T. Guasti (eds). 1995. *Small Clauses. Syntax and Semantics 28*. New York: Academic Press.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Collins, C. 2005. A smuggling approach to the passive in English. *Syntax* 8: 81–120.

Di Sciullo, A. M., and C. Boeckx (eds). 2011. *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford Studies in Biolinguistics, Oxford and New York: Oxford University Press.

Fodor, J. 1970. Three reasons for not deriving 'kill' from 'cause to die'. *Linguistic Inquiry* 1: 429–438.

Harris, Z. 1951. *Methods in Structural Linguistics*. Chicago: University of Chicago Press.

Kayne, R. S. 1981. Unambiguous paths. In R. May and J. Koster (eds), *Levels of Syntactic Representation*, Dordrecht: Foris, pp. 143–183. [Reprinted in Kayne (1984, Ch. 7).]

Kayne, R. S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, R. S. 2008. Antisymmetry and the lexicon. In J. Rooryck and J. van Craenenbroeck (eds), *Linguistic Variation Yearbook*, Amsterdam and Philadelphia: John Benjamins, vol. 8, pp. 1–31. [Reprinted in Kayne (2010, Ch. 9) and in Di Sciullo and Boeckx (2011, Ch. 14).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Lees, R. B. 1960. *The Grammar of English Nominalizations*. The Hague: Mouton.

Marantz, A. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In A. Dimitriadis, L. Siegel, C. Surek-Clark, and A. Williams (eds), *University of Pennsylvania Working Papers in Linguistics*, vol. 4.2, pp. 201–225.

Nunberg, G., I. Sag, and T. Wasow. 1994. Idioms. *Language* 70: 491–538.

Williams, E. 1975. Small clauses in English. In J. Kimball (ed.), *Syntax and Semantics 4*, New York: Academic Press, pp. 249–273.

Williams, E. 1981. Argument structure and morphology. *The Linguistic Review* 1: 81–114.

4

Conditions on Transformations

Noam Chomsky

1973

4.1 Introduction

The framework of Transformational Grammar used in the 1960s and early 1970s made use of a set of phrase structure rules to generate an underlying Deep Structure, and a set of transformational rules that modified this structure in various ways. While this framework generated a wide range of productive work in characterizing the grammar of particular languages, there was a concern that the system was too powerful in that it could express both natural and unnatural patterns. *Conditions on Transformations*, along with other work in the 1970s, provided an answer to this concern, uncovering and formulating constraints on the application of transformational rules.

The general argument in the paper is that in a configuration like (1), a transformational relation between X and Y is ruled out if certain conditions on X, Y, and the context intervening between X and Y hold.

(1)   . . .X. . . [α. . . Y. . .

For example, the TENSED-S CONDITION states that no rule that involves the variables X and Y can apply to the structure in (1) if there is a constituent intervening between X and Y (that is, containing Y but not X) that is a tensed clause. To see how this works, consider (2a), where the embedded clause is tensed, and (2b), where it is not. The Tensed-S Condition predicts that a movement rule such as passivization should yield an ungrammatical result when applied to (2a), where α is a tensed clause, but not when applied to (2b). This prediction is correct, as shown in (3):

(2) a.   I believe [α the dog is hungry].
     b.   I believe [α the dog to be hungry].
(3) a.   *The dog is believed [α is hungry].
     b.   The dog is believed [α to be hungry].

The Superiority Condition states that when a rule can apply to more than one element in a given structure (because those elements are similar in some respect), it must pick the element that is structurally higher one (that is to say, the one that is closer). For example, consider a *wh*-movement rule that could move either *who* or *what* to the front of the sentence in (4). The Superiority Condition forces it to apply to *who*, which is structurally higher than *what*.

(4) . . . do you think [who [read what]]

This idea is in many ways the predecessor of *Minimality* (Chomsky 1986) and *Relativized Minimality* (Rizzi 1990).

In this paper, Chomsky also introduces the Subjacency Condition (the word *sub-jacency* is a blend of *subordination* and *adjacency*). The condition says that a transformation cannot relate elements separated by more than one cyclic node. For Chomsky, the cyclic nodes are the clause, S, and the noun phrase, NP. To generate the dependency found in long-distance *wh*-questions like (5), movement must proceed in multiple steps (to so-called COMP positions), crossing at most one cyclic node at a time, first from Z to Y, then from Y to X, and finally from X to the front of the sentence.

(5) Who do [$_S$ you think X [$_S$ Alice said Y [$_S$ we saw Z]]

As explored extensively in Chomsky (1977), the Subjacency Condition provides an account of several of the island conditions on movement identified by Ross (1967). For instance, a *wh*-phrase cannot move out of a complex NP, from X to the front of the sentence in (6), because such movement would necessarily cross both an S and an NP (on the assumption that NP does not provide a COMP position through which the *wh*-phrase could move).

(6) *What did [$_S$ you hear [$_{NP}$ the claim X [$_S$ John said Y]]]

The basic intuition here expressed as subjacency, namely that movement is subject to locality conditions, is a central tenet within the field of syntax, and has been expressed in a number of different ways since this paper was published. Chomsky's own reformulation, roughly thirty years after the writing of *Conditions on Transformations*, is in terms of the Phase Impenetrability Condition, formulated and discussed in Chomsky (2000, 2001).

4.2   From "CONDITIONS ON TRANSFORMATIONS"

1. From the point of view that I adopt here, the fundamental empirical problem of linguistics is to explain how a person can acquire knowledge of language. For our purposes, we can think of a language as a set of structural descriptions of sentences, where a full structural description determines (in particular) the sound and meaning of a linguistic expression. Knowledge of a language can be expressed in the form of a system of rules (a grammar) that generates the language. To approach the fundamental empirical problem, we attempt to restrict the class of potential human languages by setting various conditions on the form and function of grammars; the term "universal grammar" has commonly been used to refer to the system of general constraints of this sort. With a narrow and restrictive formulation of the principles of universal grammar, it may become possible to account for the remarkable human ability, on the basis of limited and degenerate evidence, to select a particular grammar that expresses one's knowledge of language and makes possible the use of this knowledge.

For heuristic purposes we may distinguish two aspects of universal grammar: (*a*) conditions on form, and (*b*) conditions on function – that is, (*a*) conditions on the systems that qualify as grammars, and (*b*) conditions on the way the rules of a grammar apply to generate structural descriptions. In the terminology of Chomsky (1965, Chapter 1) and earlier work, these are, respectively, conditions on the class $G_1, G_2, \ldots$ of admissible grammars and on the function f that assigns the structural description $SD_{f(i,j)}$ to the sentence S_i generated by the grammar G_j. The distinction is one of convenience, not principle, in the sense that we might choose to deal with particular phenomena under one or the other category of conditions. The distinction might be carried over to particular grammars as well. That is, while it has generally been assumed that particular grammars contain specific rules whereas conditions on the functioning of rules are assigned to universal grammar, there is no logical necessity to make this assumption. It is possible that particular grammars differ in conditions of application, just as it is possible that some specific rules actually belong to universal grammar.

To illustrate, we can consider the enumeration of distinctive features or the specification of the form of phonological rules to be conditions of the first sort, that is, conditions on the form of grammars. Or consider the definition of a grammatical transformation as a structure-dependent mapping of phrase markers into phrase markers that is independent of the grammatical relations or meanings expressed in these grammatical relations. This definition makes certain operations available as potential transformations, excluding others. Thus an operation converting an arbitrary string of symbols into its mirror image is not a grammatical transformation, and transformations generally apply to phrase markers that meet some condition on analyzability with no regard to other associated properties.

To take a standard example, the Passive transformation (reducing it to essentials) applies to any phrase marker that can be "factored" into five successive substrings in such a way that the second and fourth are noun phrases, the third a verb of a particular category (perhaps determined by some semantic property), and the first

and fifth anything at all (including nothing). Thus the structural condition defining the transformation can be given in the form $(Z, \text{NP}, V_x, \text{NP}, Y)$. The transformation rearranges the noun phrases in a fixed way. It will, therefore, apply to the phrase markers underlying the sentences of (1), converting them to the corresponding passive forms:

(1) a. Perhaps–John–read–the book–intelligently
     b. John–received–the book
     c. John–regards–Bill–as a friend
     d. John–painted–the wall–gray
     e. John–expects–the food–to be good to eat

Evidently, the semantic and grammatical relation of the main verb to the following noun phrase varies in these examples (there is no relation at all in (e)), but these relations are of no concern to the transformation, which applies blindly in all cases, producing *Perhaps the book was read intelligently by John, The book was received by John, Bill is regarded as a friend by John, The wall was painted gray by John, The food is expected to be good to eat by John*. By requiring that all transformations be structure-dependent in this specific sense, we limit the class of possible grammars, excluding many imaginable systems.

I will presuppose here, without further discussion, a set of additional conditions on the form of grammar constituting what I have called the "extended standard theory" (see Chomsky (1970b; [1972])). Other conditions on the choice of possible transformations that also seem to me plausible and suggestive, if controversial, are outlined in Emonds (1970).

The conditions on the form of grammar mentioned so far are quite abstract and still permit much too wide a range of potential grammars. One might therefore look for much more specific restrictions. An example, to which I return, is the "Complementizer Substitution Universal" in (2):

(2) Only languages with clause-initial COMP permit a COMP-substitution transformation

This principle presupposes that COMP is a universal element that may appear in various sentence positions and asserts that an item can be moved into COMP position only when COMP is initial. In particular, "*wh*-words" – the relativized constituents in relative clauses or questioned constituents in interrogatives – can be moved only to the left, such movement being permitted only when there is an initial COMP in the phrase to which the transformation is being applied.

It would be quite natural to explore further along these lines. Thus one might try to enumerate "major transformations" such as Question Formation, Imperative, and so on from which languages may draw, with some permitted variation and minor "housekeeping rules" (Bach (1965; 1971)). It may well be that transformations fall into various categories meeting quite different conditions. By constructing a more intricate, more highly articulated theory of grammar in such ways as these, we can perhaps move toward a solution of the fundamental empirical problem.

A second approach attempts to constrain the functioning of grammatical rules and thereby to limit the generative power of grammars of a given form. The earliest suggestions appear in Chomsky (1964[a]) (namely, the condition of "Recoverability of Deletion," the "*A*-over-*A* Condition," and so on). Another example, to which I will return, is the Insertion Prohibition suggested in Chomsky (1965), which prevents transformations from inserting morphological material into sentences that have already been passed in the cycle. Many examples are discussed in a very important study by Ross (1967), where a number of specific conditions are proposed. These conditions are formulated in such a way as to restrict severely the operation of the rules of grammars while not affecting their form. Thus such conditions contribute toward a solution of the fundamental empirical problem.

In this paper I want to consider conditions on the functioning of grammars once again, specifically, conditions on how transformations apply. As noted, I assume here the general framework of the extended standard theory and, in particular, the lexicalist theory of base structures and nominals discussed in Chomsky (1970a). The work leading to the extended standard theory suggested constraints on base structures and on the relations between derivations and semantic representations but said little about transformations. Here, I will explore some conditions on the application of transformations within the framework of the extended standard theory.

[. . .]

Let us restrict our attention initially to rules of extraction that move an item to the left (as in the case of Passive) and to rules of insertion that move an item from the left into an embedded phrase. With this restrictive assumption, we can generalize (19) to (20), incorporating the Insertion Prohibition; we henceforth refer to (20) as the "Tensed-S Condition":[16]

(20)   No rule can involve X, Y in the structure
           $\ldots X \ldots [_\alpha \ldots Y \ldots] \ldots$
           where α is a tensed sentence

To understand the application of the Insertion Prohibition as a special case of this principle, consider the sentences in (21):

(21)   a.   The candidates each hated the other(s)
           b.   The candidates each expected the other(s) to win
           c.   The candidates each expected that the other(s) would win

Dougherty (1970) has argued that such a sentence as *The men hated each other* derives from *The men each hated the other(s)* (ultimately, from *Each of the men hated the other one(s)*) by a rule that moves *each* into the determiner position in *the other(s)*. Assuming this, note that the sentences of (21) should be transformed into those of (22):

(22)   a.   The candidates hated each other
           b.   The candidates expected each other to win
           c.   *The candidates expected that each other would win

Only the first two cases are permitted; (22c) is blocked, as required, by the Tensed-S Condition.

Before turning to other examples, let us consider some facts that lead us to a supplementary principle. Suppose that (23b) derives from the underlying form (23a):

(23) a. John expected [$_s$PRO to win]
 b. John expected to win

Now notice that from (24a) we can derive (24b), whereas from (25a) we cannot derive (25b):

(24) a. The candidates each expected [$_s$PRO to defeat the other]
 b. The candidates expected to defeat each other
(25) a. The men each expected [$_s$the soldier to shoot the other]
 b. *The men expected the soldier to shoot each other

To account for this difference, let us postulate a second principle, the "Specified Subject Condition" (26), where by "specified subject" we mean a subject NP that contains either lexical items or a pronoun that is not anaphoric:

(26) No rule can involve X, Y in the structure
 $\dots X \dots [_\alpha \dots Z \dots - WYV \dots] \dots$
 where Z is the specified subject of WYV in α

We shall return to this principle later to give a more careful formulation. As set forth here, it suffices to distinguish (25), with the specified subject $Z = $ *the soldier* in the embedded sentence α, from (23) and (24), which have no specified subject in that position.[18]

 [. . .]

3. Consider next the sentence (49) which, we assume, derives from (50) by *wh*-Placement (on *something*), *wh*-Movement, and Auxiliary Inversion:

(49) What did you tell me that Bill saw
(50) COMP you told me [$_s$COMP Bill saw something]

The rule of *wh*-Movement in this case appears to violate both the Tensed-S Condition and the Specified Subject Condition.

Before turning to the problem posed by *wh*-Movement, let us consider the notion "transformational cycle" somewhat more carefully. The Insertion Prohibition, now sharpened as a special case of the Tensed-S and Specified Subject Conditions, is a step toward a stricter interpretation of the cycle: it asserts that once a stage of the cycle has been passed, we cannot introduce material into it from the outside under the stated conditions. To further sharpen the notion "transformational cycle," suppose that we impose the general condition (51):

(51)   No rule can apply to a domain dominated by a cyclic node A in such a way
       as to affect solely a proper subdomain of A dominated by a node B which
       is also a cyclic node

In other words, rules cannot in effect return to earlier stages of the cycle after the
derivation has moved to larger, more inclusive domains. We will refer to (51) as the
"Strict Cycle Condition."

From this condition it follows that *wh*-Movement must be a cyclic rule, since it
applies in indirect questions and relatives. The condition (51) seems fairly natural,
and we will proceed to investigate its consequences.

Returning now to (50), we first assign *wh* and apply *wh*-Movement on the
innermost cycle, which gives (52):

(52)   COMP you told me [$_S$[$_{COMP}$what]Bill saw]

On the next cycle, we want to move *what* to the COMP position of the matrix
sentence, to give (49). The Specified Subject Condition is no longer a barrier, but
we are left with a violation of the Tensed-S Condition. An investigation of the
conditions of the violation indicates that they are quite narrow: an item can
"escape" from a tensed sentence if it has been moved into the COMP position on
an earlier cycle and is moving into the COMP position on the present cycle.
Furthermore, in no case does an item in COMP position move to anything other
than the COMP position. These specific properties of COMP may be considered
alongside the property formulated as the Complementizer Substitution Universal.
With the appropriate reformulation of our conditions (which we give as (55)),
wh-Movement can apply to (52), giving (53), which becomes (49) by Auxiliary
Inversion and *that*-Insertion:

(53)   What you told me [$_S$COMP Bill saw]

Suppose now that we replace some of the base rules in (16) [in the full article]
to obtain the more detailed analysis (54) (following Bresnan (1970)):

(54)   S → COMP S′
       S′ → NP Aux VP
       ⋮

Suppose further that we continue to take S (but not S′) to be the domain of
cyclic rules. Under this assumption we can reformulate the Tensed-S and Specified
Subject Conditions, together with the narrow restrictions on COMP, as in (55):

(55)   No rule can involve X, Y in the structure
       $\ldots X \ldots [_\alpha \ldots Z \ldots -WYV \ldots] \ldots$
       where (a) Z is the specified subject of WYV
           or (b) Y is in COMP and X is not in COMP
           or (c) Y is not in COMP and α is a tensed S

This modification of the conditions in effect asserts that an item can be extracted from a tensed sentence or across a specified subject only if there is a rule that moves it into the COMP position. Thus a *wh*-word can be extracted, as in (49)-(50), but the subject of the embedded sentence cannot be passivized in *I believe the dog is hungry*. Notice, however, that *wh*-Movement will not be permitted across a specified subject in [. . .] (56), to give the ungrammatical **Who did you see John's pictures of*:

(56)   COMP you saw [NP John's pictures of who]

The relevant difference between (56) and (50) is that (56) has no COMP node in an NP. Therefore the *wh*-word in (56) cannot escape from the NP.

It is observed in Chomsky (1964) that *wh*-Movement can be applied only once to a constituent of the form S. We cannot, for example, question (or relativize) an item that is within an indirect question to derive (57) from (58):

(57)   *What did he wonder where John put
(58)   COMP he wondered [S COMP John put what where]

To derive (57) from (58), we must first place *where* in the COMP position of the embedded sentence. But in that case, *what* cannot enter the COMP position, which is filled by *where*, and thus cannot be extracted on the next cycle. The principles of the cycle presupposed so far in this discussion permit no other ordering of rule applications to give (57).

4. As the rules and conditions now stand, we can derive (59) from (60) because the embedded S is not tensed:

(59)   What crimes does the FBI know how to solve
(60)   COMP the FBI knows [S COMP PRO to solve what crimes how]

The item *how* is moved into COMP position in the internal cycle, but *what crimes* can be extracted on the next cycle. This is a dubious result, however. Though judgments vary, there seem to me to be severe restrictions on cases such as (59). Thus, (61) seems to me unacceptable, surely much less acceptable than (59); and from (62) the predicted derived sentences (63) and (64) both seem unacceptable, though the immediately underlying forms (65) and (66) are all right:

(61)   *What crimes does the FBI know whether to solve
(62)   COMP John knows [S COMP PRO to give what books to whom]
(63)   *What books does John know to whom to give
(64)   *To whom does John know what books to give
(65)   John knows what books to give to whom
(66)   John knows to whom to give what books

Notice also that (67) does not derive as predicted from (60), but rather only from (68), analogous to (69a) from (69b):

(67)   How does the FBI know what crimes to solve
(68)   COMP the FBI knows [$_S$COMP PRO to solve what crimes] how
(69)   a.   How does the FBI know the code
        b.   The FBI knows the code how

It may be, then, that the *know how to* examples such as (59) are unique in permitting further *wh*-Movement from the embedded sentence and that the general case is that the conditions on transformations prevent movement of a *wh*-phrase over a *wh*-COMP.

In fact, the elaboration of the Specified Subject Condition that we will develop later on will suffice to prevent *wh*-Movement in the cases considered here. However, it may be that there is an independent condition that suffices to prevent *wh*-Movement in these cases. Support for this conjecture is provided by the fact that example (70) must be ruled out as ungrammatical:

(70)   *John knows what who saw

The source, *John knows who saw what*, is grammatical, but (70) and other examples like it indicate that *wh*-Movement cannot move a *wh*-phrase across a *wh*-subject (just as it cannot move a *wh*-phrase across a *wh*-COMP). Notice, however, that *wh*-Movement over a *wh*-phrase P is permissible if P is contained in the predicate phrase, that is, to the right of the verb in the clause in question, as we can see from such examples as (71) and (72):

(71)   John remembers where Bill bought which book
(72)   John remembers to whom Bill gave which book

These examples, to which we shall return, are of a type discussed in Baker (1970). Though judgments vary slightly, such forms are surely more satisfactory than (70). (See also (65), (66).) Thus (71) would have the interpretation (roughly) "John remembers that Bill bought the *i*th book in the *i*th place," and similarly for (72).

Speculating, we might propose a further condition on transformations to accommodate these judgments. The obvious suggestion is based on the observation that the subject NP is "superior" to any phrase in the predicate in the sense that it is closer to the root of the tree structure. More precisely, we say that the category A is "superior" to the category B in the phrase marker if every major category dominating A dominates B as well but not conversely. Suppose that we then add the stipulation (73) to the set of general conditions that we are considering:

(73)   No rule can involve X, Y in the structure
        $\ldots X \ldots [_\alpha \ldots Z \ldots -WYZ \ldots] \ldots$
        where the rule applies ambiguously to Z and Y and Z is superior to Y

The condition requires that a rule must select the superior term where that rule is ambiguous in application, that is, where the structure given in (73) will

satisfy the structural condition defining the rule in question with either Z or Y selected as the factor satisfying a given term of this condition. Like the A-over-A Condition, (73) restricts the ambiguity of rule application. Like (51), it provides a stricter interpretation of the notion of the cycle. It should be noted that in all of the cases of the general form given in (73) that we are considering, the category X is superior to Y in the sense just defined; and where Z is the specified subject, it is superior to Y.

The condition (73) blocks (70) while permitting (71) and (72). Furthermore, the condition (73) suffices to block the examples (61)–(64), independently of other constraints. Thus consider (62), which we restate here as (74):

(74)   COMP John knows [$_s$COMP PRO to give what books to whom]

On the first cycle, *wh*-Movement gives either (75) (= (65)) or (76) (= (66)):

(75)   John knows what books to give to whom
(76)   John knows to whom to give what books

But on the next cycle condition (73) prevents movement of the embedded phrase *to whom* in (75), just as it prohibits movement of the embedded phrase *what books* in (76). The reason, in both cases, is that the *wh*-phrase in COMP of the embedded cycle is superior to these categories, not being dominated by the major category S . If *wh*-Movement moves the superior *wh*-phrase *what books* of (75) or *to whom* of (76) to initial position in the sentence, then the sentence will be marked ungrammatical by virtue of the unfilled COMP of the embedded clause and related considerations that we will examine later. This COMP position cannot be filled by subsequent application of *wh*-Movement by virtue of the general condition (51) that provides the strict interpretation of the cycle. Thus all possibilities are excluded except (75), (76).

To further explore structures of the general form given in (73), let us say that if X is superior to Y in a phrase marker P, then Y is "subjacent" to X if there is at most one cyclic category $C \neq Y$ such that C contains Y and C does not contain X. Thus, if Y is subjacent to X, either X and Y are contained in all the same cyclic categories (and are thus considered at the same level of the transformational cycle) or they are in adjacent cycles. In the sentences of (77), *who* is subjacent to both nodes COMP, but in (78) and (79) it is subjacent only to the node COMP of the embedded sentence:

(77)   a.   COMP he believes [$_S$ COMP John saw who]
        b.   COMP he wonders [$_S$ COMP John saw who]
(78)   a.   COMP he believes [$_{NP}$the claim [$_S$COMP John saw who]]
        b.   COMP he considered [$_{NP}$the question [$_S$ COMP John saw who]]
(79)   a.   COMP he believes [$_S$COMP John saw [$_{NP}$a picture of who]]
        b.   COMP he wonders [$_S$COMP John saw [$_{NP}$a picture of who]]

From (77a) we can derive *Who does he believe that John saw* by iteration of *wh*-Movement. From (77b) we can derive *He wonders who John saw*. From (79a) we can

derive *Who does he believe that John saw a picture of*, again by iteration of *wh*-Movement; from (79b) we derive *He wonders who John saw a picture of*. In the case of (77) and (79), *who* can move to any of the COMP positions; *who* moves ultimately to the external COMP position in (77a) and (79a) and to the internal COMP position in (77b) and (79b). The other four possibilities are excluded, namely, **He believes who John saw*, **Who does he wonder that John saw*, **He believes who John saw a picture of*, and **Who does he wonder that John saw a picture of*. We return to this matter when we consider contextual features of lexical items.

Turning now to (78), we observe that only the embedded COMP position can be occupied by *who*. Thus we can derive from (78b) the sentence *He considered the question who John saw*. No such operation is possible in the case of (78a) because of contextual features of the lexical item *claim*, as we shall discuss at a later point.

Given appropriate contextual features for nouns and verbs, we might account for all of these facts by adding the condition (80), which would restrict rules to adjacent cycles or the same cycle:

(80)   No rule can involve X, Y, X superior to Y, if Y is not subjacent to X

Certain examples that we will consider suggest that (80) does not apply to *each*-Movement. The examples, however, are somewhat marginal, and it may further-more be possible to account for them, under the assumption (80), by a sharpening of the notion "cyclic category." Let us tentatively stipulate, however, that condition (80) applies only to extraction rules, that is, to rules that move some item from the position Y to the superior position X.

5. We can combine conditions (51) and (80) as (81) (slightly reformulating (80)):

(81)   Consider the structure α, where α is a cyclic node:
         $[_\alpha \ldots X \ldots]$

         Suppose that the structural description Σ of a transformation T applies to α, where X is the maximally superior category that satisfies some term of Σ. Then T applies to α only if α is the only cyclic category containing X (condition (51)). If, furthermore, T is an extraction rule moving some category in $\ldots$ to the position X, then some constant term of Σ must hold of a category subjacent to X in α for T to apply to α (modification of condition (80))

We will tentatively suppose that condition (81) is a general property defining cyclic application of transformations.

The examples of (77) and (79) are not affected by the condition (80) incorporated in (81), and we can thus apply iterated *wh*-Movement as desired. But in (78), though *wh*-Movement can apply on the innermost cycle, it cannot apply on the next cycle since NP does not have a COMP. Furthermore, *wh*-Movement cannot apply on the outermost cycle because of the condition (80) incorporated in (81). In this way, we can explain many of the examples that fall under the Complex Noun Phrase Constraint (Ross (1967)); many others fall under the *A*-over-*A* principle.

Notice that this argument is similar to the one that explains the distinction between (50) and (56). For clarity, we repeat the essential point. The sentences (82) and (83) derive from (84) and (85), respectively:

(82)   Who did he expect Bill to see
(83)   *Who did he find Bill's picture of
(84)   COMP he expected [$_S$COMP Bill to see who]
(85)   COMP he found [$_{NP}$Bill's picture of who]

Apart from the element COMP, (84) and (85) are alike, phrase by phrase, from the point of view of rule applicability: *Bill* is the "subject" of *see who* in (84) and of *picture of who* in (85), in our extended sense of the term "subject," and *who* is the "object" in both cases. But (82) is derivable from (84) by iteration of *wh*-Movement, whereas *wh*-Movement cannot apply in (85), on the first cycle because there is no COMP in NP and on the second cycle by virtue of the Specified Subject Condition (55a).

If we are correct in assuming the condition (81), which restricts extraction to adjacent cycles, it follows that although *wh*-phrases can be extracted from such structures as *a picture of_____*, *stories about_____*, *requests for_____*, as in (86), it will not be possible to extract a *wh*-phrase when one of these structures is embedded in another, as in (87), because of the absence of a COMP node in noun phrases. On the other hand, on the same assumptions the forms in (88) are permitted since no extraction rule is involved:

(86)   a.   Who did you see a picture of_____
        b.   Who did you hear stories about_____
        c.   What do you write articles about_____
        d.   What do you generally receive requests for_____

(87)   a.   *Who did you hear stories about a picture of_____
        b.   *What do you receive requests for articles about_____

(88)   a.   We heard stories about pictures of each other
        b.   We received requests for articles about each other

Judgments are insecure, but the conclusion seems to me plausible. On the other hand, Ross (1967) cites such examples as (89) as grammatical:

(89)   What books does the government prescribe the height of the lettering on_____

Examples (87) and (89) appear to be parallel from the point of view of rule applicability. I see no obvious explanation for an apparent difference in degree of acceptability.

From the same assumptions it follows that such sentences as (90) involve repeated cyclic application of *wh*-Movement:

(90)   What did John believe that Bill asked Mary to give her sister to read

It also follows that the surface structure position of the *wh*-phrase will be significant for interpretation and for determination of acceptability, as we might expect under the general assumptions of the extended standard theory.

In Chomsky (1968, Chapter 2, note 23) it is suggested that the *A*-over-*A* principle might be extended to the effect that a transformation must select the minimal phrase of the type S as well as the maximal phrase of the type *A* contained in S. This would account, for example, for the fact that from *John was convinced that Bill would leave before dark* we can derive *John was convinced that before dark Bill would leave* but not *Before dark John was convinced that Bill would leave*. Notice that the Subjacency Condition (81) on extraction rules in effect achieves the same results as the proposed extension of the *A*-over-*A* Condition.

[…]

Notes

16   A weaker assumption would be that α is a <u>language-specific parameter</u> in the condition. In this exploratory study I will do no more than suggest a number of possibilities and investigate their consequences in English.

   Notice that one rule that obviously does not satisfy the condition is Coreference Assignment (however it is formulated). Thus the pronoun can be anaphoric in *John said that he would leave*, for example. The same rule also applies within coordinate structures (for example, *John said that he and Bill would leave*) and others that block various other types of rules.

18   Helke (1971) observes that *each*-Movement is not permitted in such cases as *The candidates each expected the others to clash* (**The candidates expected each other to clash*), *The candidates each expected the others to work together* (**The candidates expected each other to work together*). However, it seems that this results from the operation of independent rules that also exclude **The men walked between each other* from *The men each walked between the others*. What seems to be involved is a restriction on *each*-Movement in the case when the NP *the other* has the features [+totality, –individual] in the system of Dougherty (1970), where some relevant examples are discussed.

4.3 Questions pertaining to Chomsky (1973)

1   Chomsky mentions Emmon Bach's (1965, 1971) idea of a set of "major transformations" that a given language would draw from. To what extent could this idea of Bach's be recast in parametric terms? How would it compare to current ideas about parameters?

2   How would you go about updating Chomsky's (1964b) "Recoverability of Deletion" principle? To what extent might this update apply differently to different subtypes of silent elements? Bring in van Riemsdijk (2002).

3   What are the similarities and differences between Chomsky's (1964b) "A-over-A Condition" and his (1995, 2001) "Minimal Link Condition"?

4   What does the Strict Cycle Condition from this paper have in common with Chomsky's (2001) Phase Impenetrability Condition? What are some important differences? How similar are the notions "escape hatch" and "edge"?

5   How might one reinterpret the "superiority" condition of this paper in terms of the more recent notion of "search space"?

6   To what extent could one interpret Chomsky's subjacency account of Ross's (1967) Complex NP Constraint as an instance of minimalist work in the sense of Chomsky (1995), etc.?

7   In discussing subject-to-object raising of the sort often deemed relevant (cf. especially Postal 1974) to sentences like *We consider John to be honest*, Chomsky suggests that such raising might be prohibited by the requirement that every movement operation change the terminal string. To what extent would such a requirement be plausible from a contemporary perspective?

8   English sentences like *They're trying to make John out to be a liar* look like good candidates for subject-to-object raising across the particle *out*. What would it take to see sentences like *We should send John up a couple of sandwiches* as making the same point?

9   Chomsky's discussion of sentences like *Why are John and Mary letting the honey drip on each other's feet?* suggests that the apparent violation of the Specified Subject Condition (as regards the relation between *John and Mary* and *each other*) is due to the fact that *the honey* is not semantically an agent. To what extent, though, might "agent" be considered to (also) be a syntactic notion? Bring in the unaccusative hypothesis, as well as more recent analyses using v.

10 Note 49 of the full article, which introduces the notion "trace of a movement operation" in a general way, assumes the validity of the rule of agent-postposing in passives, the trace of which would subsequently be filled by the preposing of the object. What exactly makes this kind of derivation implausible in current terms?

11 In the same note, mention is made of sentences like *The offer was sent by John with great glee* and of the fact that agent-postposing makes such sentences compatible with the general restriction that such adverbials must relate to subject position. What later developments in syntactic theory reinforce Collins's (2005) analysis of agents, which dispenses with agent-postposing? (Extra credit: How would Collins's analysis have to be tweaked to make it compatible with Kayne's (1999) proposal that prepositions like *by* come together with their "objects" via internal merge (movement) rather than via external merge?)

12 Chomsky proposes deriving the pseudo-cleft-like *The only people they really like are each other* via movement of *each other* into post-copula position from an original position as object of *like*. What is the relation between this proposal and his (1995, Ch. 3) discussion of reconstruction effects? How could such a derivation be made compatible with Kayne's (1994) antisymmetry?

13 Chomsky seems to agree with Bever (1970) that the exclusion of sentences like **John is intelligent is obvious* (as opposed to *That John is intelligent is obvious*) has to do with perceptual strategies that are sensitive to the presence of *that* or to morphological marking on the verb. In what way do English subjunctives suggest that this approach is insufficiently general?

14 At a number of points, Chomsky calls upon surface filters, which are developed further in Chomsky and Lasnik (1977). What properties of these filters have led to their no longer being part of current syntactic theory?

15 Explicate the relation between subjacency and successive cyclicity.

References

Bach, E. 1965. On some recurrent types of transformations. In C. W. Kreidler (ed.), *Sixteenth Annual Round Table Meeting Linguistics and Language Studies*, vol. 18 of *Monograph Series on Languages and Linguistics*, Washington, DC: Georgetown University Press, pp. 3–18.

Bach, E. 1971. Questions. *Linguistic Inquiry* 2: 153–166.

Baker, C. L. 1970. Notes on the description of English questions: The role of an abstract question morpheme. *Foundations of Language* 6: 197–219.

Bever, T. G. 1970. The influence of speech performance on linguistic structure. In G. B. Flores d'Arcais and W. J. M. Levelt (eds), *Advances in Psycholinguistics*, Amsterdam: North Holland, pp. 4–30.

Bresnan, J. 1970. On complementizers: Toward a syntactic theory of complement types. *Foundations of Language* 6: 297–321.

Chomsky, N. 1964a. *Current Issues in Linguistic Theory*. The Hague: Mouton.

Chomsky, N. 1964b. The logical basis of linguistic theory. In H. G. Lunt (ed.), *Proceedings of the Ninth International Congress of Linguists*, The Hague: Mouton, pp. 914–978. [Reprinted as *Current Issues in Linguistic Theory*.]

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. 1968. *Language and Mind*. New York: Harcourt Brace Jovanovich.

Chomsky, N. 1970a. Deep structure, surface structure, and semantic interpretation. In R. Jakobson and S. Kawamoto (eds), *Studies in General and Oriental Linguistics* (Commemorative Volume for Dr. Shiro Hattori), Tokyo: TEC Corporation for Language Research, pp. 52–91.

Chomsky, N. 1970b. Remarks on nominalization. In R. Jacobs and P. Rosenbaum (eds), *Readings in English Transformational Grammar*, Waltham, MA: Ginn, pp. 184–221.

Chomsky, N. 1972. Some empirical issues in the theory of transformational grammar. In S. Peters (ed.), *Goals of Linguistic Theory*, Englewood Cliffs, NJ: Prentice-Hall, pp. 63–130.

Chomsky, N. 1977. On Wh-movement. In P. Culicover, T. Wasow, and A. Akmajian (eds), *Formal Syntax*, New York: Academic Press, pp. 71–132.

Chomsky, N. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2000. Minimalist inquiries: The framework. In R. Martin, D. Michaels, and J. Uriagereka (eds), *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, Cambridge, MA: MIT Press, pp. 89–155.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, pp. 1–52.

Chomsky, N., and H. Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8: 425–504.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. 1968. *Language and Mind*. New York: Harcourt Brace Jovanovich.

Chomsky, N. 1970a. Deep structure, surface structure, and semantic interpretation. In R. Jakobson and S. Kawamoto (eds), *Studies in General and Oriental Linguistics* (Commemorative Volume for Dr. Shiro Hattori). Tokyo: TEC Corporation for Language Research.

Chomsky, N. 1970b. Remarks on nominalization. In R. Jacobs and P. Rosenbaum (eds), *Readings in English Transformational Grammar*, Waltham, MA: Ginn, pp. 184–221.

Chomsky, N. 1972. Some empirical issues in the theory of transformational grammar. In S. Peters (ed.), *Goals of Linguistic Theory*, Englewood Cliffs, NJ: Prentice-Hall, pp. 63–130.

Collins, C. 2005. A smuggling approach to the passive in English. *Syntax* 8: 81–120.

Dougherty, R. 1970. A grammar of coordinate conjoined structures: I. *Language* 46: 850–898.

Emonds, J. 1970. *Root and structure-preserving transformations*. Doctoral dissertation, MIT.

Helke, M. 1971. *The grammar of English reflexives*. Doctoral dissertation, MIT. [Published 1979 by Garland Publishers, New York.]

Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. S. 1999. Prepositional complementizers as attractors. *Probus* 11: 39–73. [Reprinted in Kayne (2000, Ch. 14).]

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Postal, P. 1974. *On Raising*. Cambridge, MA: MIT Press.

van Riemsdijk, H. 2002. The unbearable lightness of *GOing*. The projection parameter as a pure parameter governing the distribution of elliptic motion verbs in Germanic. *Journal of Comparative Germanic Linguistics* 5: 143–196.

Rizzi, L. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Ross, J. R. 1967. *Constraints on variables in syntax*. Doctoral dissertation, MIT. [Published 1986 as *Infinite Syntax!* by ABLEX, Norwood, NJ.]

5

On Grammatical Relations and Clause Structure in Verb–Initial Languages

Stephen R. Anderson and Sandra Chung

1977

5.1   Introduction

Scholars of language of all traditions have felt the need to distinguish between sub-jects and objects. How is this distinction encoded in a view of grammar for which sentences are not simply strings of words, but rather have hierarchical structure? For languages in which the object is the noun phrase (NP or DP) next to the verb, it can be seen as forming a constituent with the verb, to the exclusion of the subject: $[_{VP} \text{V NP}]$ (or $[_{VP} \text{NP V}]$, the standard view for OV word order at the time this paper was written). But what about VSO languages, in which the subject occurs between the verb and the object? They give rise to (at least) the two following questions:

(1)   Is there empirical evidence suggesting that the subject and the object are syntactically distinct?
(2)   If so, how should that distinction be expressed?

In this paper, Anderson and Chung provide a clear and compelling, empirically grounded answer to the first question, and outline possible ways of answering the second one. In so doing, they provide crucial evidence in support of the existence of hierarchical structure across languages, and against the hypothesis that some languages might have a flat structure.

Anderson and Chung show that, in a number of VSO languages from different families, the subject and the object systematically exhibit distinct behavior, in that there are certain syntactic operations that target the former but not the latter. For example, take cases in which the subject of an infinitival clause can be null when it is coreferential with one of the arguments of the matrix clause. This case, which is

called CONTROL in current terms, at the time of this article was seen as resulting from a rule of EQUI-NP DELETION, which deleted the subject of the embedded clause when it was identical with the closest noun phrase in the matrix clause. Now, suppose that subjects and objects were not structurally distinct: we would then expect that, in a VSO language, this rule would target the NP that immediately follows the verb in an infinitival clause, regardless of whether it is the subject or the object. However, this is not the case: drawing examples from a number of unrelated languages, it is shown that only subjects can be null in this environment, not objects, even when they are indistinguishable in terms of linear order. This kind of contrast between subjects and objects is replicated in a number of other cases.

Given that the distinction between subjects and objects exists even in VSO languages, where the subject is closer to the verb than the object in linear order, should it be expressed as a primitive of grammar, or cast in structural terms? Anderson and Chung's article favors the second solution, though it does not make a precise proposal concerning how the verb comes to be separated from the object. Since then, the hypotheses that the verb can raise from the head of VP to a higher functional head (cf. Emonds 1978), and that the subject starts in a position lower than the landing site of the raised verb have made it possible to capture the central intuitions of this paper: that objects always differ from subjects, and that this distinction should be expressed in structural terms, by viewing objects as originating closer to the verb than subjects, regardless of the linear word order in which they may appear.

5.2 From "ON GRAMMATICAL RELATIONS AND CLAUSE STRUCTURE IN VERB-INITIAL LANGUAGES"

1 Introduction

In one of the fundamental papers of the Generative Semantics literature, McCawley (1970) argued that the underlying word order in English clauses is not verb medial (SVO), as might be expected on the basis of surface structure, but rather verb initial (VSO).

[. . .]

McCawley's position is criticized at length and in detail by Berman (1974). Her arguments are essentially of two sorts. First, she attempts to demonstrate that there is in fact no reason to posit underlying VSO structures for English at all, since the rule simplifications claimed by McCawley are either illusory or incorrect. [. . .] We will not be further concerned here with these issues, internal to the study of English.

Berman goes on to give a second class of arguments, however, which are of more general interest. By showing that there are some syntactic processes in English (and perhaps in other languages for which the same issue could arise) for which such an order is actually necessary, these points are intended to establish a positive case for verb-medial structures. The arguments of this second type which Berman provides have the following form, in outline: In a verb-initial structure, both subject and

object NP are directly dominated by the node S. There is, therefore, no structural difference between these NPs apart from their relative order. As a consequence, a rule which removed one or the other of them would produce a structure which would be formally indistinguishable from that of a basic intransitive clause. That is, there would be no way to identify whether the remaining NP in such a reduced transitive clause was originally subject or object. Thus on this hypothesis the single remaining NP in such a clause ought to behave, for the purpose of further syntactic operations, exactly the same as does the subject of an intransitive clause. Berman's arguments then take the form of showing that this conclusion is false. Certain rules apply to a clause which consists of just a verb and a single NP differentially, depend-ing on whether that NP represents an original subject or the object of a reduced clause. If this is correct, and if, furthermore, transformations are to be stated in terms of the structural properties of phrase markers, then it must follow that sub-ject and object are structurally distinct in some way that persists even after the subject has been removed from a transitive structure. On the hypothesis of VSO order, there is no obvious structural property which can perform this function, as was just noted. There is one obvious property, however, if SVO structures are basic: Subjects are preverbal, while objects are postverbal, regardless of whether the subject is still present. Reduced transitive clauses are distinguished from basic intransitives as VO versus SV, a property to which syntactic rules can have access.

It should be clear that Berman's arguments, even if correct in every detail, do not really establish a case for SVO order. That is, the argument outlined in the preceding paragraph makes no reference whatsoever to the actual order of elements in the clause, except inferentially: While it would certainly be possible to distinguish subject from object in terms of position relative to the verb, this is by no means the only possibility. For instance, if the verb and its object(s) form a subconstituent of the clause (namely, a VP), of which the subject is not a member, then the required differentiation between subject and objecw could be carried out on the basis of constituent structure alone, without reference to order. This is essentially the approach to grammatical relations taken by Chomsky (1965); it represents a posi-tion which enjoys wide acceptance among contemporary syntacticians. On this approach, underlying SOV order would be just as satisfactory as would SVO from the point of view of the argument developed by Berman. Yet another possibility would be simply to treat 'subject' and '(direct) object' as primitives of clause struc-ture, not defined in terms of any other properties, but rather the basic relations that determine the makeup of clauses. This is, of course, the position of Relational Grammar. On this line, any order whatsoever of the basic constituents of the clause (or indeed a totally unordered structure) would be equally satisfactory, since the problems raised by Berman would be resolved directly by a sort of deus ex machina which is independent of clause-internal ordering.

Insofar as Berman's argument is valid, then, it demonstrates the need for some mechanism in terms of which the grammatical relations borne by NPs within a clause can be distinguished, without determining the form of such a mechanism. Berman notes in her paper that an important source of further information on this score would be provided by the investigation of languages for which there is no reason to doubt the correctness of underlying VSO order. In such a language,

where VSO structures are not to be derived from some other order at more abstract levels of representation, it would apparently be impossible for the internal structure of clauses to involve a subconstituent of the type VP, since the verb and its object(s) are separated by the subject, and hence are noncontiguous. If, in such a language, facts of the same sort as those underlying Berman's arguments from SVO languages could be shown to obtain, this would show that internal constituent structure of the usual type could not form the basis for the required differentiation. In the absence of evidence to the contrary, we assume that in VSO languages there is no basis for positing a sort of pseudo-VP consisting of the verb and the subject, excluding the object(s), which could serve the same function. Constituent structure would apparently not be adequate to resolve our problem, then. Clearly, order relative to the verb (or anything else?) would be no more satisfactory; this was the basis of Berman's argument against a VSO analysis for English. Arguments of this type from VSO languages, then, would seem to be intimately related to the claims of Relational Grammar that grammatical relations are primitives of clause structure which are not to be defined derivatively in terms of other aspects of phrase-marker structure. If such arguments can be found, they would support the need for such a notion. On the other hand, if such arguments are lacking precisely for the class of VSO languages, this would indicate that another device – such as order relative to the verb (available only for SVO languages) or clause-internal constituent structure (available for SVO, SOV, and VOS but not for VSO languages) – is a more appropriate way to define grammatical relations. A negative demonstration of this sort, to the effect that no arguments of a given type can be found within a certain domain, is, of course, next to impossible to construct; fortunately, we do not have to do this.

It is the aim of this paper to present some arguments of the type originally provided by Berman,[2] taken from languages whose VSO typology is not seriously in question. The languages at issue are Tongan and Samoan, from the Polynesian family, and Breton, a Celtic language. [. . .]

2 Examples from Polynesian languages

We consider first an argument from the interaction of an Equi-NP Deletion rule and Subject-to-Object Raising in Samoan. The Equi rule in Samoan optionally deletes an embedded subject under identity with a controller NP in the next higher clause. The problems of selecting a controller are much the same as the familiar ones which arise in English and other languages; the NP deleted in the lower clause, however, is always a subject. The rule need not apply, as in the sentences of (1S):[3]

(1S) a. *'Ua mānana'o tagata e mālō lātou i le pālota.*
          perf want-pl people fut win they in the election
          'People wanted to win in the election.'
    b. *'Ua māfaufau Tupu e fa'alogo 'oia i le lāuga.*
          perf decide Tupu fut listen she to the sermon
          'Tupu decided that she would listen to the sermon.'

If Equi applies to these sentences, it results in the sentences of (2S):

(2S) a.   *'Ua mānana'o tagata e   mālō i le pālota.*
          perf want–pl   people fut win in the election
          'People wanted to win in the election.'

      b.   *'Ua māfaufau Tupu e   fa'alogo i le   lāuga.*
          perf decide   Tupu fut listen   to the sermon
          'Tupu decided to listen to the sermon.'

The Samoan rule of Subject-to-Object Raising is also analogous to the English rule. This rule optionally raises the subject of an embedded clause, which becomes the direct object of the higher clause. In (3S), the rule has not applied:

(3S) a.   *Sā   mana'o Tupu e   'emo le   uila.*
          past want   Tupu fut flash the lightning
          'Tupu wanted the lightning to flash.'

      b.   *E mānana'o tagata 'ia   manuia le   pālota.*
          fut want–pl   people irreal be–well the election
          'People want the election to turn out well.'

When Subject-to-Object Raising applies, the sentences of (4S) result:

(4S) a.   *Sā   mana'o Tupu i le   uila   e   'emo.*
          past want   Tupu at the lightning fut flash
          'Tupu wanted the lightning to flash.'

      b.   *E mānana'o tagata i le   pālota 'ia   manuia.*
          fut want–pl   people at the election irreal be–well
          'People want the election to turn out well.'

Both Equi and Raising are subject to some further conditions and qualifications (cf. Chung 1976 for details), but the two rules are clearly part of the grammar of Samoan, and distinct from one another.

    Significantly, there are a few verbs in Samoan which govern both Subject-to-Object Raising and Equi. *Mana'o* (plural *mānana'o*) 'want', for instance, allows either rule to affect its complement [cf. (2Sa) versus (4Sb), for example]. Now both of these rules affect subjects; since Samoan has an underlying VSO word order, both would be stated so as to affect the first NP following the verb (in the absence of any further internal structure). In that case, however, the application of one of these rules to a transitive complement ought to produce a reduced clause superficially like a basic intransitive. If this is indeed true, it ought to be possible now to apply the other of the two rules. From a sentence to which Equi has applied, like (5Sa), it should be possible to derive (5Sb) by Raising:

(5S) a.   *Sā   mānana'o tagata e   pu'e le gaoi.*
          past want–pl   people fut catch the burglar
          'People wanted to catch the burglar.'

      b.   **Sā   mānana'o tagata i le gaoi   e   pu'e.*
          past want–pl   people at the burglar fut catch
          'People wanted the burglar to catch.'

Sentence (5Sb), however, is not possible: Subject-to-Object Raising cannot apply to objects, regardless of whether the subject is still present.

The converse of the facts in (5S) is also true of Samoan. Where Raising has applied to remove the subject NP from a complement, as in (6Sa), we might expect to be able to apply Equi to delete the NP that is left in the reduced clause:

(6S) a. *E le'i   mana'o Tupu i leoleo e pu'e 'oia.*
         fut not–yet want   Tupu at police fut catch her
         'Tupu didn't want the police to catch her.'
      b. **E le'i   mana'o Tupu i leoleo e   pu'e.*
          fut not–yet want   Tupu at police fut catch
          'Tupu didn't want the police to catch.'

The ungrammaticality of (6Sb) shows, however, that this is not the case: Objects cannot be deleted by Equi even when the subject is no longer present in the clause.

The impossibility of constructions like (5Sb) and (6Sb) cannot be accounted for if linear order relative to one another is the only parameter which distinguishes subject from object structurally in a VSO language like Samoan. On this view, there would be no natural way of preventing Equi and Subject-to-Object Raising from applying to one another's outputs on a purely structural basis. The fact that the two rules do not interact in such a way would follow immediately, however, from a proposal which distinguished subject and object in some other terms. In such a framework, the rules would not be stated to affect the NP following the verb, but rather the NP which has the structural properties of a subject. If one rule removes the subject, there is no reason to expect that another NP in the clause (the object) would "inherit" the subject properties. Thus, neither rule could apply to the output of the other.

[. . .]

Other, similar arguments could be adduced, but the above should be sufficient to establish our conclusion. [. . .] Subject cannot be distinguished from object, then, by order relative to the verb. Since the argument is directly parallel to that of Berman for English, we must conclude that, in general, arguments of this sort establish the need for a structural difference between subject and object, and are irrelevant to the question of whether English has basic SVO order.

[. . .]

3 Conclusion: On clause structure in VSO languages

We have demonstrated above that Berman's conclusion concerning the need for a structural difference between subject and object can be shown to follow in languages such as Samoan, Tongan, and Breton, where the basic order is VSO. It thus cannot be the case that her argument demonstrates the need for underlying SVO order in English, though of course there may well be (indeed, we feel there are) independent arguments establishing this conclusion. If subject and object are to be structurally distinguished in a VSO structure, though, it clearly cannot be by order relative to the verb (i.e. with subject specified as $NP/\_\_V$, and object as $NP/V\_\_$).

It would appear to be the case that a constituent VP, including the verb and its object(s), is impossible, since the subject intervenes between these elements. Thus, there does not seem to be a natural constituent structure division of the clause that will accomplish the desired effect. We would appear to have arrived at a strong argument in favor of the basic position of Relational Grammar, to wit, that grammatical relations must be taken as primitives of clause structure, neither ignored nor defined solely in terms of other structural properties of clauses such as linear order or internal constituency and domination.

As the alert reader will no doubt have perceived from our choice of words in the preceding paragraph, however, the evidence does not seem unambiguous in support of this conclusion. In particular, there is some evidence from Breton that a constituent VP, consisting of just the verb and its object(s), may need to be posited despite the fact that this unit could not form a continuous constituent of the usual sort.

As we have discussed at some length above [in the full article], Breton has a rule of Topicalization which can apply in main clauses to front any single constituent. But Topicalization cannot apply to more than one constituent in a single clause:

(39B)   *_Hiziv e Kemper e tebro Yannig krampouezh._
            today in Quimper prt will-eat Johnny crepes
            'Johnny will eat crepes **in Quimper today**.'

[. . .] Either *hiziv* 'today' or *e Kemper* 'in Quimper' can be topicalized by itself, but together the two do not form a larger constituent, and so both cannot be fronted at the same time. On the other hand, where two elements do form a single constituent, they can be topicalized together. [. . .] We can also topicalize the head along with the possessor:

(40B)   a.   *Mab Per   a   zo klañv.*
                son Peter prt is sick
                '**Peter's son** is sick.'
        b.   *Buoc'h an den a   varvas.*
                cow   the man prt died-3sg
                '**The man's cow** died.'

We conclude that the Topicalization rule applies to exactly one constituent of a clause.

Thus far, most of our examples of Topicalization have involved the fronting of either an NP or an adverb. There is another kind of Topicalization, however, which is statistically of high frequency in the language: It is also possible to topicalize the main verb of a clause. When this happens, the topicalized verb shows up in the form of an infinitive, and in the place of the original finite verb is found an inflected form of the verb *ober* 'to do', marked for person, number and tense. This option is not limited to activity verbs: It is applicable to any verb in the language with the exception of *bezañ* 'to be' and the periphrastic verb *endevout* 'to have'. The 'do' in this case is thus closer to the *do* produced in English by Verb-Phrase Deletion (*he ought to know that, but he **doesn't***) than to the verb of *do so* and related constructions:

(41B) a. *C'hoarzhiñ a ra   ar baotred.*
          to-laugh   prt does the boys
          'The boys laugh.'

      b. *Mont a rin   ganit   d'ar   pardon.*
          to-go prt do-fut-1sg with-you to-the pardon
          'I will go with you to the pardon.'

Since the topicalization of the main verb has the effect of focusing on the action rather than on one of the participants in it, it is appropriate for 'neutral' discourse, and accordingly the construction in (41B) is much used.

    The interesting point about this construction for our purposes, though, is the following: As we have noted, constituents are subject to Topicalization. Clearly the verb itself is a constituent, and so we would expect to be able to topicalize it. There is also another possibility, however: When the main verb is itself transitive, it is possible to topicalize just the verb, as in (42Ba) below, or alternatively to topicalize the verb with its object(s):

(42B) a. *Deskiñ a   reomp Brezhoneg.*
          to-learn prt do-lpl Breton
          'We are learning Breton.'

      b. *Deskiñ Brezhoneg a   reomp.*
          to-learn Breton   prt do-lpl
          'We are learning Breton.'

      c. *Lenn   eul levr brezhoneg a ran   bemdez.*
          to-read a   book Breton   prt do-1sg every-day
          'I read a Breton book every day.'

How are we to account for this possibility? It seems clear that the construction in (42Bb–c), which is in no way marginal but rather is perfectly productive in the language, must be taken to imply that the verb together with its object(s) form a constituent of some sort. While we will not describe all of these possibilities here, the material that can be treated as a 'verb phrase' for the purposes of Topicalization in Breton is essentially the same as that which is usually taken to make up the verb phrase in English: The verb together with direct and indirect objects, reduced complement clauses, and directional adverbs, but not with adverbs of time, place, manner, and others generally thought of as 'sentence adverbials' rather than 'verb phrase adverbials'.

    [. . .]

    Note that it is NOT possible to topicalize the verb and its subject, which are contiguous:

(43B) a. **Deskiñ ni a   ra Brezhoneg*
           to-learn we prt do Breton
           '**We** are **learning** Breton.'

b.   **Lenn   Yannig a   ra   eul levr brezhoneg bemdez.*
      to-read Johnny prt does a   book Breton   every-day
      '**Johnny reads** a Breton book every day.'

There are several ways we could envision the creation of a constituent with the appropriate properties. We might, for example, analyze all sentences in which Verb Topicalization is possible as having their apparent main verbs embedded as complements of a matrix verb *ober* 'to do', with this structure undergoing Equi followed by a sort of incorporation rule which, if Topicalization has not taken place, would replace the inflected form of *ober* with a form of the embedded verb. Mechanically, this is probably the most satisfactory alternative, but in the absence of any further evidence for the analysis of virtually all clauses in Breton as involving such embedding, we are somewhat uneasy about it.

Another solution to the problem would have to involve extending our notion of constituent structure in some way. We might simply allow discontinuous constituents, an option which has occasionally been suggested in the structuralist literature, but seldom taken seriously within the framework of generative grammer [sic]. Or, on the other hand, we might allow transformational rules to make reference to 'constituents' that are unitary only in some sort of semantic structure.

[. . .]

There is really no completely satisfactory solution to our problem, On the one hand, we must be able to treat the verb and its complements as a constituent of some sort for the purposes of Topicalization; on the other hand, no satisfying and well-motivated way to do this has been proposed thus far. Nonetheless, we can conclude that the verb and its complements do form a unit at some levels of structure in Breton, a unit which does not include the subject; and from this it follows that there may well be a way to distinguish subjects and objects in terms of constituent structure, at least of this (possibly extended) sort. In that case, it might no longer be necessary to take the relational terms as primitive. It would be interesting to see whether whatever device is appropriate here can be extended naturally to other languages in which a configurational definition of the fundamental relations in clause structure is apparently not available.

[. . .]

Regardless of the resolution of this issue, however, the thrust of our arguments above is the following: In VSO languages, as well as in other types of languages, rules exist which are sensitive to a structural difference between subjects and objects. The difference in question is independent of the linear position in the clause of the NP affected, and persists after the deletion of one or the other of the NPs involved in a transitive construction. From this it must follow that subject and object are structurally distinct; and in the nature of things, this distinction cannot be taken to refer to position relative to the verb in a VSO language. This, in turn, suggests that the distinction is probably not (solely) in terms of position relative to the verb in a language like English, either. Whether the required relational distinctions can be coded in terms of some notion of hierarchical constituent structure, or should be taken to be unanalyzable primitives, must remain at present an open question.

Notes

2   The arguments below are presented in the framework of "classical" generative syntactic description. In particular, we ignore the possible implications for our analyses of recent proposals for a "trace theory of movement rules." Such a theory, since it allows a degree of 'global' reference to earlier stages of a derivation (cf. Chomsky 1975 for some discussion), would have extensive implications for the sort of rules discussed here.

3   The languages from which the examples are taken are indicated by letters: e.g. (1S) is from Samoan, (7T) from Tongan, (18B) from Breton, and (44E) from English.

5.3   Questions pertaining to Anderson and Chung (1977)

1   Chomsky's (1995) copy theory of movement came in well after the writing of this paper. Had the copy theory come in before this paper, how might it have affected Anderson and Chung's arguments?

2   Discuss some similarities and some differences between this paper of Anderson and Chung's and Massam (2000, Ch. 29 of this volume).

3   Anderson and Chung make use of V-raising, but make no explicit use of VP-raising. Had they wanted to use VP-raising and been able to look ahead, what might have been the implications of Haddican (2007)?

4   What aspects of Anderson and Chung's paper are directly supported by Kayne (1981)? With respect to what aspects of their paper is his paper neutral?

5   How does the discussion of hierarchy vs. linear order in Anderson and Chung compare with Kayne (2011)?

6   One set of verb-initial languages is the Polynesian languages. Those languages appear to lack any overt copula. Test the hypothesis that all verb-initial languages lack an overt copula.

7   To what extent is the general usefulness of the term 'verb-initial' affected by the consideration of embedded sentences?

8   How might the positioning of subjects in Irish nonfinite sentences be related to Pollock (1989, Ch. 15 of this volume)?

9   Where might one expect to find adverbs in verb-initial languages? Bring in both Cinque (1999) and Pearson (2000).

10   Anderson and Chung note that Tongan has clitic placement for subject pronouns, which end up between the initial tense morpheme and the verb, but no clitic placement for object pronouns. Find as many languages as you can that do have clitic placement for object pronouns.

11   What kind of parameter might underlie the difference between those verb-initial languages that have clitic placement for object pronouns and those that do not have it? Bring in Roberts and Shlonsky (1996).

12   As discussed by Anderson and Chung, direct object pronouns in Breton must be preceded by a preposition that does not otherwise appear with direct objects. Find as many other languages as you can that have this property.

To what extent might Spanish direct objects be seen as having a related property? (Extra credit: Discuss the implications of these preposition facts for the theory of accusative Case.)

13   What aspects of the Breton verb topicalization discussed by Anderson and Chung are affected by den Besten and Webelhuth's (1990) later work? What aspects of Breton verb topicalization are relevant to Haddican (2007)?

14   Noonan (1993) discusses the fact that Irish lacks transitive stative verbs (relating it to Irish lacking a transitive verb corresponding to English *have* – cf. Harves and Kayne 2012). Evaluate the conjecture that the only languages to lack transitive statives entirely are verb-initial languages. To what extent is this conjecture relatable to Mahajan (1994)?

15   Most Niger-Congo languages are VO languages, that is, languages in which the verb normally ends up preceding the object. Yet there appears to be not a single verb-initial Niger-Congo language. From the perspective of Massam's paper in this volume, how might one account for this property of Niger-Congo? (Extra credit: Bring in the postverbal subject clitics in Wolof discussed by Torrence 2005).

16   Evaluate Baker and McCloskey's (2007) suggestion as to why VSO languages are rarer than SVO or SOV languages. Bring in Woolford (1991, section 2.1).

References

Baker, M., and J. McCloskey. 2007. On the relationship of typology and theoretical syntax. *Linguistic Typology* 11: 273–284.

Berman, A. 1974. On the VSO hypothesis. *Linguistic Inquiry* 5: 1–38.

den Besten, H., and G. Webelhuth. 1990. Stranding. In G. Grewendorf and W. Sternefeld (eds), *Scrambling and Barriers. Linguistik Aktuell/Linguistics Today 5*, Amsterdam and Philadelphia: John Benjamins, pp. 77–92.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. 1975. *Reflections on Language*. New York: Pantheon Books.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chung, S. 1976. *Case marking and syntactic relations in the Polynesian languages*. Doctoral dissertation, Harvard University.

Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York and Oxford: Oxford University Press.

Emonds, J. E. 1978. The verbal complex V′-V in French. *Linguistic Inquiry* 9: 151–175.

Haddican, B. 2007. On *egin: do*-support and VP-focus in Central and Western Basque. *Natural Language and Linguistic Theory* 25: 735–764.

Harves, S., and R. S. Kayne. 2012. Having 'need' and needing 'have'. *Linguistic Inquiry* 43: 120–132.

Kayne, R. S. 1981. Unambiguous paths. In R. May and J. Koster (eds), *Levels of Syntactic Representation*, Dordrecht: Foris, pp. 143–183. [Reprinted in Kayne (1984, Ch. 7).]

Kayne, R. S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, R. S. 2011. Why are there no directionality parameters? In M. B. Washburn, K. McKinney-Bock, E. Varis, A. Sawyer, and B. Tomaszewicz (eds), *Proceedings of the 28th West Coast Conference on Formal Linguistics*, Somerville, MA: Cascadilla Proceedings Project, pp. 1–23.

McCawley, J. 1970. English as a VSO language. *Language* 46: 286–299.

Mahajan, A. 1994. The ergativity parameter: *have–be* alternation, word order and split ergativity. In M. Gonzàlez (ed.), *Proceedings of NELS 24*, Amherst: GLSA, University of Massachusetts, vol. 1, pp. 317–331.

Massam, D. 2000. VSO and VOS. Aspects of Niuean word order. In A. Carnie and E. Guilfoyle (eds), *The Syntax of Verb Initial Languages*, New York and Oxford: Oxford University Press, pp. 97–117.

Noonan, M. 1993. Statives, perfectives, and accusativity: The importance of being HAVE. In J. Mean (ed.), *Proceedings of the 11th West Coast Conference on Formal Linguistics*, Stanford: CSLI, pp. 354–370.

Pearson, M. 2000. Two types of VO languages. In P. Svenonius (ed.), *The Derivation of VO and OV*, Amsterdam and Philadelphia: John Benjamins, pp. 327–363.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Roberts, I. G., and U. Shlonsky. 1996. Pronominal enclisis in VSO languages. In I. G. Roberts and R. D. Borsley (eds), *The Syntax of the Celtic Languages*, Cambridge: Cambridge University Press, pp. 171–199.

Torrence, H. 2005. *On the distribution of complementizers in Wolof*. Doctoral dissertation, University of California, Los Angeles.

Woolford, E. 1991. VP-internal subjects in VSO and nonconfigurational languages. *Linguistic Inquiry* 22: 503–540.

6

On Wh–Movement

Noam Chomsky

1977

6.1 Introduction

Chomsky's *On Wh-Movement* marked the beginning of a new conception of syntactic theory, one in which grammatical possibilities are determined through the interaction of general principles rather than construction-specific rules.

Syntactic theory in the early 1970s widely assumed that *wh*-questions were derived from a rule displacing a *wh*-phrase to the COMP position at the front of the clause proposed in Bresnan (1970). In *On Wh-Movement*, Chomsky observes that this *wh*-movement rule, in addition to involving COMP, is also characterized by a certain set of other properties. First, the *wh*-movement transformation abides by the Complex NP and *wh*-island constraints explored in Ross (1967) and Chomsky (1973). Second, it can violate other locality constraints (such as the Specified Subject constraint and Subjacency), though only when there is a certain kind of embedding environment, known as a "bridge." Finally, the *wh*-movement transformation has the property of leaving a gap.

Chomsky raises the question of whether we might gain understanding into the nature of grammar by taking this set of properties to constitute the signature of a deeper process that underlies not only *wh*-questions, but also other superficially disparate constructions. His discussion in this article focuses on constructions that exhibit the *wh*-movement signature properties. Take for example comparative clauses, as seen in (1):

(1) a. John is taller than Mary is.
  b. *John is taller than I heard the claim that Mary is.
  c. John is taller than you think Mary is.

In comparatives, the gap corresponds to the adjective phrase missing after the copula. And though it is not obvious that these examples involve displacement to COMP, Chomsky points out that there are varieties of English that exhibit comparatives involving overt *wh*-expressions, such as *John is taller than what Mary is*. Further, comparatives are subject to the complex NP constraint, as shown by (1b), but can overcome other locality constraints, such as Subjacency, when there is an embedding bridge verb, such as *think* in (1c). Chomsky observes that this pattern extends to a range of constructions including relative clauses (*The students that we admire*), topicalization (*Those politicians, I would rather not see at the party*), clefts (*It is out of spite that the students refused to hand in their assignments*), *easy-to-please* sentences (*John is easy for us to please*), and infinitival complements of adjectival phrases (*John is tall enough for us to see*).

Prior work had derived each of these constructions from a distinct transformational rule; on such a view, the fact that they share a set of properties is unexplained. This paper suggests instead that whenever we find a construction that involves the signature of properties mentioned above, we should attribute its formation to the same abstract rule of *wh*-movement. Residual differences among these various sentence types are derived from independent properties and constraints. On this view, the fact that these various kinds of sentences share certain characteristics is principled, and the rule system of the grammar can be simplified.

The fundamental impact of *On Wh-Movement* consists, on the one hand, in raising the plausibility of analyses involving movement in cases where such movement is not immediately apparent, and, on the other hand, in giving rise to a conception of syntactic analysis with very general processes, whose applicability is subject to certain constraints. This represented an important step toward the Principles and Parameters framework, which has characterized research in generative syntax since the early 1980s. In that framework, rules as general as *wh*-movement were further generalized to move-α, with a more articulated set of constraints deriving the properties of landing sites and gaps. Many of the questions explored in current syntactic theory (for example, is a certain movement, say clitic climbing, an instance of head movement, A-movement, or A′-movement?) presuppose the diagnostic method laid out in this article. The importance of *On Wh-Movement* is reflected and echoed in many works since, and in particular in Cheng and Corver (2006), a volume dedicated to exploring its impact on the field. Chomsky (2001, Ch. 30 of this volume) offers an insight into some of the same issues from a perspective that is a quarter century more recent.

6.2   From "ON WH-MOVEMENT"

I will presuppose, in this paper, the general framework of the extended standard theory (EST), as outlined, for example, in Chomsky (1972, 1975b) and references cited there; and more specifically, the assumptions explored in Chomsky (1971, 1973, 1974, 1975b,c) and related work cited in these references. I want to examine some proposals put forth tentatively in the work cited and in so doing, to revise and extend some of the particular analyses and principles investigated.

I will first review and somewhat reformulate some of the background assumptions drawn from earlier work and then apply them to several questions in English syntax.

I assume that a grammar is a theory of competence and that universal grammar (UG) is in essence a system of principles specifying the nature of linguistic representations and the rules that apply to them, and the manner in which these rules apply. A grammar (strongly) generates a set of structural descriptions and (weakly) generates a language, assigning one or more structural descriptions to each sentence of the language (and, in principle, to all potential sentences). A structural description of a sentence consists of a representation of the sentence on each linguistic level (cf. Chomsky, 1955). I assume that two of these levels are the levels of phonetic representation (PR) and what I will call "logical form" (LF), meaning by the latter the level that expresses whatever aspects of semantic representation are determined by properties of sentence-grammar. Cf. Chomsky (1975a,b,c) for discussion. Thus a grammar assigns to each sentence, in particular, a pair of representations (*pr, lf*), where *pr* is drawn from PR and *lf* from LF.

In accordance with EST, I assume here that a grammar consists of base rules, transformational rules, phonological rules and (semantic) interpretive rules. The base consists of a categorial component and a lexicon, the former satisfying the principles of some version of the X-bar theory (for recent discussion see Hornstein, 1975, Selkirk, 1975; Halitsky, (1975); Emonds (1976); Bresnan, [1976]; Jackendoff, [1977]), and the latter of the general character developed in Aronoff (1976). The base generates an infinite class of deep structures (initial phrase markers). I assume that thematic relations in the sense of Jackendoff (1972) and related work are determined by interaction of lexical properties and configurations of deep structures. The transformational component of the grammar generates derivations $D = (K_1, \ldots, K_n)$, where K_1 is a base-generated deep structure, K_{i+1} is formed from K_i by a transformation, and no obligatory transformation is applicable to K_n.

The derivation D must be related to PR and LF. I will have little to say here about the relation to PR. As for LF, I assume that it is determined by interpretive rules applying to K_n. Under this assumption, it must be that thematic relations are properly expressed in K_n, though determined at K_1. I will assume that this is the case, in accordance with trace theory, as outlined in the references cited above. If so, then interpretive rules extend the derivation D, carrying K_n to a representation in LF. These interpretive rules are the rules SI-1 of Chomsky (1975b,c). It is in fact misleading to call these "rules of semantic interpretation," as in these references and elsewhere; they are more properly described as rules concerned with the syntax of LF. Note that K_n will not be surface structure in the familiar sense. It is more "abstract," by virtue of trace theory, and may be subject to nontransformational rules (e.g., "scrambling"). Some crucial aspects of PR may be determined by the extended derivation from K_n to LF. Thus, as noted first by Lees (1960), deletion seems sensitive to some aspect of semantic representation, and under the present theory that means that the possibilities of deletion are in part fixed by properties of representations at LF

or between K_n and LF. Cf. Sag [1976 a,b] for an analysis of such rules as VP-deletion and gapping along these lines.

This outline is extremely sketchy, and the analyses cited are not even mutually compatible in detail. I present it only so as to locate the following discussion within a familiar general framework.

I will be concerned now with a kind of "core grammar" for English consisting of a few general rules and some general conditions governing the operation of these rules. The rules in question include two transformational rules (1) and three interpretive rules (2):

(1) a. Move NP
     b. Move *wh*-phrase
(2) a. Reciprocal rule: assign to *each other* the feature [+anaphoric to *i*] in a structure containing NP_i
     b. Bound anaphora: assign to a pronoun the feature [+anaphoric to *i*] in a structure containing NP, in the context $[_{NP}\text{—Possessive—}N_x]$
     c. Disjoint reference: assign to a pronoun the feature [–anaphoric to *i*] in a structure containing NP_i

The rules of (2) are among those that Kenneth Hale has called "rules of construal" (cf. Hale, 1976). An informal explanation of their meaning will do for now. Let us assume that there is some standard method for indexing nonterminal symbols in deep structures, in particular, NPs; transformations will preserve the property that all nonterminals are indexed, in ways to be discussed. If *each other* is assigned the feature [+anaphoric to *i*], then the structure . . . NP_i . . . *each other* . . . (or . . . *each other* . . . NP_i . . .) is assigned the appropriate reciprocal interpretation, whatever this may be (for discussion, see Fiengo and Lasnik 1973; Dougherty 1974). A pronoun marked [+anaphoric to *i*] will be interpreted in LF as anaphoric to NP_i; the relevant choice of N_x will be essentially as discussed in Helke (1970), including, for English, *self*, so that English (nonemphatic) reflexive is understood as bound anaphora. A pronoun marked [–anaphoric to *i*] will be understood as disjoint in reference to NP_i; cf. Chomsky (1973); Lasnik [1976]. I assume that this rule falls under a more general rule of disjoint reference applying (in somewhat different ways) to all NPs. To make these vague remarks explicit, it is necessary to explain what is meant by the term "anaphoric." I assume that there is a procedure for introducing variables for NPs in LF, including pronouns, and that the notions "anaphoric," "nonanaphoric" will be understood as determining the choice of variables as the same or different. For present purposes, nothing much depends on how rules (2) are implemented, so I will not pursue the matter; as far as I can see, nontrivial questions arise in the case of (2a) and plural pronouns, the latter, a special case of problems concerning the semantics of plurality. I will assume that the rules (2) and others ultimately give representations in LF in a rather conventional form, with quantifiers and variables, for some empirical arguments, cf. Chomsky (1975c).

I assume that the rules (1) and (2) meet the following conditions:

(3)   Cycle: transformational rules, e.g., (1), meet the condition of the (strict) cycle; the subjacency condition is a property of cyclic rules, i.e., part of the definition of the cycle.
(4)   Propositional-island condition (PIC)
(5)   Specified subject condition (SSC)

I understand the notion of the cycle here in the sense of Chomsky (1973, (51)), with the qualifications given there. Assuming that transformational rules are either cyclic or postcyclic, it follows from this formulation that the rules (1), specifically (1b), are cyclic, since they apply in embedded structures.[1] I will understand the subjacency condition as holding that a cyclic rule cannot move a phrase from position Y to position X (or conversely) in (6):

(6) $\ldots X \ldots [_\alpha \ldots [_\beta \ldots Y \ldots] \ldots] \ldots X \ldots$, where α and β are cyclic nodes

For the present, I will take the cyclic nodes to be $\bar{S}$ and NP; on the effect of other choices, see below.

The subjacency condition applies to cyclic rules only; hence to cyclic transformational rules but not to interpretive rules or to postcyclic transformational rules. Thus for many people (myself included), such examples as (7) and (8) are fully acceptable:

(7)   *we want very much* [$_{\bar{S}}$ *for* [$_{NP}$ *pictures of each other*] *to be on sale*]
(8)   *the men expected* [$_{\bar{S}}$ *that* [$_{NP}$ *pictures of each other*] *would be on sale*]

Similarly, a postcyclic rule such as the major case of French clitic movement (cf. Kayne, 1975) need not, on these assumptions, meet the condition of subjacency.

It follows that rightward-movement rules are "upward bounded" (cf. Ross, 1967; Akmajian, 1975). But I am assuming that the same is true of "lowering rules" such as quantifier movement, and leftward-movement "raising" rules. It is easy enough to find phenomena that appear to violate the subjacency condition. Consider, e.g., the sentences (9), (10), where there is a relation between the phrase in bold face and the position marked by *t*, "violating" subjacency under the assumption that the rule in question is a movement rule:

(9)   *John seems* [$_{\bar{S}}$ *to be certain* [$_{\bar{S}}$ *t to win*]]
(10) *who did Mary hope* [$_{\bar{S}}$ *that Tom would tell Bill* [$_{\bar{S}}$ *that he should visit t*]]

Putting the matter more carefully, a proposed condition on rules, such as subjacency, cannot be confirmed or refuted directly by phenomena of this (or any other) sort. A condition on rules can be confirmed or refuted only by rules, which observe or violate it, respectively. If the rule of NP-movement that yields (9) applies successive cyclically, as often assumed, then the rule will observe subjacency. If, as I have argued in the references cited, the rule of *wh*-movement applies

successive cyclically, then it too will observe subjacency, giving (10). To find evidence to support or to refute a proposed condition on rules, it does not suffice to list unexplained phenomena; rather, it is necessary to present rules, i.e., to present a fragment of a grammar. The confirmation or refutation will be as convincing as the fragment of grammar presented. This is a simple point of logic, occasionally overlooked in the literature. The status of conditions on rules is empirical, but evidence can only be indirect and the argument, one way or another, is necessarily rather abstract and "theory bound."

The conditions (4) and (5) (PIC and SSC) refer to structures of the form (11), where α is a cyclic node:

(11) $\ldots X \ldots [_\alpha \ldots Y \ldots] \ldots X \ldots$

As in the case of subjacency, I will take $\overline{S}$ and NP to be the cyclic nodes, delaying the discussion of other choices until later. PIC (the "tensed-S condition" of the references cited) asserts that no rule can "involve" X and Y where α is a finite clause (tensed-S). SSC asserts that no rule can "involve" X and Y where α contains a specified subject, i.e., a subject not containing Y and not controlled by X (I modify an earlier formulation here; I assume that Y contains Y). If α contains a subject, then only the subject is accessible to rule, if the subject is specified in the defined sense.

The term "involved in" was left deliberately vague in the exploratory studies cited above, as was the category of rules to which the conditions are relevant. We may sharpen the formulation somewhat to include the desired cases and exclude unwanted ones. Let us restrict attention to rules specified in terms of a structural condition and a structural change, in the usual sense of transformational grammar (cf. Chomsky, 1955, 1961; Chomsky and Miller, [1958, 1964]; Peters and Ritchie, 1973). We furthermore restrict attention to structural conditions of the elementary form (12), where α_i is a constant or $\alpha_i = vbl$, and each constant may be either a single element of the X-bar system or a terminal string (perhaps only a single symbol):

(12) $(\alpha_1, \ldots, \alpha_n)$

A terminal string with the successive factors $x_1, \ldots, x_n$ and the phrase marker K is subject to the structural change, with these factors, just in case $(x_1, \ldots, x_n)$ is analyzable as (12) with respect to K; i.e., x_i is an α_i with respect to K, where an arbitrary string is a vbl. Cf. references cited, and Chomsky (1975c).

We now say that a transformational rule *involves* X and Y when it moves a phrase from position X to position Y and a rule of construal *involves* X and Y when it assigns Y the feature [±anaphoric to i], where X has the index i (or conversely, in both cases). The two cases will be unified below.

Following a suggestion of Jean-Roger Vergnaud, we modify the definition of PIC, stipulating that α is the cyclic node immediately dominating the category of Y. Then rule (2b), giving (8), will not violate PIC. For discussion of the effect of PIC and SSC on postulated rules of grammar, see Chomsky (1971, 1973, 1974, 1975b,c); Lasnik and Fiengo (1974); Kayne (1975); Fiengo and Lasnik (1976); Quicoli [1976a,b; 1980]; Pollock (1976).

Plainly, rules can vary from language to language within the constraints imposed by UG, but it is often assumed that conditions on rules must be invariant. This assumption is somewhat arbitrary; cf. Ross (1967); Bresnan (1972); Chomsky (1973). There is no a priori reason not to assume the opposite, and in fact, a very high level of explanatory adequacy might well be attained by a theory of UG that permitted either rules or conditions to vary, within fixed limits. To consider a case in point, Kim (1976) observes that rules of anaphora in Korean meet a condition rather like PIC, but with a somewhat different condition on α of (11). There is no formal distinction in Korean between tensed and nontensed clauses, but there is a category of embedded clauses that are not islands, much like the infinitival clauses of English and the Romance languages: namely, the complements of a certain class of "assertive" verbs. It is interesting that these verbs are very close in meaning to the verbs that in English take infinitives. Thus we can formulate a variant of PIC for Korean, with the condition on α modified, and we can suggest a somewhat more abstract formulation of PIC of which English and Korean are special cases. In the absence of more extensive work on rule systems in other languages, I am reluctant to suggest anything further. Note again that evidence bearing on questions of this degree of abstractness requires a fairly credible grammatical analysis, since only rules, not phenomena, have bearing on the validity of conditions on rules.

[. . .]

[W]e assume that when a phrase moves by a transformation, its category remains as an "unfilled node," and that the moved phrase and the original position have the same index. The unfilled node labelled i is $t(i)$, the trace of P_i, the phrase moved from position i. The trace will invoke SSC and is available for assignment of thematic relations. PRO and trace are identified; they differ only with respect to the origin of the index. The position of trace may be filled by a phrase containing a variable, by expansion of a quantifier. There may be phonetic effects of trace in the latter case.

The rules and conditions given so far permit *wh*-movement within a clause, giving such sentences as (40), but not extraction of *wh*-phrases from a clause, as in (41):

(40)   *who did Mary meet t*
(41)   *who did you tell Mary that she should meet t*

The two cases are in fact quite different in character. Many languages permit the first but not the second (e.g., Russian, German). Furthermore, whereas *wh*-movement within a clause is unconstrained, extraction from a clause is lexically governed, as has frequently been remarked. Thus we have such examples as (42):

(42)   a.   **what did John complain that he had to do this evening*
       b.   **what did John quip that Mary wore*
       c.   *?who did he murmur that John saw*

Just what property of the matrix VP permits it to be a "bridge" (in the sense of Erteschik, 1973), permitting escape of the *wh*-phrase from the $\overline{S}$ "island," is unclear, Some proviso is necessary, however.

Suppose that we formulate the basic rule of *wh*-movement essentially as (43):

(43)   move *wh*-phrase into COMP

The rule will apply freely clause-internally, but will not yet move the *wh*-phrase over a bridge. We may then formulate a language-specific COMP-COMP movement rule (44):

(44)   move *wh*-phrase from COMP to a higher COMP over a bridge

The structural description of this rule (subject to modifications about placement in COMP to be discussed) will be approximately (45):

(45)   (COMP, *X*, *wh*-phrase, *vbl*), where *X* contains a VP with certain special properties

If we incorporate the "bridge" properties in (45), then the rule will not fall strictly within the format we have proposed for transformational rules. Moreover, under the relative interpretation of conditions discussed before, it might be argued that the conditions are inapplicable; more precisely, it is easy to see how "involved in" can be sharpened so as to make them inapplicable, along the lines discussed earlier. Suppose, alternatively, that we dispense with (45) and interpret the "bridge" conditions as conditions on rules of interpretation. Then COMP-COMP movement by (43) will be blocked by the conditions. We must therefore introduce a language-specific proviso in (11), for English, namely, (46):

(46)   where *Y* is not in COMP

Which of these approaches is preferable is unclear. I will assume the latter, without much reason. Thus we add the language-specific proviso (46) to (11), permitting COMP-COMP movement, and we assume that the "bridge" conditions fall within the interpretive rules, either SI-1 or SI-2 (cf. Chomsky, 1975b, c; Erteschik, 1973).

Sentence (41) will be formed, as in the references cited, by successive-cyclic application of *wh*-movement, now understood to be reapplication of (43). The rule is subject to all of the conditions on movement rules, so that we have the consequences already noted.

Continuing to adopt the framework of the references cited, as modified above, I will assume that the rule (43) places a *wh*-phrase within the COMP node to the left of [±WH], which is realized phonetically as *that, for*, or null. There are a number of apparently rather idiosyncratic rules that determine the phonetic realization of the items in COMP. A formulation given in Chomsky (1973) can be considerably improved and extended, but I will not go into the matter here. One general rule for Modern English is that sequences of the form *wh*-phrase + complementizer are not

permitted, as they were in earlier stages of the language. Thus we will have rules such as (47), (48):

(47)   *wh*-phrase becomes null
(48)   a.   *that* becomes null
        b.   *for* becomes null

One of the three must apply. By general conditions on recoverability of deletion, which we may assume to exist though they are not understood in detail, (47) will be inapplicable when the *wh*-phrase contains actual lexical content (e.g., preposi-tions, possessives, etc.). The rules (48) apply more broadly; e.g., *that* can be deleted under certain circumstances in nonrelatives, *for* is deleted immediately following verbs of the *want* category and under certain circumstances before *to*, etc.

I will assume that the *wh*-phrase moved by the rule is as determined by Bresnan's relativized A-over-A principle (cf. Bresnan, [1976]; Woisetschläger, 1976, Sag, [1976a] for somewhat different versions).

The rule of *wh*-movement has the following general characteristics:

(49)   a.   it leaves a gap
        b.   where there is a bridge, there is an apparent violation of subjacency, PIC, and SSC
        c.   it observes CNPC
        d.   it observes *wh*-island constraints

The properties (49) follow, on the theory outlined, from the assumption that *wh*-movement moves a phrase (implying (a)), observes SSC, PIC, and subjacency (implying (c) and (d)), and is permitted from COMP-to-COMP under "bridge" conditions (implying (b)).

So far, I have been recapitulating and somewhat revising earlier work. Now I want to turn to the main question of this paper, namely, (50):

(50)   Where we find the configuration (49) in some system of data, can we explain it on the assumption that the configuration results from *wh*-movement?

In other words, does the configuration (49) serve as a kind of "diagnostic" for *wh*-movement. That it may has been suggested, quite tentatively and without elaboration, in earlier work. I now want to investigate the plausibility of the contention. The following remarks, then, have a narrower and a broader aim. The narrower aim is to provide evidence that certain examples with the configuration (49) may in fact plausibly be understood as cases of *wh*-movement. The stronger aim is to suggest that this may be true in general. By the logic of the question, the stronger proposal cannot be demonstrated but only suggested.

I will assume, following the analysis in the references cited, that *wh*-movement is what underlies restrictive and nonrestrictive relatives and direct and indirect questions. There are, of course, some distinctions among these cases. Some of them can be accounted for by considering the contexts in which the *wh*-movement

rule applies. E.g., questions but not relatives can have *wh*-movement of adjective phrases, but this distinction will obviously follow from the rule of relativization, whether it is a raising rule (cf. Vergnaud, 1974) or an interpretive rule. In other cases, stipulation may be necessary to distinguish some types from others (though this is not obvious), but if so, there seems no compelling reason to suppose that the stipulation is a condition on the *wh*-movement rule itself, though even if it were, it would not materially affect the point at issue.

Apart from these cases, the best-studied relevant example is the case of comparatives. It has been frequently noted (first, I believe, by David Vetter) that comparatives essentially have the properties (49), and it was therefore proposed in Chomsky (1973) and Vergnaud (1974) that "comparative deletion" is in reality a case of *wh*-movement. The contrary position is argued by Bresnan in an important article (Bresnan, 1975), which, together with Bresnan (1972, 1973), constitutes the most extensive and illuminating study of comparatives available. The issue is complex. Let me try to sort it out.

First, is there evidence for a *wh*-movement rule underlying comparatives? For some dialects of English, there is direct evidence for such a rule, as noted in Bresnan (1972). Thus many dialects of American English normally have such comparatives as (51):

(51) a. *John is taller than what Mary is*
      b. *John is taller than what Mary told us that Bill is*

For such dialects, the comparative rule is virtually identical to the general rule of *wh*-movement. Subject to the qualifications given above, it seems that the rule postulated for relatives and questions can simply extend to comparatives, with essentially no change. The properties (49) will then follow directly.

But there is evidence (Richard Kayne, personal communication) in support of a *wh*-movement analysis for other dialects of English as well. Consider the sentence (52), where brackets bound internal cyclic nodes:

(52) a. *Mary isn't the same as* [*she was five years ago*]
      b. *Mary isn't the same as* [*John believes* [*that Bill claimed* [*that she was five years ago*]]]
      c. *Mary isn't the same as* [*John believes* [*Bill's claim* [*that she was five years ago*]]]
      d. *Mary isn't the same as* [*I wonder* [*whether she was five years ago*]]

This construction has the properties (49). The "gap" is an adjective phrase, just as in comparatives; we can replace "the same as" by "taller than" throughout. There are similar constructions in which even the phrase *the same* does not appear, as in (53), etc.:

(53) a. *Mary is (more or less) as she was five years ago*
      b. *Mary is rather like John thought she was* [in colloquial English]
      c. *Mary isn't as John believes that Bill claimed that she was five years ago*

In these cases, a deletion analysis, if possible at all, seems rather artificial, since in contrast with comparatives, there is no overt matrix phrase that can trigger and control the deletion. We can easily account for (52–3) by a *wh*-movement rule of the sort postulated for the dialects that permit (51). The rule will give (54a), just as it gives (54b) in the dialects that have an overt *wh*-form in comparatives:

(54) a. *Mary isn't (the same) as* [*what she was five years ago*]
      b. *Mary isn't taller than* [*what she was five years ago*]

Sentence (54b), for dialects that do not permit it, can be regarded as the structure underlying (55) by a rule of *wh*-phrase deletion, falling under (47):

(55) *Mary isn't taller than she was five years ago.*

The same rule will give (52–3). The dialects differ, then, in obligatoriness of *wh*-phrase deletion; as noted, this and related rules are subject to a variety of apparently rather idiosyncratic conditions.

According to this analysis, the sentences of (52)–(53) are regarded as analogous to those of (56):

(56) a. *Mary isn't different than* [*what she was five years ago*]
      b. *Mary isn't different than* [*what John believes* [*that Bill claimed* [*that she was five years ago*]]]
      c. **Mary isn't different than* [*what John believes* [*Bill's claim* [*that she was five years ago*]]]
      d. **Mary isn't different than* [*what I wonder* [*whether she was five years ago*]]

Examples (56c,d) are ruled out by subjacency, PIC, and SSC. Under the analysis that presupposes (54a) underlying (52a), (53a), the same is true of (52c, d), etc.

Proceeding, we may treat *as*, *than* as prepositions, analogous to *than* in (56). This seems reasonable anyway; it means that such sentences as (57) will be analyzed as having final prepositional phrases of the form P NP, rather than being derived by deletion of *be* from (58):

(57) *John is taller than Bill*
(58) *John is taller than Bill is*

Cf. Hankamer (1973) for arguments supporting this analysis of (57).

The analysis of (52–3) along these lines seems natural and perhaps compelling. If it is correct, then all dialects that permit (52–3) have a rule of *wh*-movement forming comparatives. Therefore, there is no need for a new rule of comparative deletion.

If this is correct, we might propose further that there do not exist rules of "deletion over a variable." Thus the category of permissible rules is reduced, always a welcome step. Furthermore, we have some support for a positive answer to the question (50). Correspondingly, we have some evidence that the island constraints

of (50iii, iv) [sic] can be explained in terms of general and quite reasonable "computational" properties of formal grammar (i.e., subjacency, a property of cyclic rules that states, in effect, that transformational rules have a restricted domain of potential application; SSC, which states that only the most "prominent" phrase in an embedded structure is accessible to rules relating it to phrases outside; PIC, which stipulates that clauses are islands, subject to the language specific "escape hatch" (46)). If this conclusion can be sustained, it will be a significant result, since such conditions as CNPC and the independent *wh*-island constraint seem very curious and difficult to explain on other grounds. Whether or not these further consequences prove tenable, it seems clear that a strong argument would be required to show that English has a second rule of comparative deletion that gives exactly the same forms as the independently motivated and quite general *wh*-movement rule (subject, again, to the qualification on p. [00]). It would be rather paradoxical for a language to contain a general rule of *wh*-movement forming all comparatives (and much else), along with a second rule (comparative deletion) that is extensionally identical (as a mapping) with the first over the subdomain of structures such as (58).

[. . .]

Bresnan notes that comparatives have the cross-over properties discussed by Postal, Wasow and others. She then argues that cross-over properties are not a diagnostic for movement rules, on her assumption that comparatives are formed by a deletion rule. If she is correct, it would follow that the explanation for cross-over suggested in Wasow (1972) and in another form in Chomsky (1975b,c) is incorrect or at least incomplete, since it would seem that this explanation could not be extended to deletion rules. But if comparatives are formed by *wh*-movement, as suggested above, it follows at once that they should have exactly the cross-over properties of relatives and questions; the proposed explanations would directly cover the cases that Bresnan cites, with no changes. It seems to me fair to take this as an indirect but significant additional argument in favor of the hypothesis that comparatives are formed by *wh*-movement. The argument is, in this case, that under this hypothesis we retain a fairly general, and, I believe, rather convincing explanation for cross-over phenomena.

The cross-over cases that Bresnan cites are (essentially) the following:

(60) a. *more students flunked than–thought they would (flunk)*
      b. *more students flunked than they thought–would (flunk)*

Students is the understood subject of *think* in (a) and *flunk* in (b). But in (a), *they* can refer to the students, whereas in (b) it cannot.

According to a *wh*-movement analysis, the structure of (a) and (b) after *wh*-movement will be approximately (61a), (61b), respectively:

(61) a. *more students flunked than* [[*wh-many* (*students*)] [*t thought* [*they would flunk*]]]
      b. *more students flunked than* [[*wh-many* (*students*)] [*they thought* [*t would flunk*]]]

The structures of (61) are analogous in relevant respects to the direct questions (62a), (62b):

(62) a. *how many* (*students*) [*t thought* [*they would flunk*]]
      b. *how many* (*students*) [*they thought* (*did they think*) [*t would flunk*]]

The analysis proposed in the references cited accounts for all of these cases, in what seems to me a very natural way, on the basis of fairly general principles. It remains to be determined whether all cases of cross-over in comparatives fall so readily under the analysis developed for *wh*-movement.

[. . .]

Let us turn now to another example of a grammatical process that gives the configuration (49), namely, topicalization. To begin with, topicalization does yield this configuration. Thus we have (63):

(63) a.   *this book, I really like*
      b.   *this book, I asked Bill to get his students to read*
      c.   **this book, I accept the argument that John should read*
      d.   **this book, I wonder who read*

Before proposing an analysis of topicalization, let us consider again left-dislocation as in (64) [. . .]):

(64)   *as for this book, I think you should read it*

Plainly in this case, there can be no transformational analysis in our terms since no transformation can "create" the structure "as for this book" or even more complicated phrases that can appear in this position. Suppose, then, that we postulate the base rule R1 in addition to Bresnan's R2, already assumed:

(65)   R1: $\overline{\overline{S}} \to \text{TOP} \, \overline{S}$
       R2: $\overline{S} \to \text{COMP S}$

In addition, we assume the semantic rule of predication already discussed informally in connection with (24) [in the full article].

As Sag observes, structures such as (64) can be embedded, with varying degrees of acceptability, as in (66):

(66)   *I informed the students that as far as this book is concerned, they would definitely have to read it*

To accommodate such cases, let us revise rule R2 to (67):

(67)   R2: $\overline{S} \to \text{COMP} \begin{Bmatrix} \overline{\overline{S}} \\ S \end{Bmatrix}$

These rules will allow recursions, giving such sentences as (68):

(68)   *as for John, as far as this book is concerned, he will definitely have to read it*

If such structures are to be permitted, the rule of predication will have to be extended in an obvious way.

Let us now return to topicalization. Suppose that the analysis is just like left-dislocation, except that in the TOP $\overline{S}$ structure, $\overline{S}$ is a *wh*-clause – in effect, a kind of free relative, as in comparatives. Thus (63b) will derive from (69), which in turn derives from (70):

(69) $[_{\overline{\overline{S}}} \, [_{\text{TOP}} \, this \, book] \, [_{\overline{S}} \, [_{\text{COMP}} \, what] \, [I \, asked \, Bill \, to \, get \, his \, students \, to \, read \, t]]]$
(70)   *this book, I asked Bill to get his students to read what*

To form (63b) from (69) we use the obligatory rule of *wh*-phrase deletion already motivated for comparatives.

On these assumptions, (63b) is analogous to such sentences as (71):

(71)   a.   *this book is what I asked Bill to read*
        b.   *it is this book that I asked Bill to read*

From the point of view of the semantics as well as the syntax, the analogy seems appropriate.

In (69) the rules already discussed introduce a bound variable, giving (72):

(72) $[_{\overline{\overline{S}}} \, [_{\text{TOP}} \, this \, book] \, [_{\overline{S}} \, [_{\text{COMP}} \, what \, x] \, [I \, asked \, Bill \, to \, get \, his \, students \, to \, read \, x]]]$

Deletion of the *wh*-phrase leaves an open sentence, which we may assume to be interpreted by the predication rule that applies in the case of left-dislocation and relatives.

It follows from these assumptions that topicalizations, like left-dislocation, should be possible with varying acceptability within embedded clauses, as in (73):

(73)   *I informed the students that this book, they would definitely have to read*

[It] seems to me that (73) is about on a par with the formally analogous (66).

It also follows that topicalization should have the properties of (49), as was illustrated in (63).

Before we leave this topic, let us consider some further consequences of the analysis. Notice that although topicalization is possible within *that*-clauses, as in (73), it is impossible within relatives or questions. Thus we cannot have (75) corresponding to (74):

(74)   *John gave away the books to some friends*

(75)   a.   *_to whom the books did John give away_ (_to whom did the books John give away_)
        b.   *_whom the books did John give away to_
        c.   *_the boy to whom the books John gave away_
        d.   *_the boy whom the books John gave away to_

The structure underlying, e.g., (75c,d) would on our assumptions be (76):

(76)   _the boy_ [$_{\overline{S}}$ COMP [$_{\overline{\overline{S}}}$ [$_{TOP}$ _the books_] [$_{\overline{S}}$ COMP _John gave away which to whom_]

The structure (76) is generable by the base rules. Furthermore, _wh_-movement can apply to _which_ in the embedded sentence, placing it in the internal COMP position and leaving a trace. If the dominating $\overline{S}$ were within a _that_-clause instead of a relativized NP, we would then derive (77):

(77)   _I believe that the books, John gave away to some friends_

While (77) is not very elegant, it is surely far better than (75c,d), which would derive from (76) by still another application of _wh_-movement, namely to (_to_) _whom_, placing it in the position of the higher COMP.

The problem with (75) does not seem to be just a surface difficulty; compare the sentences (78), which seem much better than (75) and more or less on a par with (77):

(78)   a.   _I believe that this book, you should read_
        b.   _I believe that this book, you should give away_
        c.   _I believe that his friends, John gave some books away to_

We can explain the impossibility of the sentences (75) by essentially the same line of argument that accounts for the _wh_-island constraint. Movement of (_to_) _whom_ to the internal COMP is blocked, because the internal COMP is already filled by _which_ under the _wh_-movement analysis of topicalization. Movement of (_to_) _whom to the higher_ COMP node is impossible because it would violate SSC and PIC (and, if $\overline{S}$ is a cyclic node, subjacency). Even if the already moved _which_ could move by COMP–COMP movement to the higher COMP, freeing the lower one, subsequent movement of (_to_) _whom_ to the lower COMP would be excluded by strict cyclicity. Since the trace left by movement of _which_ is (when replaced by a variable) taken to be satisfied by _the books_ under the predication rule, there is no possible interpretation of (76) or of any of the sentences of (75). Thus there are a number of reasons why (75) are ungrammatical, on the _wh_-movement analysis of topicalization. In effect, we can form (75) only by extraction from a _wh_-island.

[. . .]

Over a considerable range, then, analysis of topicalization as _wh_-movement seems quite reasonable. The proposal is that in the TOPIC position there is a base-generated structure and that the associated proposition, which is an open sentence except for some cases of left-dislocation, says something about it. There

are in principle two ways to derive an open sentence: by *wh*-movement (and *wh*-phrase deletion; but cf. note 25 [in the full article]) or with an uninterpreted pronoun. Both of the available ways are used. The first gives topicalization; the second, left-dislocation.

[. . .]

To summarize, I have suggested that we can eliminate from the grammar rules of comparative deletion, topicalization, clefting, object-deletion and "tough movement," rules for adjective and adjective-qualifier complements, and others, in favor of the general rule of *wh*-movement that also yields direct and indirect questions (finite and infinitival) and finite and infinitival relative clauses, several rather general rules of interpretation, and some language-specific properties of base and surface structures. If this analysis proves tenable, we can drastically reduce the grammatical apparatus for the description of English; but more important, we can drastically limit the class of possible rules. Some curious and otherwise unexplained phenomena fall into place quite naturally, under this simplification of grammatical theory and the description of English. The properties (49), which appear (with the provisos noted) in a wide range of cases, fall together naturally, as a consequence of independent and, I think, rather natural conditions on rules: the subjacency condition, which in effect limits the "memory" available to transformational rules; SSC, which selects a most prominent NP in an embedded cyclic category that is alone accessible to rules if it is present; and PIC, which immunizes a certain category of propositions from rule application, subject to the language-specific proviso that permits COMP-COMP movement over a "bridge." Each of these conditions may be thought of as a limitation on the scope of the processes of mental computation that ultimately determine phonetic and logical form.

[. . .]

6.3   Questions pertaining to Chomsky (1977)

1   There is a sharp contrast in English between *I assure you John to be among the best students in the class* and *(?)John, who I assure you to be among the best students in the class, will arrive shortly*, where *wh*-movement substantially improves acceptability. Discuss from a minimalist perspective the pros and cons of Kayne's (1980) way of accounting for this in terms of Case in Comp.

2   What exactly does the phenomenon of the preceding question have in common with the contrast between *John think those kids should be invited to the party*, with *think* rather than *thinks*, and the much improved (for many speakers) ?*Which kids do John think should be invited to the party?*, again via *wh*-movement? Bring in Kayne (2003) and references cited there.

3   Many French speakers allow, in some cases, quantifier extraction out of a finite clause, as in *Il faut tous qu'ils partent* ('it is-necessary all that they leave'), in apparent violation of the Propositional-island condition. Chomsky considers the possibility of allowing such extractions through the use of a certain complex transformational rule. What are the advantages and disadvantages of his approach as compared with those of Kayne (1981) and Cinque (2002)?

4   In his note 12 Chomsky suggests that resumptive pronouns cannot be introduced by transformation, recalling his more recent (1995: 228) inclusiveness condition. This leaves open the question whether resumptive pronouns could be merged as part of a doubling structure (cf. Kayne 1972, section 3, and Uriagereka 2001) and then stranded by movement of their double (cf. Boeckx 2003). Discuss the relevance to this question, bringing in Kayne (2002b) on Condition C, of the English contrast between the reasonably acceptable *Which rumors$_i$ about a famous politician$_j$ is his$_j$ chief of staff unaware of the fact that they$_i$'re all over the internet?* and the impossible **Which rumors$_i$ about a famous politician$_j$ is he$_j$ unaware of the fact that they$_i$'re all over the internet?*

5   English allows relatives with no visible relative pronoun, for which Chomsky suggests deletion of a *wh*-phrase, noting at the same time that a *wh*-phrase preceded by a preposition cannot be deleted, as seen in *the chair *(on which) we were sitting*. He suggests an account in terms of recoverability of deletion. How might one allow for languages that do have counterparts of **the chair we were sitting*? Bring into the discussion Klein (1995).

6   Rizzi (1997) appears at first glance to constitute a step away from Chomsky's unified approach to *wh*-movement, insofar as Rizzi attributes different landing sites to *wh*-movement in relatives vs. questions, in particular by having *wh*-phrases landing higher in relatives than in questions. What independent difference between relatives and questions might account for this difference in landing site, and how would it? Bring into the discussion Bianchi (1999).

7   Chomsky himself notes a discrepancy between *wh*-movement in relatives and questions, insofar as (headed) relatives seem not to have any counterpart to adjectival questions such as *How tall is she?* But what exactly is it that underlies the absence of **He's not half the courageous that his father was*, as opposed to the acceptable *He doesn't have half the courage that his father had*? Or could there be a sense in which comparatives like *He's not half as courageous as his father was* are actually relatives (above and beyond their both involving *wh*-movement). Evaluate in particular the possibility that such comparatives involve raising of the adjective in a way parallel to the head-raising analysis of relatives developed by Vergnaud (1974). Bring to bear reconstruction effects, as in **We're not half as proud of him$_i$ today as John$_i$ was yesterday* vs. *We're not half as proud of him$_i$ today as his$_i$ wife was yesterday*.

8   Contrary to unacceptable headed relatives like **the courageous that his father was*, headless relatives in English are possible with adjectives, as in *However courageous he is*, . . . As a means of understanding this contrast, evaluate the following proposal: Headless relatives (at least of the *ever*-sort) are actually embedded questions, close in structure to *No matter how courageous he is*, . . . , except that *no matter* is not pronounced and *ever* is.

9   Chomsky takes English topicalization to involve *wh*-movement (with the *wh*-phrase then being obligatorily deleted), while simultaneously taking the initial (non-adverbial) topic phrase not to be moved from any lower position. Yet in English, at least, there are reconstruction effects, as in *His youngest daughter every man is particularly proud of*. Discuss, against the background of Chomsky (1995, Ch. 3), the implications of such reconstruction effects, bringing in Bianchi (1993) on Italian CLLD.

10   Extend the preceding question to clefts.
11   Chomsky takes English *easy-to-please* sentences to involve *wh*-movement, but
     to have the surface subject generated/merged in situ. But again there seem to
     be reconstruction effects, as in the fairly acceptable idiom chunk example
     Tabs are easy to keep on people who are unsuspecting. Assume this example is
     representative and discuss the possible consequences.
12   As Chomsky notes, the involvement of *wh*-movement in *easy-to-please* sen-
     tences is in part supported by the acceptability of long-distance examples like
     *That book would be easy to get John to read, That book would be easy to convince
     John that he should read*. Such examples are not possible, however, in either
     French or Italian. How might one execute the idea that this English vs.
     French/Italian difference is due to French and Italian having pronominal
     clitics? Bring into the discussion Kayne (1989) and Cinque (2002).
13   Chomsky mentions the fact that alongside *John left early but Bill didn't*,
     English also allows *John left early but Bill didn't leave early*, with an unstressed
     leave early. Why might it be that this does not hold of *John didn't break the
     window, Bill did*, which has no counterpart **John didn't break the window, Bill
     did break the window*?
14   Chomsky takes *for*-deletion in *We would like (for) you to help us* to illustrate
     the "deletion of designated terminals." Pick a language other than English
     and find at least ten subtypes of that kind of deletion (or non-pronunciation).
     (Hint: Look at the entry "silence" in the index of Kayne (2005).)
15   Chomsky's discussion of sentences like *Who did they find a picture of?* rests on
     the assumption that in *They found a picture of Mary*, "a picture of Mary" is a
     constituent. If Kayne (2002a) is right to contest that assumption, what will
     some of the consequences be?

References

Akmajian, A. 1975. More evidence for the NP cycle. *Linguistic Inquiry* 6:
     115–129.
Aronoff, M. 1976. *Word Formation in Generative Grammar*. Linguistic Inquiry
     Monograph No. 1. Cambridge, MA: MIT Press.
Bianchi, V. 1993. An empirical contribution to the study of idiomatic expressions.
     Rivista di Linguistica 5: 349–385.
Bianchi, V. 1999. Consequences of Antisymmetry: Headed Relative Clauses.
     Berlin: Mouton de Gruyter.
Boeckx, C. 2003. *Islands and Chains: Resumption as Stranding*. Amsterdam and
     Philadelphia: John Benjamins.
Bresnan, J. 1970. On complementizers: Toward a syntactic theory of complement
     types. *Foundations of Language* 6: 297–321.
Bresnan, J. 1972. *The theory of complementation in English syntax*. Doctoral disser-
     tation, MIT. [Published 1979 by Garland Publishers, New York.]
Bresnan, J. 1973. Syntax of the comparative clause construction in English.
     Linguistic Inquiry 4: 275–343.

Bresnan, J. 1975a. Comparative deletion and constraints on transformations. *Linguistic Analysis* 1: 25–74.

Bresnan, J. 1975b. Transformations and categories in syntax. In R. Butts and J. Hintikka (eds), *Proceedings of the Fifth International Congress of Logic, Methodology, and Philosophy of Science*, London, ON: University of Western Ontario.

Bresnan, J. 1976. On the form and functioning of transformations. *Linguistic Inquiry* 7: 3–40.

Cheng, L. L.-S., and N. Corver (eds), 2006. *Wh-movement: Moving on*, Cambridge, MA: MIT Press.

Chomsky, N. 1955/1975. *The Logical Structure of Linguistic Theory*. New York and London: Plenum Press. [Also 1985, University of Chicago Press.]

Chomsky, N. 1961. On the notion 'rule of grammar'. *Proceedings of the Symposium on the Structure of Language and its Mathematical Aspects*, American Mathematical Society 12: 6–24. [Reprinted in Fodor and Katz (1964).]

Chomsky, N. 1971. *Problems of Knowledge and Freedom*. New York: Pantheon Books.

Chomsky, N. 1972a. *Language and Mind*. New York: Harcourt Brace Jovanovich.

Chomsky, N. 1972b. Some empirical issues in the theory of transformational grammar. In S. Peters (ed.), *Goals of Linguistic Theory*, Englewood Cliffs, NJ: Prentice-Hall, pp. 63–130.

Chomsky, N. 1972c. *Studies on Semantics in Generative Grammar*. The Hague: Mouton.

Chomsky, N. 1973. Conditions on transformations. In S. Anderson and P. Kiparsky (eds), *A Festschrift for Morris Halle*, New York: Holt, Rinehart, and Winston, pp. 232–286.

Chomsky, N. 1974. *The Amherst Lectures*. (Lectures given at the 1974 Linguistic Institute, University of Massachusetts, Amherst, June 1974.) *Documents Linguistiques*, Paris: Université de Paris VII, Département de Recherches Linguistiques.

Chomsky, N. 1975a. Questions of form and interpretation. *Linguistic Analysis* 1: 75–109.

Chomsky, N. 1975b. *Reflections on Language*. New York: Pantheon.

Chomsky, N. 1976. Conditions on rules of grammar. *Linguistic Analysis* 2: 303–351. [Reprinted in R. Cole (ed.), 1977, *Current Issues in Linguistic Theory*, Bloomington: Indiana University Press. Based on a lecture to the Linguistic Institute, South Florida, June 1975.]

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, pp. 1–52.

Chomsky, N., and G. Miller. 1958. Finite state languages. *Information and Control* 1: 91–112. [Also in D. Luce, R. R. Bush, and E. Galanter (eds), 1964, *Readings in Mathematical Psychology*, New York: Wiley.]

Cinque, G. 2002. A note on 'restructuring' and quantifier climbing in French. *Linguistic Inquiry* 33: 617–636. [Reprinted in Cinque (2006, Ch. 4).]

Cinque, G. 2006. Restructuring and Functional Heads: *The Cartography of Syntactic Structures* vol. 4. New York and Oxford: Oxford University Press.

Dougherty, R. C. 1974. The syntax and semantics of *each other* constructions. *Foundations of Language* 12: 1–47.

Emonds, J. E. 1976. *A Transformational Approach to English Syntax: Root, Structure-preserving, and Local Transformations*. New York: Academic Press.

Erteschik, N. 1973. *On the nature of island constraints*. Doctoral dissertation, MIT.

Fiengo, R., and H. Lasnik. 1973. Logical structure of reciprocal sentences. *Foundations of Language* 9: 447–468.

Fiengo, R., and H. Lasnik. 1976. Some issues in the theory of transformations. *Linguistic Inquiry* 7: 82–191.

Fodor, J. D., and J. J. Katz. 1964. *The Structure of Language: Readings in the Philosophy of Language*. Englewood Cliffs, NJ: Prentice-Hall.

Hale, K. 1976. Linguistic autonomy and the linguistics of Carl Voegelin. *Anthropological Linguistics* 18: 120–128.

Halitzky, D. 1975. Left branch S's and NP's in English; a Bar Notation analysis. *Linguistic Analysis* 1: 279–296.

Hankamer, J. 1973. Why there are two than's in English. In C. Corum, T. C. Smith-Stark, and A. Weiser (eds), *Papers from the Ninth Regional Meeting of the Chicago Linguistics Society*, Chicago: University of Chicago, pp. 179–188.

Helke, M. 1971. *The grammar of English reflexives*. Doctoral dissertation, MIT. [Published 1979 by Garland Publishers, New York.]

Hornstein, N. 1975. S and the $\overline{\mathrm{X}}$ convention. *Recherches Linguistiques à Montréal / Montreal Working Papers in Linguistics* 4.

Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jackendoff, R. 1977. *X-Bar Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Kayne, R. S. 1972. Subject inversion in French interrogatives. In J. Casagrande and B. Saciuk (eds), *Generative Studies in Romance Languages*, Rowley, MA: Newbury House, pp. 70–126.

Kayne, R. S. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

Kayne, R. S. 1979. Extensions du liage et du marquage du cas. *Lingvisticae Investigationes* 3: 29–55.

Kayne, R. S. 1980. Extensions of binding and Case-marking. *Linguistic Inquiry* 11: 75–96. [Revised version of Kayne (1979). Reprinted in Kayne (1984, Ch. 1).]

Kayne, R. S. 1981. Binding, quantifiers, clitics, and control. In F. Heny (ed.), *Binding and Filtering*, London: Croom Helm, pp. 191–211.

Kayne, R. S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, R. S. 1989. Null subjects and clitic climbing. In O. Jaeggli and K. J. Safir (eds), *The Null Subject Parameter*, Dordrecht: Kluwer, pp. 239–261. [Reprinted in Kayne (2000, Ch. 4).]

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2002a. On some prepositions that look DP-internal: English *of* and French *de*. *Catalan Journal of Linguistics* 1: 71–115. [Reprinted in Kayne (2005, Ch. 7).]

Kayne, R. S. 2002b. Pronouns and their antecedents. In S.D. Epstein and T.D. Seely (eds), *Derivation and Explanation in the Minimalist Program*, Oxford: Blackwell, pp. 133–166. [Reprinted in Kayne (2005, Ch. 6).]

Kayne, R. S. 2003. Some remarks on agreement and on Heavy-NP-Shift. In M. Ukaji, M. Ike-Uchi, Y. Nishimura, and H. Nakajima (eds), *Current Issues in English Linguistics,* vol. 2 of *Special Publications of the English Linguistic Society of Japan*, Tokyo: Kaitakusha, pp. 67–86. [Reprinted in Kayne (2005, Ch. 11).]

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kim, W. C. 1976. *The theory of anaphora in Korean syntax*. Doctoral dissertation, MIT.

Klein, E. C. 1995. Toward Second Language Acquisition: A Study of Null-prep. Dordrecht: Kluwer.

Lasnik, H. 1976. Remarks on coreference. *Linguistic Analysis* 2: 1–22.

Lasnik, H., and R. Fiengo. 1974. Complement object deletion. *Linguistic Inquiry* 5: 535–571.

Lees, R. B. 1960a. *The Grammar of English Nominalizations*. The Hague: Mouton.

Lees, R. B. 1960b. A multiply ambiguous adjectival construction in English. *Language* 36: 207–221.

Peters, S., and R. Ritchie. 1973. On the generative power of transformational grammars. *Information Sciences* 6: 49–83.

Pollock, J.-Y. 1976. Théorie des traces et syntaxe du français: quelques problèmes. Unpublished mimeo, Université de Paris. [Published 1977 in *Recherches Linguistiques de Vincennes 4*: 209–263.]

Quicoli, C. 1976a. Conditions on clitic movement in Portuguese. *Linguistic Analysis* 2: 199–223.

Quicoli, C. 1976b. Conditions on quantifier movement in French. *Linguistic Inquiry* 7: 583–608.

Quicoli, C. 1980. Clitic movement in French causatives. *Linguistic Analysis* 6: 131–185.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 281–337.

Ross, J. R. 1967. *Constraints on variables in syntax*. Doctoral dissertation, MIT. [Published 1986 as *Infinite Syntax!* by ABLEX, Norwood, NJ.]

Sag, I. 1976a. *Deletion and logical form*. Doctoral dissertation, MIT. [Published 1980 by Garland Publishers, New York.]

Sag, I. 1976b. A logical theory of verb phrase deletion. In S. Mufwene, C. Walker, and S. Steever (eds), *Papers of the Twelfth Regional Meeting of the Chicago Linguistics Society*, Chicago: University of Chicago, pp. 533–547.

Selkirk, E. 1975. Determiners, quantifiers and measure phrases in English. Unpublished ditto, University of Massachusetts, Amherst.

Uriagereka, J. 2001. Doubling and possession. In B. Gerlach and J. Grijzenhout (eds), *Clitics in Phonology, Morphology and Syntax. Linguistik Aktuell/ Linguistics Today 36*, Amsterdam and Philadelphia: John Benjamins, pp. 405–431.

Vergnaud, J.-R. 1974. *French relative clauses*. Doctoral dissertation, MIT. [Revised version in Vergnaud (1985).]

Vergnaud, J.-R. 1985. *Dépendances et niveaux de représentation en syntaxe*. Amsterdam: John Benjamins.

Wasow, T. 1972. *Anaphoric relations in English*. Doctoral dissertation, MIT.

Woisetschläger, E. 1976. Conditions on mixed terms. Unpublished mimeo, MIT.

7

Why Subject Sentences Don't Exist

Jan Koster

1978

7.1 Introduction

The notion of subject has a long history both in linguistic and in philosophical studies, across theoretical frameworks. This is in part because subjects exhibit, with remarkable and intriguing crosslinguistic consistency, unique properties that distinguish them from other arguments of a predicate. From a syntactic point of view, as pointed out in Keenan (1976) and McCloskey (1997), subjects are unique in several respects. For example, subjects of finite clauses can bind reflexives and reciprocals in other argument positions, but cannot themselves be bound by other arguments. Similarly, they can license a Negative Polarity Item (NPI) in other argument positions, but cannot themselves be licensed by appropriate licensers in other argument positions in the same clause. Some languages require that the sole argument of an intransitive predicate be a surface subject, whereas no language seems to impose a similar requirement on other argument positions. Whereas other arguments can be noun phrases or prepositional phrases, subjects are over-whelmingly noun phrases.

In *Why Subject Sentences Don't Exist*, Koster furthers our understanding of the complex nature of subjects by addressing the issue of their predominantly nominal nature, in particular by asking whether a clause interpreted as a subject occupies the same structural position as a nominal subject. Consider an example like (1a), where the clause *that the doctor came* is interpreted as the subject of the sentence:

(1) a. That the doctor came surprised me.
     b. My sister surprised me.

Is the so-called sentential subject in (1a) in the same structural position as the nominal subject *my sister* in (1b)? Koster discusses a number of asymmetries between sentential and nominal subjects (some already mentioned in Rosenbaum 1967, and in Emonds 1970), and reaches the conclusion that the answer is negative.

There exist several contexts where a finite sentential subject is not possible, but a nominal subject is. For example, a finite sentential subject cannot occur in matrix questions following an auxiliary or modal (2a), whereas a nominal subject can (2b):

(2)   a.   *Did *that John showed up* please you?
      b.   Did *that* please you?

A finite sentential subject cannot occur when another constituent has been fronted (e.g., *such things* in (3a)), whereas a nominal subject can, (3b):

(3)   a.   *Such things *that he reads so much* doesn't prove.
      b.   Such things *it* doesn't prove.
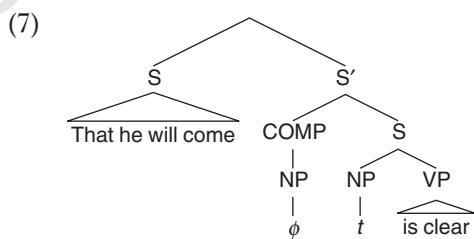
A sentential subject cannot always serve as the subject of a subordinate clause (4a). This restriction does not apply to nominal subjects (4b):

(4)   a.   *Although *that the house is empty* may depress you, it pleases me.
      b.   Although *this outcome* may depress you, it pleases me.

Koster concludes from these asymmetries that a sentential subject does not occupy the same structural position as a nominal subject. Rather, he proposes, it occupies a higher structural position, and it binds a phonetically null nominal element in the canonical subject position. This relation is analogous to the relation between a topic and an overt pronoun within the clause (as in (5)), or a sentence and a coreferential pronoun, as in (6):

(5)   My father, *he* won't come today.
(6)   He will come; *that* is for sure!

Koster suggests that the relation between the sentential subject and the pronominal in subject position is mediated by movement: the pronominal element is said to move from the subject position to COMP, where it gets deleted (this null element is represented as ϕ in (7)); the "satellite" clause binds this null nominal in COMP:

(7)

This proposal accounts for the asymmetries between sentential and nominal subjects: a sentential subject cannot occur in matrix questions following an auxiliary or modal (2a), or in sentences with a fronted constituent (3a) because these elements raise past the canonical subject position, but not past the "satellite" clause position in (7). Finite sentential subjects do not occur in subordinate clauses because of a general restriction that "satellite sentences" cannot occur at the left periphery of an embedded clause.

In addition to accounting for these and other asymmetries, this article is one piece of an important body of work that tries to reach a deeper understanding of the relation between the constituent(s) that we interpret as the subject of a predicate and the structural position that we identify as the subject position. The results of this body of work have been showing that this position is not one where subjects are generated, but one to which subjects are connected either by a movement or by a binding relation (cf. Koopman and Sportiche 1991, Ch. 19 of this volume).

7.2 From "WHY SUBJECT SENTENCES DON't EXIST"

1 The anomalies of extraposition

Since the appearance of Rosenbaum (1967), many linguists believe that (1b) is derived from (1a) by Extraposition:

(1) a. That the doctor came surprised me.
     b. It surprised me that the doctor came.

This rule of Extraposited is very problematic, as pointed out by Emonds (1970, [1972]). The anomalies of Extraposition (partly noticed by Rosenbaum himself) can be summarized as follows:

(i) *While Extraposition is optional in main clauses like (1), it is usually obligatory in subordinate clauses*:

(2) a. *That *for Bill to smoke* bothers the teacher is quite possible.
     b. *Although *that the house is empty* may depress you, it pleases me.

(ii) *Extraposition has to be obligatory after Subject Aux Inversion*:

(3) a. *Did *that John showed up* please you?
     b. *What does *that he will come* prove?

(iii) *Extraposition has to be obligatory after preposings like Topicalization* (cf. (4) and (5)):

(4) a. It doesn't prove such things.
     b. Such things it doesn't prove.

(5) a. *That he reads so much* doesn't prove such things.
    b. *Such things *that he reads so much* doesn't prove.

(iv) *Extraposition gives the wrong result with bisentential verbs like* prove, imply, *etc.*:

(6) a. That John has blood on his hands proves that Mary is innocent.
    b. *It proves that Mary is innocent that John has blood on his hands.

These four anomalies indicate that there is something wrong with Extraposition. They have received relatively little attention, probably because of the appeal of Ross's Internal S Condition:

(7) *Internal S Condition*
    Grammatical sentences containing an internal NP which exhaustively dominates S are unacceptable.

This ad hoc principle, which seems to have some plausibility in terms of performance, reduces the first three anomalies to a common factor. Unfortunately the principle does not work, being both too strong and too weak. Kuno (1973) gives examples like (8), where the Ss (in italics) are clearly internal:

(8) a. Believing *that the grapes are sour* gives one some solace.
    b. John proved *that the earth is round* when he was fifteen.

Higgins (1973) has given examples in which the unextraposed subject sentences are not internal at all:

(9) *How likely is *that John will come*?

Kuno makes a generalization which comes very close to the truth:

(10) Subject sentences can only appear in sentence initial position.

This "constraint", as Kuno calls it, is the very fact that I would like to explain in this article.
    [. . .]

4 The satellite hypothesis

My proposal implies that subject sentences do not exist. Instead there are satellite sentences binding the (phonologically zero) NP subject of the main sentence. In this conception, the main sentence is taken as an open sentence, satisfied by the satellite. Thus, the traditional structure (18) has to be given up in favor of (19) (where the arrow indicates the binding relation, and [$_{NP}$e] is an empty NP):

(18) [$_{NP}$S]   VP   (19) S   [$_{NP}$e]   VP
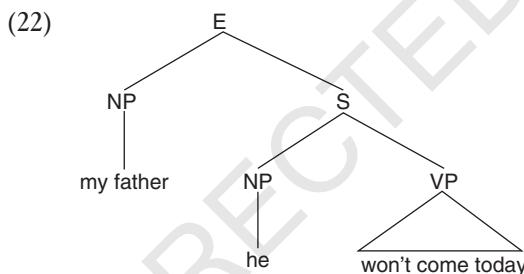
This solution fits in a framework that is needed for independent reasons. In perhaps all languages a root sentence can be preceded by material introducing the topic of the sentence. For instance, sentence (20) contains an open sentence about (the introduced topic) *my father*:

(20)   My father, *he* won't come today.

The subject of the sentence (*he*) is bound by *my father*. Such sentences are often considered a result of a movement rule called Left Dislocation (cf. Ross (1967)). But Van Riemsdijk and Zwarts (1974) have argued that Left Dislocation is not the most general solution for such sentences. They give several arguments to show that an extension of the base rules is more appropriate. The "dislocated" NP (*my father* in (20)) is introduced as a satellite of the main sentence:

(21) $E \rightarrow NP \; \bar{S}$

In this rule, E (= expression) is a symbol that dominates the satellite NP together with the root sentence:

(22)

Such satellite structures occur in most (and possibly in all) languages. A common feature of satellites is the obligatory binding of an NP in the main sentence. This NP is often fronted to a position next to the satellite. In English, this is usually a *wh*-word, which is obligatorily deleted in the case of Topicalization. According to Chomsky (1977), a sentence like (23) is derived as in (24):

(23)   This book, I asked Bill to read.
(24)   a. $[_E[_{NP}$ this book] $[_{\bar{S}}$ $[_{COMP}$what] [I asked Bill to read t]]]
        b. $[_E[_{NP}$ this book] $[_{\bar{S}}$ $[_{COMP}\phi]$ [I asked Bill to read t]]]

The *wh*-word is moved to the COMP position of the root sentence, where it is obligatorily deleted. Obligatory deletion of *wh*-words in COMP position has already been motivated for other constructions like comparatives. Something similar happens in Relative Clause Formation in English, where the *wh*-phrase is

obligatorily fronted to the COMP adjacent to the antecedent. In this case, deletion of the *wh*-phrase is optional:

(25) a.   the book (which) I read . . .
      b.   the book [$_\bar{S}$ [$_{COMP}$which] [I read t]]
      c.   the book [$_\bar{S}$ [$_{COMP}$$\phi$] [I read t]]

The options in Dutch are somewhat different. Relative clauses like the one in (25) are derived by movement of a so-called *d*-word which cannot be deleted:

(26) a.   de man die ik ken . . .
            the man who I know . . .
      b.   *de man ϕ ik ken

But contrary to English, the pronoun *die* is optional in the case of Topicalization:

(27) a.   Die man die ken   ik.
            that man that know I
            'That man I know.'
      b.   Die man ϕ ken ik.

Constructions involving Topicalization generally have an optional pronoun in Dutch. Thus, other major categories like APs (28) or PPs (29) also appear as satellites, binding an optional pronoun (*dat* for adjectives, *daar* for locative PPs):

(28) a.   Knap, *dat* is ze zeker.
            clever that is she certainly
            'Clever she certainly is.'
      b.   Knap ϕ is ze zeker.

(29) a.   In Amsterdam, *daar* logeerde hij niet.
            in Amsterdam, there stayed   he not
            'In Amsterdam, he didn't stay.'
      b.   In Amsterdam, ϕ logeerde hij niet.

It doesn't come as a surprise, therefore, that a fourth major category, $\bar{S}$, also occurs as a pronoun-binding satellite:

(30) Dat hij komt, *dat* is duidelijk.
      that he comes that is clear
      'That he will come is clear.'

Again, the pronoun is optional. If we drop it, the result contains an apparent subject sentence:

(31) Dat hij komt, ϕ is duidelijk.

Recall now that deletion of the *wh*-phrase is obligatory with Topicalization in English. The only possibility is (32):

(32)   That he will come ϕ is clear.

To derive this sentence, we need *Wh* Movement together with the following rules:

(33)   *Base Rule*

$$E \rightarrow \left(\left\{ \begin{matrix} NP \\ \overline{S} \end{matrix} \right\} \right) \quad \overline{S}$$

(34)   *Deletion in COMP*

$$[_{COMP} \ . \ . \ X \ . \ .] \Rightarrow [_{COMP} \ . \ . \ \phi \ . \ .]$$
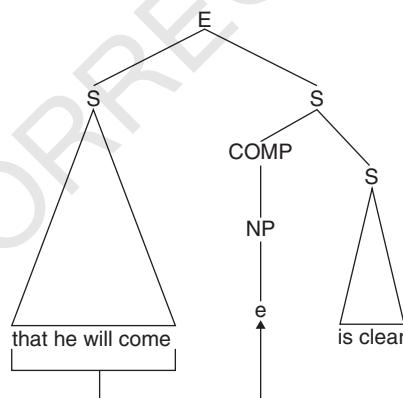
Rule (34) is a very general rule deleting material in COMP. Optional or obligatory application differs from construction to construction and depends on a system of filters as proposed by Chomsky. In the case of Topicalization, rule (34) is optional for Dutch and obligatory for English. With these rules, sentences like (32) are derived in the following way:

(35)   a. $[_E[_{\overline{S}}$ that he will come$]\,[_{\overline{S}}\,[_{COMP}[\qquad\qquad]]\,[[_{NP} \ . \ . \ wh \ . \ .]$ is clear$]]]$
      b. $[_E[_{\overline{S}}$ that he will come$]\,[_{\overline{S}}\,[_{COMP}[_{NP} \ . \ . \ wh \ . \ .]]\,[[_{NP}t]$ is clear$]]]$
      c. $[_E[_{\overline{S}}$ that he will come$]\,[_{\overline{S}}\,[_{COMP}[_{NP}\phi]]\,[[_{NP}t]$ is clear$]]]$
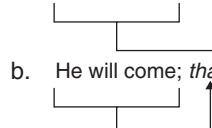
Thus, so-called subject sentences are considered satellite sentences binding an adjacent NP (in COMP):

(36)
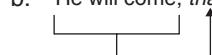
A sentence binding an NP is quite normal; this was already clear from the Dutch example (30). There are also obvious cases in English like (37a, b):

(37)   a.   He will come, *which* we regret.

      b.   He will come; *that* is for sure!

In the present analysis, so-called subject sentences are indistinguishable from base-generated topics. Object sentences can be treated in the same way. Again, there is an optional pronoun in Dutch (38a), and an obligatorily deleted *wh*-phrase in English (39).

(38) a.   Dat hij komt, *dat*   betreur ik t.
           that he comes that regret   I   t
           'That he will come I regret.'
      b.   Dat hij komt, ϕ betreur ik.
(39) That he will come ϕ I regret t.

5   The anomalies solved

The anomalies of Extraposition can now be solved under the following assumptions:

(40) a.   There is no rule of Extraposition;
      b. $\bar{S}$ s are possible satellites;
      c.   "Extraposed" sentences are generated in the base at the end of the VP
           (VP $\rightarrow$ . . . V . . S . .).

The first anomaly was connected with the ungrammaticality of sentences like (2a,b), repeated here:

(2)   a.   *That *for Bill to smoke* bothers the teacher is quite possible.
      b.   *Although *that the house is empty* may depress you, it pleases me.

These sentences are generated under my proposal, because the old subject sentences are now satellites, which are daughters of E. E is never embedded in Dutch, and only under very restricted conditions in English.

The second anomaly concerned the ungrammatical results of Subject Aux Inversion:

(3)   a.   *Did *that John showed up* please you?
      b.   *What does *that he will come* prove?

These sentences are not generated, because Subject Aux Inversion involves the NP subject of the main sentence $\bar{S}$, and not the $\bar{S}$ satellite which is a daughter of the higher E.

The same is true for the third anomaly, because Topicalization brings an NP in front of the main sentence $\bar{S}$ and not in front of E as would be the case in (5b):

(5)   b.   *Such things *that he reads so much* doesn't prove.

The fourth anomaly disappears, too. The double S problem at the end of VP does not arise, since the satellite $\overline{S}$ stays where it is (6a) and cannot be extraposed (6b):

(6) a.   That John has blood on his hands proves that Mary is innocent.
     b.   *It proves that Mary is innocent that John has blood on his hands.

What the four anomalies of Extraposition had in common was that the front position was the only position where subject sentences could occur. This fact is naturally accounted for by extending the class of front items (satellites) with the category $\overline{S}$. In this way the satellite hypothesis fits the facts, where the Extraposition hypothesis does not. But there are more advantages.

One of the most important points in favor of the present analysis is that we can maintain Emonds's idea that complement Ss are not dominated by NP, so that we can do without dubious PS rules like NP→S. If sentences dominated by the NP subject did exist, the Passive transformation would apply to them, and (41) would be predicted to be grammatical:

(41) a.   That the children are always late shows the necessity of discipline.
      b.   *The necessity of discipline is shown by that the children are always late.

The satellite hypothesis, on the other hand, correctly predicts the ungrammaticality of such sentences.

A last advantage of the satellite hypothesis is that is makes unnecessary the Sentential Subject Constraint as formulated by Ross (1967):

(42) *The Sentential Subject Constraint*
      No element dominated by an S may be moved out of that S if that node S is dominated by an NP which itself is immediately dominated by S.

This is an ad hoc statement which we do not need under the satellite hypothesis. Consider the following ungrammatical sentence (an apparent violation of (42)):

(43)   *What did [that John saw t] surprise Mary.

If "subject sentences" are satellites, this sentence is automatically excluded. Subject Aux Inversion cannot apply, and no element can be extracted from satellites to the front of E. The latter category has no initial COMP.

Merely fitting the facts is not the only thing we expect from a new hypothesis. Ideally, it should lead to new knowledge. This is largely a matter of further research, but what follows might be a case in point. Probably we can derive some new predictions from the special status of the front $\overline{S}$ as a satellite. Consider the following facts from Kuno (1973):

(44) a.   What is important?   Love is.
      b.   What is important?   *That we work harder is.

Kuno comments: "For a reason that is totally mysterious to me, copulative sentences with NP clause subjects cannot undergo deletion of postcopular elements." The satellite hypothesis may help to solve Kuno's mystery, because the observed deletions are generally incompatible with satellites:

(45) a. Who is nice?   John is.
     b. Who is nice?   *John, he is.

Therefore, not only does the satellite hypothesis fit the facts, where the Extraposition hypothesis does not; it may also lead to new insight in unexpected areas.

7.3 Questions pertaining to Koster (1976)

1 Koster's proposal that sentential subjects are not in ordinary subject position looks correct for sentential subjects whose main verb is finite (tensed). How does it fare with other kinds of sentential subjects?

2 Koster's proposal postulates a close link between *That Mary is smart is obvious* and *That Mary is smart, that's obvious*. Can you find any differences between these two kinds of sentences (in English or in any other language that has both)?

3 To an English speaker, there is an intonational difference (indicated by the comma in the first and its necessary absence in the second) between *That person, everybody likes* and *That person likes everybody*. To what extent does this bear on Koster's proposal?

4 *It's unfortunate that they said that* is readily pronounced with no intonation break, as opposed to *It's unfortunate, their having said that*. What other kinds of facts might this be related to?

5 How would you fit Koster's analysis into Rizzi's (1997) framework?

6 *That Mary is smart is obvious* resembles *The fact that Mary is smart is obvious*, yet Koster does not extend his analysis of the former to the latter. Was he right to keep them separate? Give your reasons.

7 Sentential subjects, according to Koster, are actually in a left-dislocated position, and not in ordinary subject position. Is this likely to be related to the fact that (standard) English disallows (finite) sentential objects of prepositions, as in *We're counting on that he'll win the race*? Why or why not? (Extra credit: To what extent are other Germanic languages like English here?)

8 Are extraposed sentential subjects in English in the same position as heavy-NP-shifted objects? Same question for Dutch and German. In what way are the facts of Dutch and German relevant to the proper analysis of English?

9 In English, there's a sharp difference between *It's important to everybody that you arrive on time* and *It's important that you arrive on time to everybody*. How closely is this related to Koster's discussion? Is Stowell's (1981) Case-resistance hypothesis relevant?

10 In *It was pointed out to John that he was late*, the presence of *it* is obligatory, as shown in **Was pointed out to John that he was late*. In the corresponding active sentence, on the other hand, it is not necessary and even less natural than its absence – *They pointed (?it) out to John that he was late*. Why might this be? (Extra credit: Are EPP-features, as in Chomsky 2001, the answer, or at least part of the answer? Why?)

11 Koster notes that Dutch has no embedded root phenomena of the sort seen in English in sentences like *We know perfectly well that at no time have you been willing to help us*. Do other Germanic languages pattern here with English or with Dutch? Give relevant examples. (Extra credit: How many other Germanic languages do you think there are? Hint: Take a look at Kayne 1996, Zwart 1996, van Craenenbroeck 2010.)

12 Koster observes that his proposal makes it unnecessary to have structures in which NP exhaustively dominates S. How is this point related to X-bar theory? And to bare phrase structure?

13 (Extra credit) In Paduan, a north Italian dialect that has subject clitics, a subject clitic is normally unnecessary if the sentence has a lexical DP subject. Yet when there's a nonrestrictive relative clause as part of the subject DP the presence of a subject clitic in addition to the DP becomes necessary. To what extent might this be related to Koster's discussion? (Lots of extra credit: To what extent might Koster's discussion be relevant to the fact that Hungarian focused DPs are not allowed to have a post-nominal relative clause? To what extent is Kayne 2000a relevant?)

14 Koster notes that the impossibility of extraction in sentences like **What did that John saw surprise Mary?* fits in straightforwardly with his analysis. Do you think that the same holds for the impossibility of **What did his seeing surprise Mary?* (vs. the well-formed *His seeing the mountain surprised her*)? Give your reasons.

15 English allows *That fact that it's important that we solve the problem is obvious*. It also allows *Its importance is obvious*. But it doesn't allow **Its importance that we solve the problem is obvious*. Why might that be?

16 In addition to sentential extraposition of the sort Koster discusses, English allows relative clause extraposition, as in *Somebody arrived who we used to know a long time ago*. Even possible, in colloquial English, is *Somebody's book just came out who we used to know a long time ago*. Not possible, on the other hand, is **Somebody's book who we used to know a long time ago just came out*. Do you think these facts are related to those of the previous question? Give your reasons. (Extra credit: Propose an account of this restriction on extraposition.)

17 In addition to the restriction illustrated in question 15, English also disallows **Its importance is obvious that we solve the problem*. This last restriction contrasts with the relative acceptability of the example from question 16, namely *Somebody's book just came out who we used to know a long time ago*. What might be a relevant factor, or factors?

18 English allows both *It seems clear that you're right* and *It seems that you're right*. Yet alongside *That you're right seems clear*, English does not allow **That you're

right seems. How might Koster try to relate this contrast to the difference between *That seems clear* and *That seems?

19   What expectations does Koster's analysis lead to concerning the question of *It's clear that you're right* vs. **What's clear that you're right*?

20   (Extra credit) French generally requires a preposition (*de*) to precede an extraposed infinitive, as in *Il est important de dire la vérité* ('it is important of tell the truth'). Without *de*, the sentence would be unacceptable – **Il est important dire la vérité*. Italian, on the other hand, does not require a preposition – *importante dire la verità*. What well-known difference between French and Italian might that be related to, and how?

21   (Lots more extra credit) French and Italian also differ in that in sentences corresponding to *We believe it important to tell the truth*, French has a preposition (*Nous croyons important de dire la vérité*) and Italian doesn't (*Noi riteniamo importante dire la verità*). How might this French–Italian contrast be related to the one in question 20?

References

Chomsky, N. 1977. On Wh-movement. In P. Culicover, T. Wasow, and A. Akmajian (eds), *Formal Syntax*, New York: Academic Press, pp. 71–132.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, pp. 1–52.

van Craenenbroeck, J. 2010. *The Syntax of Ellipsis: Evidence from Dutch Dialects*. New York and Oxford: Oxford University Press.

Emonds, J. 1970. *Root and structure-preserving transformations*. Doctoral dissertation, MIT.

Emonds, J. 1972. A reformulation of certain syntactic transformations. In S. Peters (ed.), *Goals of Linguistic Theory*. Englewood Cliffs, NJ: Prentice-Hall, pp. 21–62.

Higgins, R. 1973. On J. Emonds's analysis of extraposition. In J. Kimball (ed.), *Syntax and Semantics*, vol. 2, New York: Seminar Press, pp. 149–195.

Kayne, R. S. 1996. Microparametric syntax. Some introductory remarks. In J. R. Black and V. Motapanyane (eds), *Microparametric Syntax and Dialect Variation*, Amsterdam and Philadelphia: John Benjamins, pp. ix–xviii. [Reprinted in Kayne (2000b, Ch. 1).]

Kayne, R. S. 2000a. A note on prepositions, complementizers, and word order universals. In Kayne (2000b, Ch.15).

Kayne, R. S. 2000b. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Keenan, E. 1976. Toward a universal definition of subject. In C. N. Li (ed.), *Subject and Topic*, San Diego and New York: Academic Press, pp. 303–334.

Koopman, H., and D. Sportiche. 1991. On the position of subjects. *Lingua* 85: 211–258.

Kuno, S. 1973. Constraints on internal clauses and sentential subjects. *Linguistic Inquiry* 4: 363–385.

McCloskey, J. 1997. Subjecthood and subject positions. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 197–235.

van Riemsdijk, H., and F. Zwarts. 1974. Left dislocation and the status of copying rules. Unpublished paper, MIT/University of Amsterdam.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 281–337.

Rosenbaum, P. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: MIT Press.

Ross, J. R. 1967. *Constraints on variables in syntax*. Doctoral dissertation, MIT. [Published 1986 as *Infinite Syntax!* by ABLEX, Norwood, NJ.]

Stowell, T. 1981. *Origins of phrase structure*. Doctoral dissertation, MIT.

Zwart, J.-W. 1996. Verb clusters in Continental West Germanic dialects. In J. R. Black and V. Motapanyane (eds), *Microparametric Syntax and Dialect Variation*, Amsterdam and Philadelphia: John Benjamins, pp. 229–258.

8

Violations of the Wh Island Constraint in Italian and the Subjacency Condition

Luigi Rizzi

1980

8.1   Introduction

One of the central concerns in syntax is accounting for both the ways in which languages are similar and the ways in which they differ. This paper shows, within a specific empirical domain, that both goals can be achieved if our model of grammar contains general principles coupled with parameters, that is, points on which languages can differ.

The starting point of *Violations of the Wh Island Constraint* is Chomsky's (1973) proposal that *wh*-movement is constrained by a locality condition, Subjacency. This condition states that a single step of movement cannot cross more than one cyclic (or bounding) node, which Chomsky defined as S and NP. This implies that when a *wh*-phrase moves long distance, it must do so in a stepwise fashion, crossing only one S (or NP) node at a time. *Wh*-movement in (1a), for example, is grammatical because it respects Subjacency, by moving step-by-step: the *wh*-phrase *what* crosses the embedded S node, lands in a COMP position to its left, and then moves further, crossing the matrix S node, as indicated in (1b):

(1)   a.   What do you think she saw?
      b.   what$_i$ do [$_S$ you think [e$_i$ [$_S$ she saw t$_i$]]]

Assuming that there is only one COMP position per clause, Subjacency also provides an account for the ungrammaticality of sentences like (2a), in conjunction with the Strict Cycle Condition. If *where* moves into the embedded

COMP position first, *who* has to move to the matrix COMP position in one fell swoop (2b), violating Subjacency:

(2) a.   *Who was she wondering where they should seat?
     b.   who$_i$ was [$_S$ she wondering [where$_j$ [$_S$ they should seat t$_i$ t$_j$]]]

     c.   who$_i$ was [$_S$ she wondering [t$_i$ [$_S$ they should seat t$_i$ where]]]

If, on the other hand, *who* were to move into the lower COMP position first, and then into the matrix COMP position, as indicated in (2c), subsequent movement of *where* into the embedded COMP would violate the STRICT CYCLE CONDITION (Chomsky 1973: 243). Informally put, this condition states that once a rule has applied to a larger domain (the matrix S in this case), it cannot then apply to a smaller domain (the embedded S).

Assuming that Subjacency is a general principle that constrains *wh*-movement across languages, Rizzi makes an observation that is at first sight puzzling. Italian is like English in being subject to Subjacency (as seen in cases of extraction from so-called "complex NPs"). Yet, unlike English, Italian allows a relative pronoun to be extracted from an embedded indirect question (cf. (3a) versus (3b)). This is unexpected because the *wh*-phrase *a cui*, like its English counterpart *to whom*, crosses two S nodes, as shown in (4):

(3) a.   Tuo fratello, a cui mi domando che storie abbiano raccontato, . . .
     b.   *Your brother, to whom I wonder which stories they told, . . .

(4) a.   [a cui$_i$ [$_S$ mi domando [$_{S'}$ che storie [$_S$ abbiano raccontato t$_i$ . . .]]]]

     b.   *[to whom$_i$ [$_S$ I wonder [$_{S'}$ which stories [$_S$ they told t$_i$ . . .]]]]

Also replace '…' by ' …', as in other examples.

On the basis of such contrasts, Rizzi argues that there is a parametric difference between English and Italian concerning what counts as a bounding node: NP and S in English, NP and S′ in Italian. Sentence (3a) is grammatical because, given the Italian setting of the parameter, the relative pronoun *a cui* crosses only one bounding node, S′ (cf. (4a)). However, (3b) is ungrammatical because, given the English setting of the parameter, the relative pronoun crosses two bounding nodes, S (cf. (4b)). Postulating this simple difference can provide a unified account for both the similarities and the differences between languages like English and Italian within this empirical domain.

The idea that differences in extraction possibilities might be attributed to a parametric difference in what counts as a bounding node (S′ as opposed to S) was discussed in several subsequent pieces of work (e.g., Sportiche 1981 for French; Engdahl 1980 for Swedish; Torrego 1984 for Spanish, among others), and critically examined in Maling and Zaenen (1982) and Grimshaw (1986). Reinhart (1981) tackled the same set of issues (namely, variation in extraction from *wh*-islands) starting from Hebrew, and explored the idea that languages may vary in their phrase structure (some allowing two COMP positions per

clause, some only one); in her view, parametric differences can be found both in phrase structure and in the specification of the universal conditions on movement rules. Much subsequent work has also pursued a better understanding of the notion of parameter, investigating where the grammar can express parametric differences (cf. Borer 1984; Kayne 2005b, 2006, to appear; Rizzi [2011]; among others) and the possibility of distinguishing between micro- and macro-parameters (cf. Baker 2001).

8.2   From "VIOLATIONS OF THE WH ISLAND CONSTRAINT IN ITALIAN AND THE SUBJACENCY CONDITION"

0.   Ross (1967) noticed that a clause introduced by a wh pronoun is an island. This fact, generally accounted for by means of a primitive constraint (the wh island constraint), can be explained by more general principles in the framework recently developed by Chomsky.[1] In this framework, extraction from a clause is blocked by the following general conditions on rule application:

a.   Tensed S Condition (TSC): no rule can involve two elements X and Y in (1) if m is a tensed sentence.
b.   Specified Subject Condition (SSC): no rule can involve two elements X and Y in (1) if m has a subject distinct from Y and not controlled by X.
c.   Subjacency Condition: no rule can involve X and Y in (2) if both m and n are cyclic nodes.

(1) $\ldots X \ldots [_m \ldots Y \ldots] \ldots X \ldots$
(2) $\ldots X \ldots [_m \ldots [_n \ldots Y \ldots] \ldots] \ldots X \ldots$

Both TSC and SSC are superficially falsified by acceptable English sentences like the following (3), in which a wh pronoun originating in a non-subject position of an embedded tensed sentence has been moved to the main complementizer:

(3)   Who do you think Bill saw _____?

This fact is accounted for, in the references quoted in footnote 1, by specifying in the grammar of English a restricted class of language-specific exceptions to SSC and TSC. That is to say, the following proviso on SSC and TSC is specified in the grammar of English:
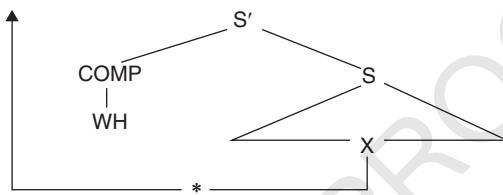
(4)   . . . unless X is in COMP and Y is in COMP.

Informally speaking, (4) means that, in English, the extraction of an element from a clause can violate TSC and/or SSC just in case the element is moved from a COMP position to a higher COMP position.

In this framework, the wh island constraint follows from the general principles just mentioned plus a plausible assumption on the syntax of the complementizer. Consider the abstract structure of a wh island: an element X cannot be extracted out of (5) because:

(1)   it cannot be moved directly to a position outside S′ because of TSC and SSC;

(2)   it cannot be moved to the COMP position (the 'escape hatch') because this position is already filled (under the assumption that a COMP cannot be filled more than once).

(5)

This analysis has been repeatedly proposed in the current literature, and I do not want to discuss it again in detail. The only point I would like to stress here is that this 'explanation' doesn't predict that the wh island constraint is universal. In fact, it makes crucial reference to the 'escape hatch' (4), which is supposed to be specific to English, at least in the framework I am adopting here (cfr. Chomsky (1977), p. 85). Given the assumption that TSC and SSC admit language-specific exceptions ('escape hatches'), as far as I can see nothing in this framework can exclude the existence of a grammar whose specific 'escape hatch' allows extraction from a wh clause.

But this apparent weakness of the theory turns out to be empirically justified: it is a fact that in several natural languages the wh island constraint can be freely violated.

1.   One of these languages is Italian. In Italian a relative pronoun can be extracted from an embedded indirect question:

(6)   a.   Il solo incarico che non sapevi a chi avrebbero affidato è poi finito proprio a te

         (The only charge that you didn't know to whom they would entrust has been entrusted exactly to you)

      b.   Tuo fratello, a cui mi domando che storie abbiano raccontato, era molto preoccupato

         (Your brother, to whom I wonder which stories they told, was very troubled)

      c.   La nuova idea di Giorgio, di cui immagino che cosa pensi, diverrà presto di pubblico dominio

         (Giorgio's new idea, of which I imagine what you think, will soon become known to everybody)

         [. . .]
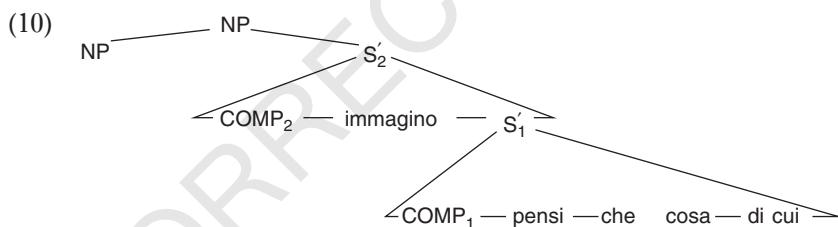
Notice that the process of relative clause formation in Italian crucially involves the movement of the relativized element: no resumptive pronoun is used in the construction, and the complex NP constraint is respected:

(9)   a.   *Questo incarico, che non sapevo la novità che avrebbero affidato a te, . . .
          (This task, that I didn't know the news that they would entrust to you, . . .)
     b.   *Tuo fratello, a cui temo la possibilità che abbiano raccontato tutto, . . .
          (Your brother, to whom I am afraid of the possibility that they told everything, . . .)
     c.   *La nuova idea di Giorgio, di cui immagino facilmente l'eventualità che Piero pensi male, . . .
          (Giorgio's new idea, of which I easily imagine the event that Piero has a bad opinion, . . .)

The unacceptability of (9) is quite relevant: if sentences (9) were acceptable, the acceptability of (6) would turn out to be of no particular interest. The construction would simply be unconstrained [. . .]. But since sentences (9) are unacceptable, an explanation is required for the asymmetry between the wh island constraint and the complex NP constraint shown in (6)–(9).

2.   If we extend to Italian the system proposed in the references of footnote 1, the derivation of sentences (6) is possible only at the cost of violating some conditions on rules, or significantly changing the syntax of the complementizer. Consider, for instance, the plausible base structure of (6) (c):

(10)

Sentence (6) (c) can be derived:

(A)   By violating the strict cyclicity condition: *di cui* is moved into $COMP_1$ at the cycle S'_1, then to $COMP_2$ at the cycle S'_2, and then *che cosa* is moved into $COMP_1$.
(B)   By violating both TSC and SSC: *che cosa* is moved into $COMP_1$ at the first cycle, and *di cui* is moved into $COMP_2$ directly from its base position at the second cycle.
(C)   By allowing a COMP to be filled more than once per cycle: *che cosa* and *di cui* are both moved into $COMP_1$ at the first cycle, and then *di cui* is moved to $COMP_2$ at the second cycle.

Solution (A) seems to me highly suspect: allowing a grammar to violate the strict cyclicity condition amounts to emptying the construct 'transformational

cycle' of any empirical content. Therefore, as long as this construct is believed to play any role in linguistic theory, solution (A) doesn't seem to be a viable one.
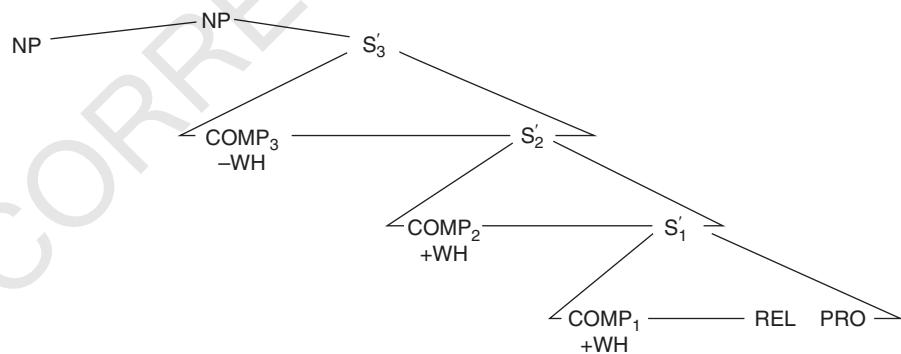
Solution (B) could appear, at first glance, to have more disruptive consequences than solution (C) for the theory I am adopting here. But upon careful consideration this impression turns out to be incorrect. Recall that, under the assumption of footnote 2, the explanation of the wh island constraint in English crucially depends on the language-specific definition of the possible exceptions to SSC and TSC. Therefore, nothing prevents the construction of a grammar whose specific 'escape hatch' permits extraction from a wh clause. For instance, the following rough formulation would give the desired result for Italian:

(11)   Movement into COMP is not constrained by SSC and TSC.

As far as I can see, there is no theoretical reason why (B) would be more costly than (C) (but see paragraph 6): both solutions require some specific hypotheses on the grammar of Italian, and in no sense is the second a priori simpler than the first.

3. In this paragraph I will argue on empirical grounds for solution (B). With respect to structures like (10), solutions (A), (B), (C) seem to be extentionally equivalent; but if we consider more complicated structures, their predictions diverge interestingly. The following phrase marker roughly represents the relevant case:

(12)

In this structure, a relative clause (S'_3) contains an indirect question (S'_4), which in turn contains an indirect question (S'_t). The relative pronoun is base generated in the most deeply embedded indirect question. It is easy to verify that, according to solutions (A) and (C), an acceptable sentence should be derivable from a base structure like (12). According to solution (A), the relative pronoun can climb up cycle after cycle, and then the interrogative complementizers can be

filled by the respective interrogative pronouns. According to solution (C), the relative pronoun can climb up with the respective interrogative pronouns on cycles S_1' and S_2', and then it can move alone to the relative complementizer at the S_3' cycle.

On the contrary, solution (B) predicts that no grammatical surface structure can be derived from (12). Putting aside the derivations which are ruled out via strict cyclicity, only one possibility is left: that $COMP_1$ and $COMP_2$ are filled at the respective cycles by the respective interrogative pronouns, and then, at the S_3' cycle, the relative pronoun is moved directly from its base position to the relative COMP; but such a movement is blocked by the subjacency condition, since the source position and the target position are separated by two cyclic boundaries.

The systematic unacceptability of the second sentences of the following pairs shows that the prediction of solution (B) is correct, and therefore this approach is to be preferred over solutions (A) and (C) on empirical grounds:
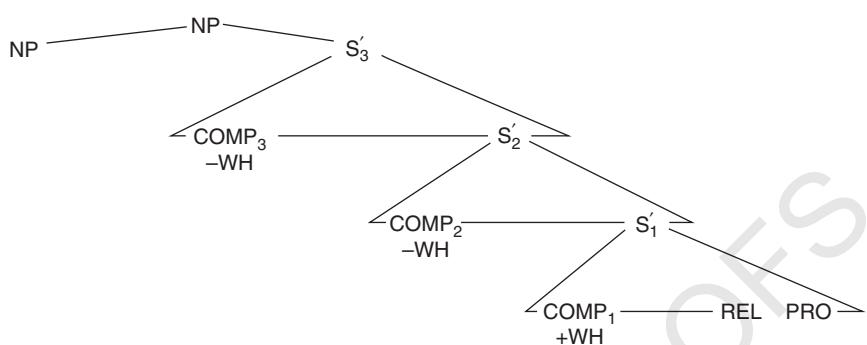
(13) a. Non so proprio chi possa avere indovinato a chi affiderò questo incarico
          (I really sdon't know who might have guessed to whom I will entrust this task)
      b. *Questo incarico, che non so proprio chi possa avere indovinato a chi affiderò, mi sta creando un sacco di grattacapi
          (This task, that I really don't know who might have guessed to whom I will entrust, is getting me into trouble)

(14) a. Non immagino quanta gente sappia dove hanno mandato Francesca
          (I do not imagine how many people know where they have sent Francesca)
      b. *Francesca, che non immagino quanta gente sappia dove hanno mandato, il giorno della partenza era disperata
          (Francesca, that I don't imagine how many people know where they have sent, was in despair)

(15) a. Mi sto domandando a chi potrei chiedere quando dovrò parlare di questo argomento
          (I am wondering whom I could ask when I will have to talk about this topic)
      b. *Questo argomento, di cui mi sto domandando a chi potrei chiedere quando dovrò parlare, mi sembra sempre più complicato
          (This topic, of which I am wondering whom I could ask when I will have to talk, seems to me more and more complicated)

[. . .]
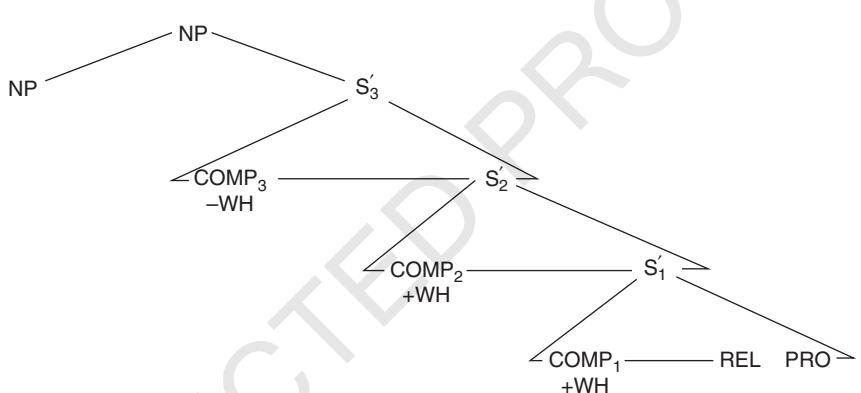
There is an even more interesting class of structures on which solutions (A), (B), (C) make empirically distinguishable predictions; this class of structures can be roughly represented in the following way:

(17)

a.

NP — NP — S'_3

COMP$_3$ —WH — S'_2

COMP$_2$ —WH — S'_1

COMP$_1$ +WH — REL PRO

b.

NP — NP — S'_3

COMP$_3$ —WH — S'_2

COMP$_2$ +WH — S'_1

COMP$_1$ +WH — REL PRO

In (17) (a), S'_3 is a relative clause, S'_2 is a declarative clause, S'_1 is an interrogative clause; in (17) (b), S'_3 is a relative clause, S'_2 is an interrogative clause, S'_1 is a declarative clause. Structures (17) (a) and (17) (b) minimally differ in the respective embeddings of the declarative clause and the interrogative clause.

It is easy to ascertain that, as before, solution (A) and solution (C) predict that an acceptable sentence can be derived in both cases. On the contrary, solution (B) predicts that an acceptable sentence can be derived from (17) (a), but not from (17) (b). The possible derivation from (17) (a) goes as follows: at the first cycle an interrogative pronoun is moved into COMP$_1$, the relative pronoun is moved directly into COMP$_2$ at the second cycle, and then into the relative COMP at the third cycle. But no well formed surface structure can be derived from (17) (b); once again, excluding all the derivations blocked by the strict cyclicity condition, one possibility is left: the relative pronoun is moved into COMP$_1$ at the first cycle, an interrogative pronoun is moved into COMP$_2$ at the second cycle, and at the third cycle the relative pronoun should be moved from COMP$_1$ to COMP$_3$; but, as before, this movement is blocked by subjacency.

The prediction of solution (B) is correct: in the following pairs of sentences, the first corresponding to structure (17) (a), the second corresponding to structure (17) (b), the second is systematically unacceptable:
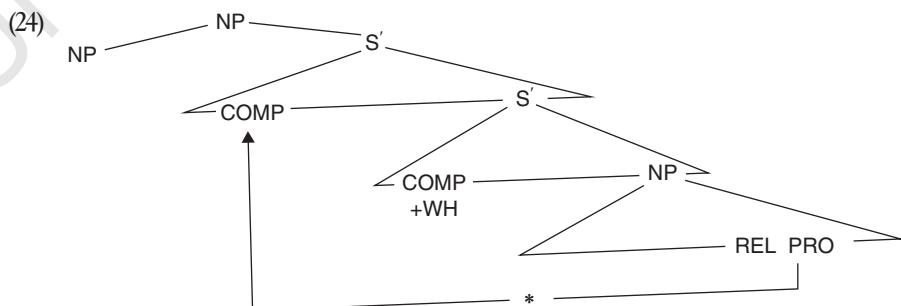
(18) a. Il mio primo libro, che credo che tu sappia a chi ho dedicato, mi è
          sempre stato molto caro
          (My first book, which I believe that you know to whom I dedicated,
          has always been very dear to me)
      b. *Il mio primo libro, che so a chi credi che abbia dedicato, mi è sempre
          stato molto caro
          (My first book, which I know to whom you believe that I dedicated,
          has always been very dear to me)
(19) a. La macchina che credo che Gianni si domandi se potrà utilizzare nel
          week end è la mia
          (The car that I believe that Mario wonders whether he will be allowed
          to use during the week end is mine)
      b. *La macchina che mi domando se Mario creda che potrà utilizzare nel
          week end è la mia
          (The car that I wonder whether Mario believes that he will be allowed
          to use during the week end is mine)

[. . .]

In conclusion, solution (B) is empirically superior to both solutions (A) and (C).
The results arrived at in this paragraph can be summarized as follows:

(1) Wh movement doesn't obey TSC and SSC; otherwise, the acceptable
     sentences (6), (18)(a)–(21)(a) would not be derivable;
(2) Wh movement does obey the subjacency condition; otherwise, the
     unacceptable sentences (13)(b)–(15)(b), (18)(b), (21)(b) would be derivable;
(3) The cyclic node which is relevant for subjacency is S′, not S; otherwise,
     there would be no way to distinguish between the acceptable sentences (6),
     (18)(a)–(21)(a), and the unacceptable sentences (13)(b)–(15)(b), (18)(b)–(21)
     (b), since none of them could be derived.

5. There is a further class of facts on which the hypothesis presented in this
paper makes non trivial predictions: it predicts that any movement out of a wh
clause should be blocked if the relative pronoun is extracted from a NP, under the
assumption that, in Italian, NP is a cyclic category (an assumption substantiated
by the fact that the complex NP constraint holds for Italian): such a movement
would violate the subjacency condition, as can be seen in the following schema:

(24)

The data which are relevant to check this prediction are the following:

(25) a. Questo autore, di cui non ricordo chi mi abbia parlato per la prima
          volta, mi sembra estremamente interessante
          (This author, of whom I do not remember who talked to me for the
          first time, seems to me very interesting)
      b. Questo autore, di cui non ricordo chi mi abbia mostrato il primo libro,
          mi sembra estremamente interessante
          (This author, by whom I do not remember who showed me the first
          book, seems to me very interesting)
      c. Questo autore, di cui ricordo che mi hai mostrato il primo libro, mi
          sembra estremamente interessante
(26) a. Francesca, di cui non so dove tu abbia sentito parlare prima d'ora, è
          una mia buona amica
          (Francesca, of whom I do not know where you have heard (someone)
          speaking before, is a good friend of mine)
      b. Francesca, di cui non so dove tu abbia conosciuto la sorella, è una mia
          buona amica
          (Francesca, of whom I don't know where you have met the sister (= whose
          sister I don't know . . .), is a good friend of mine)
      c. Francesca, di cui so che recentemente hai conosciuto la sorella, è una
          mia buona amica
          (Francesca, of whom I know that you have recently met the sister (= whose
          sister . . .), is a good friend of mine)
(27) a. Gianni, di cui immagino in che occasione ti abbiano parlato, . . .
          (Gianni, of whom I imagine in which occasion they talked to you, . . .)
      b. Gianni, di cui immagino in che occasione tu abbia visto la foto sul gior-
          nale, . . .
          (Gianni, of whom I imagine in which occasion you have seen the
          picture on the newspaper (= whose picture . . .), . . .)
      c. Gianni, di cui immagino che tu abbia visto la foto sul giornale qualche
          giorno fa, . . .
          (Gianni, of whom I imagine that you have seen the picture on the
          newspaper some days ago (= whose picture . . .), . . .)

In (a), a prepositional complement of the verb is moved out of an embedded
question; in (b) a prepositional complement of a noun is moved out of an embedded
question; in (c), a prepositional complement of a noun is moved out of an embed-
ded declarative. Now, if the extraction analysis which is accepted in this paragraph
is correct, my hypothesis predicts that the (b) sentences of the preceding triples
should be unacceptable. As the following rough derivations show, sentences (a) can
be derived, as proposed before, via direct movement of the wh phrase into the
relative COMP (cfr. (28)(a)); sentences (c) can be derived via successive cyclic
movement of the relativized PP (cfr. (28) (c)); but sentences (b) cannot be derived:
putting aside the derivation which would violate strict cyclicity, the only possibility
which is left violates subjacency, as indicated by the arrow in (28)(b):

(28)   a. $_{NP}[_{S'}$ COMP . . . $[_{S'}[_{COMP}$ chi] . . . $[_{pp}$ di cui]]]

b. $_{NP}[_{S'}$ COMP . . . $[_{S'}[_{COMP}$ chi] . . . $[_{NP}$ . . . $[_{pp}$ di cui]]]]

c. $_{NP}[_{S'}$ COMP . . . $[_{S'}[$ COMP Δ che] . . . $[_{NP}$ . . . $[_{pp}$ di cui]]]]

Even if the relevant intuitions are not extremely sharp, this prediction seems to be reasonably correct: sentences (b) of (25), (26), (27) are significantly worse than sentences (a)–(c).

[. . .]

Note

1   Cfr. Chomsky (1973, 1976, 1977).

8.3   Questions pertaining to Rizzi (1980)

1   Discuss the relevance to this paper of Rizzi's (1997) subsequent proposals about the left periphery.
2   Rizzi argues that in Italian the bounding nodes for subjacency must be S-bar (later integrated into X-bar theory as CP) and NP. Abney (1987) argued that what had previously been thought of as NP should instead be taken to be DP (with NP now contained within DP). Thinking of Rizzi's arguments in favor of taking S-bar to be a bounding node, rather than S (now IP or TP), discuss the implications of now taking DP to be a bounding node rather than NP.
3   To what extent can Chomsky's (2001) Phase Impenetrability Condition be used to mimic the effects of subjacency? To what extent is the question of which projections count as phases the same as the question of which nodes count as bounding nodes for subjacency?
4   Judgments on extraction from *wh*-islands in English are sometimes less sharp than in Italian. Yet of the following two sentences, the first is clearly more acceptable, to at least some speakers, with no speaker likely to have the reverse preference:

(1)   a.   ?Who aren't you sure what to say to?
      b.   *What aren't you sure who to provide with?

How might this acceptability difference be related to the following one discussed by Kayne (2008) concerning noncontrastive PP-topicalization?:

(2)   a.   ?To so many poor children, they've given so much money!
      b.   *With so much money, they've provided so many people!

5   For at least some speakers of English, the following is perfectly acceptable (especially with stress on *when*):

(3)   Which important person aren't you sure when you can arrange for me to see?

To what extent is this expected or not expected, given Rizzi's analysis?

6   Same question for (the somewhat less acceptable):

(4)   ?For that sort of reason, I can't imagine who would have resigned.

with the relevant reading (again, with stress on *who*) being the one in which *for that sort of reason* is interpreted within the *wh*-island.

7   In English, there is a sharp subject–object asymmetry concerning extraction from a *wh*-island:

(5)   a.   ?the woman who John isn't sure whether he should invite to the party
       b.   *the woman who John isn't sure whether should invite him to the party

What does this tell us about subjacency or about other locality conditions? (Extra credit: Bring in Richards (1997) on Scandinavian.)

8   Although (5b) of the preceding question is strongly unacceptable, the following reciprocal sentence is merely somewhat marginal:

(6)   a.   ?We always like to know where each other will be staying.

as opposed to the impossible reflexive:

       b.   *We always like to know where ourselves will be staying.

In a phase-based syntax, how might one try to interpret (6a) and (6b) here as involving (to differing degrees) locality violations on movement?

9   In English, (5b) of question 7 contrasts with:

(7)   We don't know where to go.

in which there is a silent subject (PRO) within a wh-island. How does this bear on movement-based analyses of control such as those in Hornstein (1999, Ch. 28 of this volume) or Kayne (2002)? Discuss the significance of the fact that many Germanic languages other than English disallow (7).

10   Control into *wh*-islands is often said to be impossible if the *wh*-word is *why*:

(8)   a.   *I don't know why to leave.

Yet certain other examples are much better:

       b.   I can't imagine why else to invite them.

(best with stress on *else*). Look further into this question in at least three languages other than English that have control with *wh*-infinitives.

11 Chomsky's (1973) integration into subjacency of Ross's (1967) Complex NP-Constraint makes extraction out of relative clauses of direct relevance to extraction out of *wh*-islands. What, then, is the significance of Taraldsen's (1982) demonstration that Norwegian under certain conditions readily allows extraction from within a relative?

12 Chung and McCloskey (1983) show that English sometimes allows extraction from a relative, too, though less readily, it seem, than in Norwegian. Why might Norwegian be freer here than English? (Hint: Look at Richards 1997.) Why might German and Dutch be less free here than English? (Hint: Look at Koster 1987.)

13 What is the status of nonrestructuring, non-*wh* control infinitives with respect to the notion of bounding node? Why would we or would we not expect gerunds and infinitives to have the same status? (Extra credit: Bring in other kinds of nonfinite embeddings.)

14 In his note 25, Rizzi considers the possibility that English might differ from Italian with respect to the question of which nodes each language takes to be bounding nodes for subjacency. How might that possibility be made compatible with the restriction on parameters discussed by Chomsky (1995), according to which parameters are necessarily associated with functional items in the lexicon?

15 Rizzi notes that free relatives prohibit extraction as strongly as headed relatives do, and suggests assimilating free relatives to headed ones, via the presence of a silent head. Discuss the merits and demerits of an alternative analysis of free relatives that instead takes them to be instances of embedded interrogatives, that is to say, of an analysis that takes *John will buy whatever is on sale* to be closely related to *No matter what is on sale, John will buy it*, with the free relative sentence containing a silent counterpart of *no matter*.

References

Abney, S. 1987. *The English noun phrase in its sentential aspect*. Doctoral dissertation, MIT.

Baker, M. 2001. *The Atoms of Language*. New York: Basic Books.

Borer, H. 1984. *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Dordrecht: Foris.

Chomsky, N. 1973. Conditions on transformations. In S. Anderson and P. Kiparsky (eds), *A Festschrift for Morris Halle*, New York: Holt, Rinehart, and Winston, pp. 232–286.

Chomsky, N. 1976. Conditions on rules of grammar. *Linguistic Analysis* 2: 303–351. [Reprinted in R. Cole (ed.), 1977, *Current Issues in Linguistic Theory*, Bloomington: Indiana University Press. Based on a lecture to the Linguistic Institute, South Florida, June 1975.]

Chomsky, N. 1977. On Wh-movement. In P. Culicover, T. Wasow, and A. Akmajian (eds), *Formal Syntax*, New York: Academic Press, pp. 71–132.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, pp. 1–52.

Chung, S., and J. McCloskey. 1983. On the interpretation of certain island effects in GPSG. *Linguistic Inquiry* 14: 703–713.

Di Sciullo, A. M., and C. Boeckx (eds). 2011. *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford Studies in Biolinguistics. Oxford and New York: Oxford University Press.

Engdahl, E. 1980. *Wh* constructions in Swedish and the relevance of subjacency. In J. Jensen (ed.), *Proceedings of NELS 10*, Amherst: GLSA, University of Massachusetts, pp. 89–108.

Grimshaw, J. 1986. Subjacency and the S/S′ parameter. *Linguistic Inquiry* 17: 364–369.

Hornstein, N. 1999. Movement and control. *Linguistic Inquiry* 30: 69–96.

Kayne, R. S. 2002. Pronouns and their antecedents. In S. D. Epstein and T. D. Seely (eds), *Derivation and Explanation in the Minimalist Program*, Oxford: Blackwell, pp. 133–166. [Reprinted in Kayne (2005a, Ch. 6).]

Kayne, R. S. 2005a. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2005b. Some notes on comparative syntax, with special reference to English and French. In G. Cinque and R. S. Kayne (eds), *The Oxford Handbook of Comparative Syntax*, New York and Oxford: Oxford University Press, pp. 3–69. [Reprinted in Kayne (2005a, Ch. 12).]

Kayne, R. S. 2006. On parameters and on principles of pronunciation. In H. Broekhuis, N. Corver, R. Huibregts, U. Kleinhenz, and J. Koster (eds), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, Berlin and New York: Mouton de Gruyter, pp. 289–299. [Reprinted in Kayne (2010, Ch. 4).]

Kayne, R. S. 2008. Antisymmetry and the lexicon. In J. Rooryck and J. van Craenenbroeck (eds), *Linguistic Variation Yearbook*, Amsterdam and Philadelphia: John Benjamins, vol. 8, pp. 1–31. [Reprinted in Kayne (2010, Ch. 9) and in Di Sciullo and Boeckx (2011, Ch. 14).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Kayne, R. S. to appear. Comparative syntax. *Lingua*.

Koster, J. 1987. *Domains and Dynasties. The Radical Autonomy of Syntax*. Dordrecht: Foris.

Maling, J., and A. Zaenen. 1982. A phrase-structure account of Scandinavian extraction phenomena. In P. Jacobson and G. K. Pullum (eds), *The Nature of Syntactic Representation*, Dordrecht: Reidel, pp. 229–282.

Reinhart, T. 1981. A second COMP position. In A. Belletti, L. Brandi, and L. Rizzi (eds), *Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW conference*, Pisa: Scuola Normale Superiore, pp. 517–557.

Richards, N. 1997. *What moves where when in which language?* Doctoral dissertation, MIT.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 281–337.

Rizzi, L. 2011. On the elements of syntactic variation. In *Working Papers of Centro Interdipartimentale di Studi Cognitivi sul Linguaggio (CISCL)*, Università di Siena. Distributed by MIT Working Papers in Linguistics.

Ross, J. R. 1967. *Constraints on variables in syntax*. Doctoral dissertation, MIT. [Published 1986 as *Infinite Syntax!* by ABLEX, Norwood, NJ.]

Sportiche, D. 1981. Bounding nodes in French. *The Linguistic Review* 1: 219–246.

Taraldsen, K. T. 1982. Extraction from relative clauses in Norwegian. In E. Engdahl and E. Ejerhed (eds), *Readings in Unbounded Dependencies in Scandinavian Languages*, Stockholm: Almqvist & Wiksell International, pp. 205–222.

Torrego, E. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15: 103–129.

9

On Certain Differences between French and English

Richard S. Kayne

1981

9.1   Introduction

In *Violations of the Wh Island Constraint in Italian and the Subjacency Condition*, Luigi Rizzi posits a single parameter to account for a difference between the grammatical systems of Italian and English that determines whether or not extraction from certain syntactic environments ("*wh*-islands") is possible. While a parameter whose effects lie entirely in extraction phenomena might seem natural, other cases of proposed parameters have effects in a range of areas across the grammar. Such parameters are posited on the basis of linkages in the properties of what might seem on the surface to be unrelated grammatical phenomena, locating these linked differences in a single abstract grammatical property. This is what we see in *On Certain Differences between French and English*: a parameter that brings together properties that at first sight might seem unrelated.

In this paper, Kayne draws a connection among several differences between English and French. One is that English allows a certain class of verbs to be followed by an infinitival complement with an overt lexical subject, whereas French does not, as illustrated in (1). A second difference is that English allows the object of a preposition (P) to be fronted, leaving the P behind (so called "preposition stranding"), whereas French does not, as shown in (2):

(1) a.   John believes Bill to have lied.
     b. *Jean croit Bill avoir menti.
(2) a.   Which candidate have you voted for?
     b. *Quel candidat as-tu voté pour?

Chomsky (1980) suggested that verbs like *believe* take an S-bar complement and can govern the embedded subject (and assign it case) across the S-bar (so-called "Exceptional Case Marking," ECM). Such an approach, however, leaves unexplained why the same option is not available to French *croire*. Kayne abandons Chomsky's assumption and proposes instead that embedded infinitival clauses are introduced by a prepositional complementizer in both English and French. This can be *for* in English and *de* (or *à*) in French, or, in the case of *believe*-type verbs, a phonetically null complementizer Φ (cf. Chomsky and Lasnik 1977): John believes/croit $[_{S'} \Phi [_{S} \dots]]$

Kayne characterizes the relevant difference between English and French as follows: prepositional complementizers govern the subject position of the complement clause in English, but not in French. This accounts for the pattern in (1): the subject of an infinitival complement of *believe*-type verbs can be overt in English, because Φ governs it and assigns it case, but not in French. Moreover, it accounts for the fact that the subject of an infinitival clause complement of *believe*-type verbs can be null in French, but not in English, as we see in (3):

(3)   a.   *I believe to have made a mistake.
      b.   Je crois avoir fait une erreur.

Assume that the null subject of an infinitival clause is PRO, an element that cannot be governed; it then follows that PRO is possible in French, because Φ fails to govern it, but impossible in English, because Φ governs it. This approach can also make sense of the fact that, when the complement of *believe*-type verbs is not an infinitival clause, but rather a "small clause" consisting of a subject and a non-verbal predicate, as in (4), an overt subject is possible in both English and French:

(4)   a.   I believe/judge John intelligent.
      b.   Je crois/juge Jean intelligent.

Small clauses are not introduced by a complementizer; in both English and French the verb governs and assigns case to the subject of the small clause.

Finally, this proposal can account for the differences in P-stranding shown in (2). Like prepositional complementizers, English and French prepositions more generally differ in their ability to govern. In English, Ps govern "structurally," like verbs, that is, by virtue of the structural configuration in which they occur with respect to a noun phrase; in French, in contrast, they only govern a noun phrase that is their argument, unlike verbs. As a consequence, in English Ps can undergo a process of "reanalysis" (cf. Hornstein and Weinberg 1981) with the verb and govern the trace of the object, as in (2a), whereas in French, where Ps and Vs govern in different ways, they cannot, as in (2b).

In this view, a single difference in the grammatical system of English and French, namely the ability of prepositions to govern structurally, accounts for the fact that English allows P-stranding and can have overt (but not null) subjects in infinitival clauses that are the complement of *believe*-type verbs, whereas French exhibits the opposite pattern. At the same time, it can account for some of the

similarities the two languages exhibit, for example in allowing overt subjects in small clause complements. This constitutes a step forward in building a model of grammar that can derive, from a small set of abstract properties, a larger number of superficial differences.

9.2 From "ON CERTAIN DIFFERENCES BETWEEN FRENCH AND ENGLISH"

English, but not French, allows preposition stranding in *wh*-constructions and in passives:

(1) a. Which candidate have you voted for?
    b. *Quel candidat as-tu voté pour?
(2) a. John was voted against by almost everybody.
    b. *Jean a été voté contre par presque tous.

If preposition stranding involves some kind of reanalysis between verb and preposition, then it might simply be that English, but not French, has that kind of reanalysis.

English, but not French, allows verbs like *believe* to be followed by an infinitival complement with a lexical subject:

(3) a. John believes Bill to have lied.
    b. *Jean croit Bill avoir menti.

If (3a) involves exceptional Case-marking across a clause boundary, then it might simply be that French lacks such exceptional Case-marking. Yet one could ask why that should be so.

We shall attempt to show, not only that there is more to be said about both of the above differences between French and English, but also that these two differences are in fact related to one another.

1 The complementizer status of *de* and *di*

To bring out the relation between preposition stranding and exceptional Case-marking, we shall need a bridge, whose construction will depend on a comparative analysis of the complementizer systems of French and English. Following Bresnan (1970; 1972), we shall assume that English has the complementizers *that* and *for* and that occurrences of them are dominated by a node COMP. We further assume that French *que* corresponds straightforwardly to English *that* (and thus that *que* occurs under COMP).

The question is whether or not there is a complementizer in French that corresponds to English *for*. An obvious candidate is *de*:

(4)   Je lui ai dit qu'il parte.
        'I told him that he (should) leave.'
(5)   Je lui ai dit de partir.

Apart from its dative object, here a clitic, the verb *dire* takes a sentential comple-
ment, which in (4) begins with complementizer *que*. It is natural to analyze (5) in
parallel fashion: the (infinitival) sentential complement of *dire* is *de partir*, which
begins with complementizer *de*.

   That *de* in (5) is part of the sentential complement is emphasized by its absence
in (6):

(6)   a.   Je lui ai dit quelque chose.
                'I told him something.'
       b.   *Je lui ai dit de quelque chose.

   Paradigms such as (4)–(6) play an important role in Huot's (1977, chapter 3)
detailed argument in favor of the complementizer status of *de*. And it does seem
clear, in the light of (4)–(6), that the *de* of (5) is within the sentential complement
of *dire*. That conclusion, however, would not be incompatible with the claim that
de in (5) is a French equivalent of English *to*, rather than a true complementizer in
Bresnan's sense. Thus, to defend the complementizer status of *de*, we need to
diminish the plausibility of pairing *de* with *to*.

   One straightforward piece of evidence, alluded to by Long (1974, chapter 4,
note 37) and implicit in Huot (1977, 282), comes from the observation that *de*, like
for but unlike *to*, is excluded from infinitival complements headed by a
wh-phrase:

(7)   Je lui ai dit où aller.
        'I told him where (to) go.'
(8)   *Je lui ai dit où d'aller.
(9)   Elle cherche quelqu'un avec qui parler.
        'She's looking for someone with whom (to) speak.'
(10) *Elle   cherche quelqu'un avec qui de parler.

If the *de* of (5) is a complementizer, then (8) and (10) reduce to the general restric-
tion against doubly-filled COMP; that is, they are excluded for the same reason as
(11) and (12), much as in Bresnan (1972, chapter IC [sic]) (cf. also Kayne ([1981a],
section 3.3)):

(11) *I told him where for her to go.
(12) *She's looking for someone with whom for her child to speak.

Were *de* a French equivalent of *to*, the ungrammaticality of (8) and (10) would
remain puzzling, since *to* itself cooccurs uneventfully with *wh*-phrases:

(13) I told him where to go.
(14) She's looking for someone with whom to speak.

The next three pieces of evidence that we shall present in favor of the complementizer status of *de* will be enhanced if we consider Italian at the same time as English and French. With respect to the data considered so far in this section, Italian is like French:

(15)   Gli ho detto di partire.   (= (5))
        him(I) told   leave
        'I told him to leave.'
(16)   Gli ho detto qualcosa.   (= (6a))
(17)   *Gli ho detto di qualcosa.   (= (6b))
(18)   Gli ho detto dove andare.   (= (7))
(19)   *Gli ho detto dove di andare.   (= (8))
(20)   Cerca qualcuno con cui parlare.   (= (9))
(21)   *Cerca qualcuno con cui di parlare.   (= (10))

In other words, the hypothesis that French *de* is a complementizer can be naturally extended to Italian *di*.

De and *di* occur in a large number of control contexts. Some examples like the *dire* cases, with a dative controller, are given in (22) and (23):

(22)   Je lui ai interdit/suggéré/demandé de partir.
        'I prohibited/suggested (to)/asked him to leave.'
(23)   Gli ho proibito/suggerito/chiesto di partire.

There are also many cases of subject control:

(24)   Jean a essayé/oublié/décidé de partir.
        'John tried/forgot/decided to leave.'
(25)   Gianni ha tentato/dimenticato/deciso di partire.

However, the French and Italian equivalents of the (small) class of verbs treated by Rosenbaum (1967) as instances of raising to subject position are uniformly incompatible with *de/di*:

(26)   *Jean semble/paraît/se trouve/s'avère d'être parti.
(27)   *Gianni sembra/pare/risulta di essere partito.

Without *de/di*, these are all grammatical:

(28)   Jean semble/paraît/se trouve/s'avère être parti.   (= (36))
(29)   Gianni sembra/pare/risulta essere partito.

That the impossibility of *de/di* in (26)–(27) reflects a significant generalization is suggested especially by Italian *sembrare* and *parere*, which can, in contrast to (27), occur with *di* in a dative control context:

(30)   Mi sembra/pare di aver capito.
        '(It) seems/appears (to) me (that I) have understood.'

The incompatibility of *de/di* with raising as Rosenbaum describes it is understandable under the hypothesis that *de* and *di* are complementizers. If they are, then the ungrammaticality of (26) and (27) can be taken as another instance of the well-known restriction against extracting subjects across an adjacent complementizer:

(31) a. *Who would you prefer for to leave first?
     b. *Who did you think that had married her?
(32) *Qui croyais-tu que l'avait épousée?
(33) a. *They seem for to speak English.
     b. *They seem that speak English.
(34) *Ils semblent que parlent anglais.
(35) *Sembrano che parlino inglese.

More precisely, the complementizer status of *de/di* should allow us to derive the ungrammaticality of (26) and (27) from the principle(s) of grammar responsible for (31)–(35).

Were *de* and *di* not complementizers, but rather French and Italian equivalents of *to*, their incompatibility with raising would be difficult to understand, given the well-formedness of (36), where *to* appears in a raising context:

(36) John seems/appears/happens/turns out to have left.

We conclude that these data constitute evidence for the complementizer status of *de* and *di*.

[. . .]

2   A Government difference between *de* and *for*

We shall henceforth assume the main result of the previous section, namely that *de* and *di* are complementizers, but we shall not explicitly mention Italian unless there is some pertinent difference between it and French.

De and *for* now have in common their status as complementizers, in particular as infinitival complementizers. There are of course two major differences between them. First, whereas *for* can be followed by a lexical subject of the infinitive, *de* cannot.

(56) *Ce serait dommage de quelque chose lui arriver.
(57) It would be a pity for something to happen to him.

Second, *de* is compatible with control, whereas *for* is not:

(58) Ce serait dommage de partir maintenant.
(59) *It would be a pity for to leave now.

The ungrammaticality of (56) could reasonably be related to that of (60):

(60) *Quelque chose lui arriver serait dommage.   ($\simeq$ (61))

That is, lexical subjects of infinitives are normally impossible in French, and the same holds for English, in the absence of *for* or an appropriate matrix verb:

(61) *Something to happen to him would be a pity.

Following Chomsky (1980, 25), let us assume a Case filter that requires every lexical NP to be marked for Case. Then, assuming that subjects of infinitives cannot receive Case from within the infinitival S itself, both (60) and (61) are straightforward violations of the Case filter. Furthermore, we can distinguish (56) from (57), while grouping (56) with (60) and (61), by agreeing that complementizer *de* cannot assign Case to the subject of the infinitive, whereas *for* can. In other words, we can interpret (56) as a violation of the Case filter.

Since government is a necessary condition for Case assignment in Chomsky (1980, 25), it follows also that *for* must govern the adjacent subject position. Chomsky ([1981a; 1981b]) has constructed a theory within which government of some position is incompatible with control of that position (cf. also Rouveret and Vergnaud (1980, 124)). Hence, government of the adjacent subject position by *for* will preclude control, which accounts for the ungrammaticality of (59).

Let us, moreover, interpret the inability of *de* in (56) to assign Case to the subject NP as an indication that *de* in COMP does not govern the adjacent subject position. Then the possibility of control is straightforward, which accounts for (58). In other words, the two differences that we started out with, namely the contrasts between (56) and (57) and between (58) and (59), reduce to one: English *for* governs the adjacent infinitival subject position, but French *de* does not.

From here, we shall proceed as follows: we shall generalize this specific difference between English *for* and French *de*, in two steps. The intermediate generalization of section 3 will allow us to effect a connection with the *believe* facts observed in (3). The subsequent generalization of section 4 will establish a connection with preposition stranding ((1)–(2)).

3 The apparent differences between *croire* and *believe*

The promised intermediate generalization is a slight one: English prepositional complementizers govern the adjacent infinitival subject position, but French prepositional complementizers do not.

This generalization has two effects. First, if *à* can be a complementizer in French, a possibility raised in Kayne (1975, section 4.10), then *à* falls (correctly) under this generalization, since it is compatible with control and incompatible with a following lexical subject:

(62) a. *Je cherche quelqu'un à   Jean photographier.
          I   look for someone   for/to John (to) photograph
      b. *Marie est facile à   Jean contenter.
          Mary is easy for/to John (to) please

The second effect concerns *believe*. Before examining it, let us recall the basic data. The French word-for-word counterpart of English (63) is ungrammatical:

(63)   I believe/acknowledge/have determined John to be the most intelligent of all.

(64)   *Je crois/reconnais/constate Jean être le plus intelligent de tous.

[. . .]

Recalling that in Chomsky's ([1981a; 1981b]) framework, there is a (negative) relation between government and control, as discussed above toward the end of section 2, we would expect that, within that framework, with the class of verbs at issue, control with an infinitive would be on the whole impossible in English, but possible in French. And there does in fact exist such a difference:

(68)   *I believe/acknowledge/affirm to have made a mistake.
(69)   Je crois/reconnais/affirme avoir fait une erreur.

[. . .]

Let us ask why *believe* should allow government of the embedded subject position, but not *croire*. Why is it not the other way around?

Our answer will exploit the similarity, within the Case/government framework, between the *believe/croire* difference and the *for/de* difference of section 2. The latter, we recall, led to the generalization given at the beginning of this section: English prepositional complementizers govern the adjacent infinitival subject position, but French prepositional complementizers do not. The former difference led to the conclusion that the matrix V could govern the infinitival subject position in English, but not in French [. . .].

Our basic idea is that government across two S-type boundaries is uniformly impossible, and that English only appears to allow it. We can achieve this result if we integrate into the Case/government framework Chomsky and Lasnik's (1977, section 2.2.2) ϕ complementizer. In particular, let us assume with them that *believe*-type verbs take a ϕ complementizer. We assume further that ϕ is another prepositional complementizer, which differs from *for* and *de* in having no phonetic realization.

Once having made these assumptions, we can consider that *John* in *believe* (ϕ (*John to be happy*)) receives its Case from ϕ, and not directly from *believe*. Now there is no reason why French should not have a prepositional ϕ complementizer with the same class of verbs: *croire* (ϕ (*Jean être heureux*)). However, since French prepositional complementizers do not govern the adjacent subject position, ϕ will not govern *Jean*, with the result that it will receive no Case and will violate the Case filter.

Similarly, we can consider ϕ to appear uniformly in both (68) and (69). In (68), ϕ governs the embedded subject position and blocks control. In (69), ϕ, though present, does not govern the embedded subject position, and control is possible.

From this point of view, there is no essential difference between *believe* and *croire*. The apparent differences between them are rather a function of the way in which English and French treat prepositional complementizers. The contrasts

between (68) and (69) and between (63) and (64) are thus connected to the contrasts between *for* and *de* discussed in section 2, i.e. to (59) vs. (58) and to (57) vs. (56).

If this is correct, then there need be and can be no exceptional Case-marking in the strict sense of Chomsky (1980) even in English. There cannot be, since our account of French (64) depends not only on the government properties of φ, but also on the unavailability to French of direct Case-marking from *croire* to the embedded subject, an unavailability that is most simply interpreted as reflecting nonexistence in universal grammar. (This nonexistence in turn follows from the characterization of government as a relation not capable of spanning more than one S-type boundary.)

Furthermore, like *croire* in French, *believe* in English must be subcategorized for an $\overline{S}$ complement (that can be infinitival), and it must not be subcategorized for _____ NP VP. If this latter kind of subcategorization were available to *believe*, there would be no principled reason for its being unavailable to *croire*, and the French–English difference would become puzzling again.

In conclusion, study of the *believe*–lexical NP–infinitival VP construction in English can lead to the postulation of various analyses: $V - \overline{S}$ with Raising (Postal (1974)); $V - \overline{S}$ without Raising, with "exceptional Case-marking" across two boundaries (Chomsky (1980)); $V - \overline{S}$ without Raising, with "exceptional COMP Deletion" (Rouveret and Vergnaud (1980, section 1.6)); V–NP–VP with lexical-interpretive mechanisms (Bresnan (1978, 35); cf. also Dowty (1978, section 8.1) and Wasow (1980, section 3)); $V - \overline{S}$ without Raising, with Case-marking via an abstract prepositional complementizer (this article). Of these, however, we would claim that only the last is restricted enough to allow an adequate account of the absence of the corresponding construction in French.

This conclusion, that *believe*'s infinitival complement is an $\overline{S}$, is relevant to the status of Passive, since, as Bresnan ([1982], section 3.7) has noted clearly, an analysis of Passive as a purely lexical rule would not be compatible with (the passivizability of) *believe* ($_{\overline{S}}$ NP VP$_{inf}$); cf. Wasow (1977, section 2.2).

[. . .]

4 The unifying difference between English and French prepositions

The difference in preposition stranding between English and French that is illustrated in (1) and (2) has been attributed by Hornstein and Weinberg (1981) to the existence, in English only, of a Reanalysis rule that amalgamates V and P into one constituent, much as in Chomsky's (1974; 1980, 26) analysis of *take advantage of*. We shall accept the existence of a Reanalysis rule in English, while following Vergnaud's (1979) suggestion that, at least in (1), what is involved is not so much reanalysis qua constituent as reanalysis in terms of government, essentially as in Rouveret and Vergnaud's (1980) proposal for French causatives and related verbs.

[. . .]

[R]ather than interpreting the lack of preposition stranding in French as resulting from the absence of a reanalysis rule, let us state more precisely that it results from the absence of a reanalysis rule involving prepositions.

Comparing in particular the V–V reanalysis in causatives with the absence of V–P reanalysis, in French, it seems that there must be some important difference between V and P at issue. This recalls the differential proposal in Chomsky (1980, 25–26) concerning Case Assignment: P assigns (oblique) Case in the base, whereas V assigns (objective) Case elsewhere than in the base. It also suggests the following principle: reanalysis between two lexical categories is possible only if they assign Case in the same way.

The idea that the lexical categories V and P assign Case differently from one another might be expressed independently of the point of application of Case Assignment. Taking the association of subcategorization with the base as a starting point, consider the possibility that P can assign oblique Case only to an NP for which it is subcategorized, whereas V can assign objective Case somewhat more freely, in particular to any NP that it governs.

If this were so, then (in French) V, but not P, could assign Case in the configuration $\left(\left\{ \begin{matrix} V \\ P \end{matrix} \right\} (_{\overline{S}} \; NP \; X) \right)$, since government can span a single boundary of type S. For example, we have argued that V assigns Case in precisely this configuration in (39) and (65) [in the full article], but there are no such instances of cross- $\overline{S}$ Case Assignment from P.

A second relevant example is complementizer *de*, which occurs in the configuration P $(_S \; NP \; X)$. Since *de* is not subcategorized for the subject of S, we would expect that P could not assign Case to that NP. This is consonant with the data examined previously, e.g. (56).

We noted earlier that the compatibility of *de* with control means that *de* does not govern the adjacent subject position. This suggests a slight generalization of the V vs. P Case Assignment difference: V governs NP in the structural sense of Chomsky (1980) and Rouveret and Vergnaud (1980), but normally P governs NP only in the sense of subcategorization.

This suggests in turn revising the reanalysis principle: reanalysis between two lexical categories is possible only if the two govern in the same way.

We now see that the French–English contrast with respect to preposition stranding is this:

(71)   In French, P and V do not govern in the same way; but in English they
        do. (That is, in English, P can govern structurally, as well.)

This recalls our intermediate generalization from the beginning of section 3: English prepositional complementizers govern the adjacent infinitival subject position, but French prepositional complementizers do not. Since the relation between COMP and the adjacent subject position is one of structural government, but not one of subcategorization, this intermediate generalization is simply a special case of (71).

In other words, (71) covers both preposition stranding and prepositional complementizer government. Since the latter is what underlies the French–English contrast with respect to "exceptional Case-marking" (that is, *croire* vs. *believe*), as

we argued in section 3, (71) is the principle underlying that, too.[22] This is the relation promised in the opening paragraphs.

   [. . .]

9.3   Questions pertaining to Kayne (1981)

1   English robustly allows preposition stranding, as in *Who were you talking to?* French does not. To what extent is it legitimate to ask why French doesn't allow what English does allow? Same question for the absence in French of any word-for-word counterpart of *For there to be another war would be unfortunate*.

2   Do you think that the absence in French of *Who were you talking to?* and the absence in French of *For there to be another war would be unfortunate* are related phenomena? Give your reasons.

3   In trying to answer questions like (1) and (2), should you take into consideration languages other than French and English themselves? Why or why not? Should you be looking at other areas of French and English syntax, too? Why or why not? What other areas might you choose to look at?

4   Is preposition stranding in English a unified phenomenon? To what extent is it relevant that Icelandic has preposition stranding in interrogatives corresponding to *Who were you talking to?*, but not in sentences corresponding word-for-word to *They're being talked about by almost everybody* (which are called "pseudo-passives")?

5   How general is English preposition stranding? Bring into the discussion section 3 of Kayne (1998).

6   Is preposition stranding more common across languages than a complementizer like *for* or is it the other way around? Mention at least five relevant languages.

7   Sentences like *We want them to be happy*, with an embedded infinitive, are run-of-the-mill English. Why might they be absent in other Germanic languages and in Romance languages? Bring into the discussion Postal (1974) and Kayne (1981b).

8   Pick a language of your choice other than English and find five types of sentences that that language allows but that English doesn't allow. Does this set of contrasts require postulating five separate parameters? Why, or why not?

9   English allows preposition stranding both in restrictive relatives like *the person we were talking to* and in nonrestrictive relatives like *John, who we were talking to*. French allows preposition stranding in neither type of relative. Should we count that as one difference between English and French or as two differences? Give your reasons.

10   English allows E(xceptional)C(ase)M(arking) sentences like *John believes your daughter to be intelligent*. By and large French does not (for nuances, see Pollock 1985). On the other hand, English does not allow **John believes to be intelligent*, with control. In French such sentences are often acceptable, e.g., *Jean croit être intelligent*. Are these two differences between English and French reducible to a single difference, do you think? Why or why not?

11 In comparing **John believes to be intelligent* and *Jean croit être intelligent*, we glossed over the fact that the French example has no word matching English *to*. If we try to add a word like *to*, the result is unacceptable: **Jean croit de/à être intelligent*. What other facts of French does this recall?

12 (Extra credit) Italian also has an acceptable counterpart of English **John believes to be intelligent*, namely *Gianni crede di essere intelligente*, in which the prepositional complementizer *di* is obligatory: **Gianni crede essere intelligente*. In the previous question we saw that the opposite holds for French, where the prepositional complementizer is impossible: *Jean croit* (**d'*)*être intelligent*. How might this Italian–French contrast be related to the following contrast concerning subjunctive?: With verbs like *believe*, Italian allows subjunctive to appear in the embedded sentence, whereas French (apart from polarity contexts) does not.

13 French subject control sentences sometimes have a prepositional complementizer, as in *Jean essaie de gagner* ('J tries de win'), and sometimes have none, as in *Ils désirent gagner* ('they desire win'), or with *croire* ('believe') as in *Jean croit être intelligent* ('J believes be intelligent'). Object control in French almost always requires a prepositional complementizer, as in *Ils nous demandent de revenir* ('they us ask de return' = 'they're asking us to come back'). One of only two (types of) exceptions is *Il me semble avoir compris* ('it me seems have understood' = 'It seems to me that I have understood'). Why might this be an exception?

14 The other exception is *Il me faut partir* ('it me needs leave' = 'I need to leave'). Why might this one be an exception? (Hint: Take a look at the French counterparts of English modals.)

15 (Extra credit) With its counterpart of 'seem,' Italian, like French, allows both control and raising. English allows raising, as in *They seem to have understood the question*, but not control: **It seems to them to have understood the question*. Might this be related to the unacceptability in English of **They believe to have understood the question*? Justify your answer.

16 With 'seem,' Italian has both raising and control, but has a prepositional complementizer only in the latter. With raising, Italian has *Maria sembra aver capito* ('M seems have understood'), with no *di*. With (dative object) control, Italian has *Mi sembra di aver capito* ('me seems di have understood'), with *di* required. Is it an accident that Italian has *di* with control and not with raising? Could it have been the other way round? Give your reasons. (Extra credit: How does this bear on the general question of the relation between control and raising discussed, for example, in Hornstein 1999, Ch. 28 of this volume, and Landau 2003, 2006?)

17 (Extra credit) English has *They don't know if they should leave*, but doesn't allow control here: **They don't know if to leave*. Italian does: *Loro non sanno se partire*, with *se* = 'if.' French, on the other hand, is like English: **Ils ne savent pas si partir* (with *si* = 'if'). Why might this be? (Hint: Take a look at Kayne 1991, Ch. 18 of this volume.) To what extent might Italian allowing control with *se* here be related to its allowing control with *di* with *sembrare* ('seem') and with *credere* ('believe')?)

18 How do Italian *di* and French *de* and English *for* fit into Rizzi's (1997) finer-grained analysis of the complementizer system?

19   The ECM approach to sentences like *John believes Bill to be intelligent* that goes back to Chomsky (1973) competes with the raising approach to such sentences defended in detail by Postal (1974). To what extent is the choice between ECM and raising akin to the choice between Chomsky's (2001) use of Agree and his earlier use of Spec-head agreement?

20   To what extent is there a certain tension between ECM and Agree, on the one hand, and Chomsky (1995) Extension principle, on the other?

21   How might Kayne's use of a silent prepositional complementizer in English sentences like *John believes Bill to be intelligent* be transposed to a raising analysis, if ECM should turn out to be incorrect?

22   There is a contrast in English between *John, who I assure you to be among the best students in the class*, . . . and **I assure you John to be . . .*, with Wh-movement apparently saving an ECM-like sentence that would otherwise not be possible. A proposal is made in Kayne (1980) in terms of Case-assignment into Comp. How might that idea be transposed into a raising approach to such sentences. (Hint: Bring in the acceptability, for some speakers, of ?*these people, who John all think should be invited* discussed in Kayne (2003).)

23   Italian and French both disallow prepositional complementizers with raising of the *seem* sort, that is in their counterparts of *There seems to be a problem*. Yet they allow them with raising of the sort found with aspectuals, i.e. in their counterparts of *There began to be problems, There have stopped being problems*. Discuss the possible implications of this discrepancy in behavior.

References

Bresnan, J. 1970. On complementizers: Toward a syntactic theory of complement types. *Foundations of Language* 6: 297–321.

Bresnan, J. 1972. *The theory of complementation in English syntax*. Doctoral dissertation, MIT. [Published 1979 by Garland Publishers, New York.]

Bresnan, J. 1978. A realistic transformational grammar. In M. Halle, J. Bresnan, and G. A. Miller (eds), *Linguistic Theory and Psychological Reality*, Cambridge, MA: MIT Press, pp. 1–59.

Bresnan, J. 1982. The passive in lexical theory. In J. Bresnan (ed.), *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press.

Chomsky, N. 1973. Conditions on transformations. In S. Anderson and P. Kiparsky (eds), *A Festschrift for Morris Halle*, New York: Holt, Rinehart, and Winston, pp. 232–286.

Chomsky, N. 1974. *The Amherst Lectures*. (Lectures given at the 1974 Linguistic Institute, University of Massachusetts, Amherst, June 1974.) *Documents Linguistiques*, Paris: Université de Paris VII, Département de Recherches Linguistiques.

Chomsky, N. 1980. On binding. *Linguistic Inquiry* 11: 1–46.

Chomsky, N. 1981a. Markedness and core grammar. In A. Belletti, L. Brandi, G. Nencioni, and L. Rizzi (eds), *Theory of Markedness in Generative Grammar –*

Proceedings of the III GLOW Conference, Pisa: Scuola Normale Superiore di Pisa, pp. 123–146.

Chomsky, N. 1981b. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, pp. 1–52.

Chomsky, N., and H. Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8: 425–504.

Dowty, D. R. 1978. Governed transformations as lexical rules in a Montague grammar. *Linguistic Inquiry* 9: 393–426.

Hornstein, N., and A. Weinberg. 1981. Case theory and preposition stranding. *Linguistic Inquiry* 12: 55–91.

Hornstein, N. 1999. Movement and control. *Linguistic Inquiry* 30: 69–96.

Huot, H. 1977. *Recherches sur la subordination en français*. Thèse de doctorat d'état, Université de Paris VIII.

Kayne, R. S. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

Kayne, R. S. 1979. Extensions du liage et du marquage du cas. *Lingvisticae Investigationes* 3: 29–55.

Kayne, R. S. 1980. Extensions of binding and Case-marking. *Linguistic Inquiry* 11: 75–96. [Revised version of Kayne (1979). Reprinted in Kayne (1984, Ch. 1).]

Kayne, R. S. 1981a. ECP Extensions. *Linguistic Inquiry* 12: 93–133.

Kayne, R. S. 1981b. Two notes on the NIC. In A. Belletti, L. Brandi, and L. Rizzi (eds), *Theory of Markedness in Generative Grammar. Proceedings of the 1979 GLOW conference*, Pisa: Scuola Normale Superiore di Pisa, pp. 317–346. [Reprinted in Kayne (1984, Ch. 2).]

Kayne, R. S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, R. S. 1991. Romance clitics, verb movement and PRO. *Linguistic Inquiry* 22: 647–686. [Reprinted in Kayne (2000, Ch. 5).]

Kayne, R. S. 1998. Overt vs. covert movement. *Syntax* 1: 128–191.

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2003. Some remarks on agreement and on Heavy-NP-Shift. In M. Ukaji, M. Ike-Uchi, Y. Nishimura, and H. Nakajima (eds), *Current Issues in English Linguistics.* Vol. 2 of *Special Publications of the English Linguistic Society of Japan,* Tokyo: Kaitakusha, pp. 67–86. [Reprinted in Kayne (2005, Ch. 11).]

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Landau, I. 2003. Movement out of control. *Linguistic Inquiry* 34: 471–498.

Landau, I. 2006. Severing the distribution of PRO from case. *Syntax* 9: 153–170.

Long, M. E. 1974. *Semantic verb classes and their role in French predicate complementation*. Doctoral dissertation, University of Indiana. [Reproduced 1976 by the Indiana University Linguistics Club, Bloomington.]

Pollock, J.-Y. 1985. On case and the syntax of infinitives in French. In J. Guéron, H.-G. Obenauer, and J.-Y. Pollock (eds), *Grammatical Representation*, Dordrecht: Foris, pp. 293–326.

Postal, P. 1974. *On Raising*. Cambridge, MA: MIT Press.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 281–337.

Rosenbaum, P. S. 1967. The Grammar of English Predicate Complement Constructions. Cambridge, MA: MIT Press.

Rouveret, A., and J.-R. Vergnaud. 1980. Specifying reference to the subject: French causatives and conditions on representations. *Linguistic Inquiry* 11: 97–202.

Vergnaud, J.-R. 1979. Case and binding. Paper given at the Pisa Workshop.

Wasow, T. 1977. Transformations and the lexicon. In P. Culicover, T. Wasow, and A. Akmajian (eds), *Formal Syntax*, New York: Academic Press, pp. 327–360.

Wasow, T. 1980. Major and minor rules in lexical grammar. In T. Hoekstra, H. van der Hulst, and M. Moortgat (eds), *Lexical Grammar*, Dordrecht: Foris, pp. 285–312.

10

Move WH in a Language without WH Movement

C. T. James Huang

1982

10.1 Introduction

In trying to understand the structure of human language, progress can be made both by comparing languages that are very similar to one another and by comparing those that are at first sight very different: both types of comparison allow us to pursue a more precise understanding of the limits of language variation. Huang's article *Move WH in a language without WH movement* is an example of the latter type of comparison. It observes that Chinese and English differ in the way they form questions and clefts, and raises the issue of how this difference is encoded in their grammatical systems and, in particular, whether Chinese lacks *wh*-movement.

Huang points out that, although in Chinese *wh*-phrases do not appear to move in questions, they can be interpreted as if they occupied more than one structural position. For example, in (1a), the *wh*-phrase that is the subject of the embedded clause occurs in the canonical subject position. Yet the sentence has two readings: the one expressed in English by a *wh*-phrase in situ (1b), where we have a declarative clause that contains an embedded question, and the one expressed in English by fronting the *wh*-phrase (1c):

(1) a. Zhangsan zhidao shei mai–le shu.   (Chinese)
         Zhangsan knows who bought books
     b. 'Zhangsan knows who bought books.'
     c. 'Who does Zhangsan know bought books?'

Huang interprets this as evidence that Chinese has a movement operation that fronts *wh*-phrases (just as English does), which applies at LF. This raises the expectation that movement of the *wh*-phrase should be structurally constrained, just as overt *wh*-movement is. This expectation is partially met: *wh*-phrases in Chinese cannot move out of a specific noun phrase, a restriction that is captured by the Specificity Condition (Fiengo and Higginbotham 1981). On the other hand, Huang notes that certain *wh*-phrases in Chinese, notably NPs like *who* and *what*, do not seem to obey Subjacency, whereas others, like *how* and *why*, do. (The latter class also includes elements marking "A-not-A" questions, or the focused element in a cleft sentence.) He suggests that this asymmetry may perhaps hold of in-situ *wh*-expressions universally and proposes (in Huang 1982) that it follows from the ECP applying at the level of LF but Subjacency applying only in overt syntax, an account further developed in Lasnik and Saito (1984). (The status of Subjacency in LF has in turn been further explored in Nishigauchi 1986, 1990; Pesetsky 1987; Fiengo et al. 1988; Richards 1996; von Stechow 1996; Starke 2001; among many others.)

In sum, Huang accounts for certain similarities and differences between English and Chinese by arguing that, although both grammars have *wh*-movement, it applies at different levels of syntactic representation. This can be seen both as an argument for the level of syntactic representation that we call LF, and as an example of how to account for parametric variation.

Pesetsky (1987) further strengthened the argument for the existence of a level of LF, and for *wh*-movement being possible at that level, based on some insightful observations on English questions with multiple *wh*-phrases (such as *Who read what?*). The issue of the level of syntactic representation at which *wh*-movement applies, and what exactly moves, has been further explored in several pieces of work (cf. Watanabe 1992; Hagstrom 1998, 2004; among others). For example, Hagstrom analyzes *wh*-movement in Japanese and Sinhala (cf. also Kishimoto 2005), suggesting that it is the Q-particle that moves, overtly in one language and covertly in the other. Cable (2008) extends some of the ideas from these proposals to the distinction between languages with fronted and those with in-situ *wh*-phrases, arguing that the difference is not in the level at which they move, but rather in the nature of the constituent that moves.

10.2   From "MOVE WH IN A LANGUAGE WITHOUT WH MOVEMENT"

1   Introduction

1.1   The study of constraints in grammar has been one of the most persistent topics in generative studies. Among the most important results of this enterprise is a set of locality conditions, including most notably Ross' island constraints [Ross 1967] and Chomsky's Subjacency Condition [Chomsky 1973, 1977]. These locality conditions are usually defined over certain structural configurations and a rule of movement, deletion or at least some kind of dependency between two structural positions.

One type of inquiry that arises from studies of these conditions concerns their status in languages where there is no structural configuration or overt dependency between two structural positions meeting the definition of the various islands and constraints. For example, many languages do not have a WH-word fronting rule in syntax to form a WH question, and some do not move any element for the purpose of subdividing a sentence into focus and presupposition (i.e. of forming a cleft sentence). Since questions and sentences with focus are universal sentence types, it is natural to ask what the proposed structural conditions have to say about these languages. It would seem that since these languages do not exhibit the defining configurations for the conditions to be applicable, they would be exempt from these conditions. If it turns out, however, that these types of sentences do show properties as if they form islands and obey island conditions, then it might seem that the conditions as formulated in structural terms must be rejected or revised. As a point of logic, however, this fact may lead us instead to ask if it is a mistake to regard these sentence types as exhibiting absolutely no structural configurations in every possible sense. That is, although they do not show overtly the required structural properties, couldn't they be seen as actually involving them in some abstract sense?

In this paper, I will try to demonstrate that taking this latter position, instead of concluding the invalidity of the structurally based conditions, will lead us to the prediction of certain very interesting facts and provide a simple explanation of these facts. In particular, I will discuss certain properties of two major sentence types in Chinese, questions and cleft sentences, and show that, although they obviously do not involve any movement in Syntax, it is desirable to assume that they undergo movement in the Logical Form component of grammar. Thus, for example, while it is clear that a WH question like (1) below is generated in Syntax with the WH word never anywhere but in its base position, it will be argued that in the LF interpretive component the WH word is moved to a position c-commanding the sentence, leaving a trace interpreted as a variable bound to it, as in (2):

(1)   ni   xihuan shei?
       you like   who
       'Who do you like?'
(2)   [shei$_i$ [ni   xihuan e$_i$]]
        who you like

The assumption that abstract movement rules of this sort exist in UG is an important feature of some recent works within the Extended Standard Theory, and is supported by a fairly wide range of facts observed in English and several other languages. It will not be unreasonable to assume the existence of such abstract devices in Chinese simply as a consequence of UG, but it will be worthwhile to ask if there are some positive language-specific motivations for making such an assumption. The main purpose of this paper is to show that there are indeed strong language-specific motivations.

1.2.   The assumption that the LF representation of (1) is of the form (2) has as its immediate consequence the existence of:

(3)   a.   a quantifier
       b.   an empty category
       c.   a movement process

and the properties associated with each of (3a–c). The main body of this paper will be concerned with case (c), the existence of movement and its properties. Case (b) will be briefly dealt with in the Appendix [in the original article]. As regards case (a), a brief discussion in this section will give some support for the treatment of WH words as a kind of quantifier and provide some initial motivation for the rest of the paper. Consider the following sentences:

(4)   [Zhangsan wen wo [shei mai-le shu]]
                    ask   me   who bought books
       'Zhangsan asked me who bought books.'
(5)   [Zhangsan xiangxin [shei mai-le shu]]
                    believe   who bought books
       'Who does Zhangsan believe bought books?'
(6)   [Zhangsan zhidao [shei mai-le shu]]
                    know   who bought books
       a.   'Who does Zhangsan know bought books?'
       b.   'Zhangsan knows who bought books.'

The only surface difference among these sentences is in the choice of the matrix verb. In (4), *wen* 'ask' belongs to a class of verbs that require an interrogative complement. In (5), *xiangxin* 'believe' does not permit an interrogative complement. In (6), *zhidao* 'know' may optionally take an interrogative complement. As the translation shows, this single difference in the choice of the verb is responsible for the fact that (4) must be interpreted as a statement taking an indirect question, (5) must be interpreted as a direct question embedding no indirect questions, and (6) may be interpreted as either. It makes good sense to ask how the very different meanings of the virtually identical (4) and (5), as well as the ambiguity of (6), may be represented in an optimal theory. One natural approach to this question is to look at an indirect question like (4) as one in which the question word has scope over the embedded sentence, and a direct question like (5) as one in which the question word takes scope over the entire sentence, while in (6) the question word may take either scope. This amounts to postulating (7)–(9) as the logical forms for (4)–(6), respectively:

(7)   [Zhangsan wen wo [shei$_x$ [x mai-le shu]]]
                    ask   me who   bought book
       'Zhangsan asked me for which x, x bought books.'

(8) [shei$_x$ [Zhangsan xiangxin [x mai–le shu]]]
      who   believe   bought book
    'For which x, Zhangsan believes x bought books.'

(9) a. [Zhangsan zhidao [shei$_x$ [x mai–le shu]]]
                  know   who   bought book
        'Zhangsan knows for which x, x bought books.'

     b. [shei$_x$ [Zhangsan zhidao [x mai–le shu]]]
          who   know   bought book
        'For which x, Zhangsan knows x bought books.'

It is a well known property of quantifiers that they exhibit scope phenomena. Since the difference between (4) and (5) and that between the two readings of (6) are naturally seen as difference in scope of a WH word, treating a WH word as a quantifier, as is shown in (7)–(9), is not unreasonable. Obviously, this is not the only possible way to account for these facts in Chinese, but note that these are not facts perculiar [sic] to Chinese; they are paralleled by corresponding sentences in other languages, including those in which a WH word is actually moved to a quantifier position in surface form. Given that facts in UG should be treated in a uniform way for all languages and that there is already a natural representation given by the surface form of overt WH-moved languages, it is entirely reasonable to postulate the abstract representations like (7)–(9) for a language without overt movement.

     Another construction for which a case can be made for an abstract movement rule in LF is the cleft construction in Chinese. The formation of a cleft sentence in this language clearly does not involve the dislocation of any constituent in Syntax. In surface structure a cleft sentence differs from a non-cleft in that in a cleft there is a focus marker (the copula *shi*) immediately preceding the focused constituent. (In the glosses below, FM=Focus Marker.)

(10) a. shi wo mingtian yao   mai neiben shu
          FM I   tomorrow want buy that   book
          'It is I that want to buy that book tomorrow.'
     b. wo shi mingtian yao   mai neiben shu
          I   FM tomorrow want buy that   book
          'It is tomorrow that I want to buy the book.'
     c. wo mingtian shi yao   mai neiben shu
          I   tomorrow FM want buy that   book
          'I do want to buy that book tomorrow.'

Since a cleft sentence has the universal semantic property of dichotomizing a sentence into focus and presupposition, it is natural to provide a unified representation of this dichotomy in LF. One reasonable assumption is that the focused material is also regarded as a quasi-quantifier binding a variable in the presupposition, so that (10a) can be represented as (11):

(11) [(shi wo)$_x$ [x mingtian yao   mai neiben shu]]
        FM I   tomorrow want buy that   book

By convention we may then interpret the operator and the open sentence in (11) as representing its focus and presupposition, respectively. Evidently this is not the only possible representation of the meaning of a cleft sentence in Chinese; in particular, it may appear that there is no need to assume a quantifier-variable relationship which is not visible in surface structure. However, such a representation is a reasonable one as a unified formal device to represent all focalizing constructions in this language and across languages. If we adopt the quantifier-variable representation, it also offers a convenient way to state a generalization about the following observation. Consider (12):

(12)   Zhangsan shuo [Lisi shi mingtian lai]
                    say   FM tomorrow come
        'Zhangsan said that it is tomorrow that Lisi will come.'
        or "It is tomorrow that Zhangsan said that Lisi will come."

In this sentence, where the focus *tomorrow* occurs in the embedded clause, two interpretations of the focus can be distinguished. It may indicate the emphasis of the speaker of the entire sentence, or it may indicate the emphasis of the matrix subject. In the former case the embedded clause is part of an indirect speech, while in the latter it may be seen as representing a direct quotation in some sense. To obtain an intuitively correct representation for each of these two meanings, we may then allow the focus to be moved either to a position c-commanding the entire sentence or to a position c-commanding only the embedded clause.

1.3. With these initial motivations, I will postulate the existence of two LF rules for Chinese, called Move WH and FOCUS. It should be kept in mind that these are intended to be two instances of the general rule Move α, or the general rule Move WH as the term is used in Chomsky (1977), i.e. movement to operator positions. For easy reference, I will describe them as if they were two rules, using the term Move WH to refer only to cases of actual WH questions. Also, for expository purposes, I shall make the familiar assumption that in either case Move α moves an $\overline{\overline{X}}$ phrase into a clause-initial COMP, an operation that is also involved in the mapping from DS to SS in deriving relativized and topicalized constructions, and furthermore that it may apply successive-cyclically through COMP. Each category to be moved is assumed to contain either the feature [+WH] or [+Focus]. The movement will affect a proper $\overline{\overline{X}}$ phrase in accordance with (a proper version of) the A–over–A Condition (cf. Bresnan 1976, May 1977), and will leave a coindexed trace at the extraction site in accordance with the trace theory. The result of such movement will be further converted into more formal representations in the following way. Following Chomsky ([1981]), a moved focused constituent will be represented as an operator having the form "for x=… ." Thus the focus in (11) above is interpreted as the operator "for x=I". (Alternatively, the focused phrase can be simply read as "it is I (that …)," analogous to cleft sentences in English.) A moved WH phrase, on the other hand, will be interpreted as the quantifier "for which x; x … ," where "… ." is a predication indicating the domain or extension of the quantifier x from which a value may be drawn to substitute for the trace, now

interpreted as a bound variable, in fixing the truth value of the sentence. Thus, *who* is interpreted as "for which x; x a person." More precisely, it will be assumed that the predicate *a person* is in fact a bundle of features containing all features of *who* except [+WH], i.e. [+N, +animate, +human,...]. This last assumption will have non-trivial consequences in our account for sentences in which a WH word is clefted, as we shall see in Section 2.3.

2 Cleft sentences

2.1.   We have seen above that in a sentence like (12), a focused constituent may occur in a clause embedded as complement to a verb but have scope over the entire matrix sentence. However, not every embedded clause may contain a focused element. (13)–(15) show that the clefted element may not occur within a relative clause or a sentential subject (the morpheme *de*, glossed as DE in (13)–(14), is a relative clause marker):

(13)   *[wo xihuan [shi Zhangsan mai de   neizhi gou]]
         I   like   FM   buy DE that   dog
       '*I like the dog that it is Zhangsan that bought.'
(14)   *[[Zhangsan shi zuotian   mai de neiben shu] hen hao]
                      FM yesterday buy DE that   book very good
       '*The book that it was yesterday that Zhangsan bought is very good.'
(15)   *[[Zhangsan shi mingtian lai]   mei guanxi]
                      FM tomorrow come no   matter
       '*That it is tomorrow that Zhangsan will come does not matter.'

Since in these sentences the focus cannot be interpreted as indicating the emphasis of a matrix subject (see footnote 5 [in the full article]), it has to be interpreted as indicating the emphasis of the speaker. Although (12) is good with the speaker-emphasis interpretation, (13)–(15), however, are not. This shows that long distance clefting is possible from verb phrase complement positions, but not from a syntactic island like a complex NP or a sentential subject. This suggests that cleft sentence formation in Chinese, although it does not involve an overt movement rule, nevertheless has to obey Subjacency. It has been claimed (e.g. Huang 1980) that relativization and topicalization in this language generally obey Subjacency. Since relative and topicalized structures are describable as involving certain overt antecedent–gap relations, it is most natural to consider them as constituting confirming evidence for the universality of this (possibly parameterized) condition. Since cleft sentences show the same locality properties as (12)–(15) indicate, there is no reason not to assume that cleft sentences are subject to the same condition. The fact that these sentences do not exhibit overt structural dependencies, however, has posed a serious formal problem for the proposed putative universal condition in a theory with-out our LF rule of FOCUS.

   According to the hypothesis being suggested here, the problem disappears if we hypothesize that the LF rule FOCUS (=Move α), like the syntactic rules of

relativization and topicalization (also instances of Move α), is subject to Subjacency. The rule FOCUS must move the focused constituent into the highest COMP in deriving the logical forms for (13)–(15), but such an operation would have to raise an element out of a syntactic island. Thus, assuming that the bounding nodes for Subjacency in Chinese are NP and S and that sentential subjects are dominated by NP, raising the focus into the highest COMP in (13)–(15) would violate Subjacency. On the other hand, FOCUS may apply successive-cyclically in (12) from within an object complement without violating Subjacency. Given the rule and the assump-tion that it obeys Subjacency, the ill-formedness of (13)–(15) is automatically predicted.

2.2.   The proposal just made also allows us to predict that a sentence containing two focused elements is bad:

(16)   *[$_\bar{S}$ [$_s$ shi Zhangsan shi   mingtian   yao   lai]]
              FM   FM tomorrow want come
       '*It is Zhangsan that it is tomorrow that will come.'

In (16) there is only one position for COMP but two focuses that must be moved into it. Either of the following two plausible alternatives may account for its ill-formedness. On the one hand, we may stipulate that movement may take place only into an unfilled COMP (see e.g. Chomsky 1973). In this way, moving either focus in (16) will fill up the only COMP it has, and (16) will be ruled out because the other focus that has to be moved has no place to move to:

(17)   [$_\bar{S}$ [$_{comp}$ [For x=Zhangsan]] [$_S$ x   shi mingtian   yao   lai]]
                                                        FM tomorrow want come

On the other hand, one may allow movement of both focuses into the same COMP, and rule out the output at LF by general conditions of proper binding (see e.g. Freidin 1978). The result of such movement will contain a branching COMP as in (18), which is ill-formed since neither operator c-commands its variable:

(18)   [$_\bar{S}$ [$_{comp}$ [For x=Zhangsan][For y=mingtian   ]] [$_s$ x y yao   lai]]
                                                        tomorrow   want come

Similarly, sentence (19) can have only the interpretation according to which the embedded focus indicates the emphasis of the matrix subject, and not that of the speaker, unlike (12):

(19)   [$_\bar{S}$ [$_s$ shi Zhangsan shuo [$_\bar{S}$ [$_s$ Lisi shi   mingtian   lai]]]]
              FM   said   FM tomorrow come
       'It is Zhangsan that said that it is tomorrow that Lisi will come.'

The sentence contains a focus in the matrix and one in the embedded clause. If the embedded focus is moved to the matrix COMP (together with the matrix focus),

the output will be ill-formed. The theory also predicts that the sentence (20) has no possible interpretation at all:

(20)   *[$_\bar{s}$ [$_s$ Zhangsan shuo [$_\bar{s}$ [$_s$ shi Lisi shi   mingtian lai]]]]
                          said   FM   FM tomorrow come

In this sentence the embedded clause contains two focuses. If both are interpreted as indicating the speaker's emphasis, or both interpreted as indicating the matrix subject's emphasis, the ill-formedness of the sentence is already explained. Furthermore, the sentence also does not have an interpretation according to which one of the embedded focuses represents the speaker's emphasis and the other that of the matrix subject. Suppose that *Lisi* is the speaker's focus and *tomorrow* the matrix subject's. Then, depending upon whether FOCUS has applied successive-cyclically, the LF representation of (20) will be either (21) or (22):

(21)   [$_\bar{s}$ [For x=Lisi][$_s$Zhangsan shuo [$_\bar{s}$ [$_{comp}$[x][For y=mingtian]]-
                                    said   tomorrow
       [$_s$ x y lai   ]]]]
              come
(22)   [$_\bar{s}$ [For x=Lisi][$_s$Zhangsan shuo [$_\bar{s}$ [For y=mingtian][$_s$ x y lai]]]]
                                    said   tomorrow   come

Neither of these representations is well-formed since (21) contains a doubly-filled COMP and in (22) the relation between [For x=Lisi] and its variable x violates Subjacency. Alternatively, again, we may say that neither (21) nor (22) are derivable in LF because derivation of the former will involve violation of the ban on move-ment into a filled COMP and derivation of the latter will violate Subjacency as a condition on movement.

What we have witnessed here is that application of FOCUS may be blocked not only by a syntactic island like a complex NP or a sentential subject, etc., but also by what may be called a "Focus Island" formed by a previous application of the same rule. We will now show that FOCUS may also be blocked by a "WH Island" formed by a prior application of Move WH. (23) and (24) show that no element may be clefted from within a direct or an indirect question:

(23)   *[$_\bar{s}$ [$_s$ shi Zhangsan da–le shei]]?
                FM   beat who
       '*Who is it Zhangsan that beat?'
(24)   *[$_\bar{s}$ [$_s$ ta xiang-zhidao [$_\bar{s}$ [$_s$ shi Zhangsan da–le shei]]]]
                he wonder   FM   beat who
       '*He wonders who it is Zhangsan that beat.'

In (23), both the focus *Zhangsan* and the WH word *shei* 'who' have to be moved into the same COMP. Therefore (23) is excluded for exactly the same reason as (16). In (24), the WH word has to be moved to the lower COMP, as required by the main

verb *xiang-zhidao* 'wonder', and the focus must be moved either to the lower or to the higher COMP. But as is the case with (20), neither interpretation of (24) is possible, since derivation of either logical form involves violation of Subjacency or filling a COMP with more than one operator.

Note that if the focus occurs on a matrix constituent and the WH word is contained in an embedded clause, the sentence is fine just in case the WH word has scope over only the embedded clause, but bad if it has matrix scope. There is a clear contrast between (25a), which is a declarative containing an indirect question, and (25b), which is a direct question:

(25)   a.   shi Lisi xiang-zhidao [shei da-le ta]
              FM   wonder   who beat him
              'It is Lisi that wonders who beat him.'
        b.   *shi Lisi xiangxin [shei da-le ta]?
              FM   believe   who beat him
              '*Who is it Lisi that believes that beat him?'

The contrast is a direct consequence of our theory. The focus in (25a) will be moved in LF to the matrix COMP from the main clause, and the WH word will only be moved into the embedded COMP from the embedded clause (since the main verb "wonder" takes an interrogative complement). No known principle of grammar is violated in the derivation of the logical form of (25a). In (25b), however, the main verb cannot take an interrogative complement, so both the focus and the WH word have to be moved into the higher COMP. (25b) is therefore ill-formed, for reasons already seen. Likewise, we also correctly predict that, if the main verb of (25a) is changed to *zhidao* 'know,' which only optionally takes interrogative complements, the sentence will be unambiguous, with only the interpretation that the embedded clause is an indirect question:

(26)   shi Lisi zhidao [shei da-le ta]
        FM   know   who beat him
        'It is Lisi that knows who beat him.'

The formal account that we have proposed thus expresses in a simple manner what appears to be the correct generalization: no element may be clefted from within the scope of a WH word. It is not correct to say just that a focus may not co-occur with a WH word.

    [. . .]

3   WH questions

3.1.   We have seen that (a) FOCUS is subject to a full range of island conditions, and (b) both Move WH and FOCUS have the effect of forming an island. It is natural to ask if Move WH is also subject to the same island conditions. Investigation of relevant sentences in the language suggests that, in many cases, the answer seems to be yes.

(32) *[s[np[s̄ tou–le sheme de] neige ren]   bei dai–le]?
          stole   what   DE that   person by caught
      '*The man that stole *what* was caught?'

(33) *[s[np[s̄ ni   weisheme mei mai de] neiben shu] hen hao]?
          you why   not buy DE that   book very good
      '*The book that you did not buy *why* is very good?'

(34) *[s[s̄ Zhangsan tao–le shei] zhen kexi]?
                    marry who real   pity
      '*That Zhangsan married *whom* is a real pity?'

(35) *[s[s̄ Zhangsan tao–le shei], ni   zhidao–le]?
                  marry who   you know
      '*That Zhangsan married *whom*, you know?'

In these sentences, a direct question is being asked to obtain the hearer's answer to specify the value of a WH word that occurs in a relative clause (32–33), a sentential subject (34), or a sentential topic (35). As indicated, these sentences are bad. On the other hand, recall that a direct question with a WH word in a sentential *object* complement embedded arbitrarily deep in the tree is perfectly well-formed, as shown by sentences (5) and (6) above. These facts seem to suggest quite strongly that, like the rule FOCUS, Move WH in Chinese should be assumed to obey Subjacency. It turns out, however, that these facts only represent a part of the whole picture. For one thing, there are a number of counterexamples to the claim that Move WH obeys the CNPC subcase of Subjacency:

(36) [s[np[s̄ shei yao   mai de] shu] zui   gui]?
            who want buy DE book most expensive
      '*Books that *who* wants to buy are most expensive?'

(37) [s ni   xiang kan [np[s̄ ta shemeshihou pai de] dianying]]?
          you want see   he when   film DE movie
      '*You want to see movies that he filmed *when*?'

S.H. Teng (p.c.) also gives the following example:

(38) [s ni   xihuan [np[s̄ wo piping   shei de] wenzhang]]?
        you like   I   criticize who DE article
      '*You like articles in which I criticize *who*?'

In (36) the speaker is, in effect, asking which book, in terms of the identity of the person who is buying it, is the most expensive. In (37), the speaker asks which movie you want to see, in terms of the time when the movie was filmed. Similarly, (38) may be paraphrased as: "In terms of the person I criticize, tell me which articles [of mine] you like." Each of these questions can be asked, if the context of situation so allows, and an answer can be given where the WH word is given a value.

Since structurally there is no doubt that, in (36)–(38), the WH word appears within a complex NP, the hypothesis that Subjacency is the principle ruling out (32)–(35) must be considered inadequate if we do not have any independent explanation for the grammaticality of (36)–(38).

A comparison of (32)–(33) with (36)–(38), where a complex NP contains a WH word in each case, reveals one important difference: In the ungrammatical (32)–(33), the head noun of the complex NP is preceded by a demonstrative, while in the grammatical (36)–(38), there is no demonstrative occurring with the head. This difference suggests that the semantic notion of specificity is relevant. In the appendix [to the full article], I will show that the ungrammatical (32)–(33) should be ruled out by an independently motivated principle along the lines of the Specificity Condition proposed in Fiengo and Higginbotham (1981). The principle prohibits a quantifier contained within a specific NP from having a scope larger than that NP, i.e. specific NPs cannot contain free variables. (32)–(33) are out, since in each of them the complex NP containing a WH word is specific due to the presence of a demonstrative. The grammaticality of (36)–(38), on the other hand, can be accounted for if we make the stipulation that the WH words in question are "wide scope" quantifiers that need not obey Subjacency. This has the further consequence that although WH islands may block FOCUS, they do not block Move WH, a prediction supported by the fact that (39) may be answered by either (40) or (41):

(39) [ni xiang-zhidao [shei mai-le sheme]]?
       you wonder   who bought what
(40) [wo xiang-zhidao [Lisi mai-le sheme]]
       I   wonder   bought what
       'I wonder what Lisi bought.'
(41) [wo xiang-zhidao [shei mai-le shu]]
       I   wonder   who bought book
       'I wonder who bought books.'

(40) answers (39) by replacing the embedded WH subject with *Lisi* and leaves the other WH word unanswered. (41) answers (39) by fixing the value of the embedded WH object but not the WH subject. (If the question is uttered with emphatic stress on *shei* 'who' then (40) comes as a more natural answer. If *sheme* 'what' is stressed, then (41) comes more readily.) It is general practice to regard the meaning of a question as one that defines the range of acceptable answers to it. Since (40) and (41) are acceptable answers to (39), this means (39) can be a direct question containing an indirect question, with either WH word having matrix scope and the other having embedded scope. The logical forms of (39) corresponding to the answers (40) and (41) are (42) and (43), respectively:

(42) $[_{\bar{S}}$ shei$_x$ $[_S$ ni   xiang-zhidao $[_{\bar{S}}$ sheme$_y$ $[_S$ x mai-le y]]]]
       who   you wonder   what   bought
(43) $[_{\bar{S}}$ sheme$_y$ $[_S$ni   xiang-zhidao $[_{\bar{S}}$ shei$_x$ $[_S$ x mai-le y]]]]
       what   you wonder   who   bought

It is obvious that in (42) the relationship between *who*$_x$ and its variable x, and in (43) the relationship between *what*$_y$ and its variable y, violates Subjacency.

3.2. [. . .] In contrast to the two representations (42)–(43), the corresponding English sentences (47)–(48) are bad, since each WH wor in them has been moved in Syntax in violation of Subjacency.

(47)   *What did you wonder who bought?
(48)   *Who did you wonder what bought?

On the other hand, like their Chinese counterparts, (49)–(50) are good with *what* syntactically unmoved:

(49)   Who bought what?
(50)   Who wonders where we bought what?

   [. . .]

If the generalization holds true that syntactically unmoved WH words may violate island conditions, therefore, the wide-scope property of such WH words in Chinese may be a consequence of UG, and need not be learned.

   [. . .]

4   Conclusion

To summarize, I have argued that for certain syntactic constructions it is desirable to assume that they involve Move α, if not in Syntax then in LF, and that the theory of bounding obtains in both modules of grammar. I have illustrated three instances of Move α in the LF of Chinese, Move WH, FOCUS, and Move A-not-A, and shown that, like topicalization and relativization (two instances of Move α that apply in Syntax), they have the effect of forming islands. Furthermore, although the various islands do not have the same range of island effects – in particular, while islands formed in Syntax block Move α in LF, those formed in LF do not block Move α in Syntax – it has been shown that this is a direct consequence of our assumption of the bounding theory as a set of conditions on movement and of the ordering of rule application imposed by the organization of grammar.

   [. . .]

The approach taken here, if correct, has several significant implications for linguistic theory. Among others, it offers evidence for the structural basis of island constraints. It also lends some support to the organization of grammar assumed here, with Syntax feeding into LF. But the most interesting implication, in my view, is that it suggests a refreshing way of looking at certain problems in typological studies. Thus, according to the view taken here, language families do not differ in whether or not they have a particular movement rule; nor do they differ in whether or not they are subject to Subjacency, etc. Rather, languages may be considered to incorporate certain substantive universals and

formal conditions, but to differ in where these universal rules apply, in Syntax or in LF. Considerations of this kind have led us to the findings of this paper concerning several similarities and differences between two typologically different languages. This provides strong support for the view that a good theory of typology should be, in the words of Ken Hale, "the by-product of a good theory of language."

10.3   Questions pertaining to Huang (1982)

1   Discuss the ways in which Huang's discussion might have been different had he written his paper subsequent to the appearance of Rizzi (1997).
2   Discuss the ways in which Watanabe (1992) might be considered to have improved upon Huang's analysis.
3   How might Huang's analysis be recast in terms of Chomsky's subsequent probe/goal based Agree?
4   In his note 9, Huang says that "We assume that the focus marker *shi* is a copulative adverb, dominated by the node EMP, the emphatic modality." Discuss the ways in which this idea is or is not similar to Kaynes (1998) later analysis of English *only*.
5   In his note 2, Huang notes that postverbal complements in Chinese cannot be immediately preceded by the focus marker. How might one relate this fact to the syntax of Chinese *dou*, as discussed, for example, by Cheng (1995)?
6   Discuss the possibility of relating the prohibition against postverbal focus markers in Chinese to apparently comparable facts about focused phrases in Malayalam (cf. Jayaseelan 2001) and many other (so-called) SOV languages.
7   SVO languages often allow focused phrases to be postverbal, as, for example, in Chadic (cf. Tuller 1992) and in English (cf. Kayne 1998, section 4). SVO languages can also have focused phrases somewhere to the left of V, as in Rizzi (1997). Against this background and that of the previous question, consider the following conjecture: SOV languages never have postverbal focus. Try to find a counterexample. If you cannot, make a proposal about what this conjecture might follow from.
8   English allows sentences like *Yesterday, what did you do?, To John, what did you say?*, and (for some speakers) even *John what did you say to?* How do these bear on Huang's analysis?
9   English interrogative *wh*-words, as in *Which chair was he seated on?*, are bare, in the sense that they are unaccompanied by any visible determiner. Such bare *wh*-words are also found in English in relative clauses, where they are called relative pronouns, as in *the chair on which he was seated*. In some languages (cf. Kuroda 1968, Kayne 2008b), the corresponding *wh-* word used as a relative pronoun is in some cases accompanied by a visible definite article, as in French *la chaise sur laquelle il était assis* ('the chair on the-which he was seated'). Discuss the possibility of taking all *wh*-word relative pronouns, in all languages,

to be accompanied by a definite article (that would sometimes, as in contemporary English, be silent).

10 English does not allow its bare *wh*-words to be used as simple indefinites. An English sentence like *John bought what yesterday* can, with appropriate intonation, be interpreted as an interrogative, but cannot be interpreted parallel to *John bought something yesterday*. As discussed in Postma (1994), Dutch differs from English in this respect. What kind of parameter (or parameters) might underlie this English–Dutch difference?

11 Both Chinese and Japanese are like Dutch in allowing their bare *wh*-words to be used as indefinites. Watanabe (1992) proposes that interrogative bare *wh*-words in Japanese are accompanied by a silent operator (which we can take to be a silent determiner) that moves to Spec, CP even as the *wh*-word itself remains in situ. Discuss the possibility that bare *wh*-words used as indefinites are also accompanied by a silent determiner, taking into account Bruening's (2007, p. 160) point that bare *wh*-words can never act as specific indefinites (as opposed to *wh*-words accompanied by a visible determiner). (Extra credit: Bring in Tsai 2003.)

12 Although English disallows bare *wh*-words as simple indefinites, it sometimes allows *wh*-words as part of a larger indefinite phrase containing a visible determiner, as in *somewhere. Somewhere* seems very close to *someplace*, especially insofar as being followed by an adjective or *else* is concerned: *somewhere interesting, someplace interesting* vs. **some city interesting*, and similarly for *somewhere else, someplace else* vs. **some city else*. The *place* of *someplace* would readily be called a "light noun." What, then, can be said about the status of *where* in *somewhere*? Bring in Kayne (2007), as well as Leu's (2008, Ch. 6) claim that *which* is phrasal, rather than being a single morpheme.

13 How does the possible light noun behavior of *place* (akin to *thing, body*) mentioned in the previous question (e.g., *someplace else* vs. *some other place*) bear on the question of how (*some+*)*place* is interpreted? Similarly, how do the differences between *someplace* and *somewhere* mentioned in Kayne (2007) bear on how *where* is interpreted? Bring in Bruening's (2007, p. 144) suggestion (following Kamp 1981; Heim 1982) that indefinites like Chinese *wh*-words are open predicates.

14 Huang argues that the Specificity Condition of Fiengo and Higginbotham (1981) is at work in Chinese relatives, while subjacency is not, but that subjacency is relevant to relatives of the English type that clearly display overt movement. How would you fit in to Huang's discussion of relatives both Taraldsen (1981) and Chung and McCloskey (1983)?

15 Huang characterizes in terms of the notion "objectual" the difference in behavior seen with respect to *wh*-islands between *wh*-words like *who* and *what*, on the one hand, and *how* and *why*, on the other. How might this distinction have to be refined in light of Starke's (2001) study of differences between locatives and temporals? How might Starke's work account for the multiple differences (e.g., *thereof* vs. **thenof*) between *there* and *then* observed in Kayne (2008a, section 9)?

References

Bresnan, J. 1976. On the form and functioning of transformations. *Linguistic Inquiry* 7: 3–40.

Bruening, B. 2007. Wh-in-situ does not correlate with wh-indefinites or question particles. *Linguistic Inquiry* 38: 139–166.

Cable, S. 2008. Q-particles and the nature of Wh-fronting. In L. Matthewson (ed.), *Quantification: A Cross-Linguistic Perspective*. Vol. 64 of *North-Holland Linguistics Series*, Bingley: Emerald, pp. 105–178.

Cheng, L. L.-S. 1995. On *dou*-quantification. *Journal of East Asian Linguistics* 4: 197–234.

Chomsky, N. 1973. Conditions on transformations. In S. Anderson and P. Kiparsky (eds), *A Festschrift for Morris Halle*, New York: Holt, Rinehart, and Winston, pp. 232–286.

Chomsky, N. 1977. On Wh-Movement. In P. Culicover, T. Wasow, and A. Akmajian (eds.), *Formal Syntax*, New York: Academic Press, pp. 71–132.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chung, S., and J. McCloskey. 1983. On the interpretation of certain island effects in GPSG. *Linguistic Inquiry* 14: 703–713.

Fiengo, R., and J. Higginbotham. 1981. Opacity in NP. *Linguistic Analysis* 7: 395–421.

Fiengo, R., J. Huang, H. Lasnik, and T. Reinhart. 1988. The syntax of wh-in-situ. In H. Borer (ed.), *Proceedings of the Seventh West Coast Conference on Formal Linguistics*, Stanford: CSLI, pp. 81–98.

Freidin, R. 1978. Cyclicity and the theory of grammar. *Linguistic Inquiry* 9: 519–549.

Hagstrom, P. 1998. *Decomposing questions*. Doctoral dissertation, MIT.

Hagstrom, P. 2004. Particle movement in Sinhala and Japanese. In V. Dayal and A. Mahajan (eds), *Clause Structure in South Asian Languages*, Dordrecht: Kluwer, pp. 227–252.

Heim, I. 1982. *The semantics of definite and indefinite noun phrases*. Doctoral dissertation, University of Massachusetts, Amherst.

Huang, C.-T. J. 1980. Topicalization and relativization in Chinese. Mimeograph, MIT.

Huang, C.-T. J. 1982. *Logical relations in Chinese and the theory of grammar*. Doctoral dissertation, MIT.

Jayaseelan, K. A. 2001. IP-internal topic and focus phrases. *Studia Linguistica* 55: 39–75.

Kamp, H. 1981. A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, and M. Stokhof (eds), *Formal Methods in the Study of Language*, Amsterdam: Mathematisch Centrum, pp. 277–322.

Kayne, R. S. 1998. Overt vs. covert movement. *Syntax* 1: 128–191.

Kayne, R. S. 2007. A short note on *where* vs. *place*. In R. Maschi, N. Penello, and P. Rizzolatti (eds), *Miscellanea di Studi Linguistici offerti a Laura Vanelli da amici e allievi padovani*, Udine: Forum, pp. 245–257. [Also in Kayne (2010, Ch. 5).]

Kayne, R. S. 2008a. Expletives, datives, and the tension between morphology and syntax. In T. Biberauer (ed.), *The Limits of Syntactic Variation*, Amsterdam and Philadelphia: John Benjamins, pp. 175–217. [Reprinted in Kayne (2010, Ch. 6).]

Kayne, R. S. 2008b. Some preliminary comparative remarks on French and Italian definite articles. In R. Freidin, C. P. Otero, and M. L. Zubizarreta (eds), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, Cambridge, MA: MIT Press, pp. 291–321. [Reprinted in Kayne (2010, Ch. 1).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Kishimoto, H. 2005. Wh-in-situ and movement in Sinhala questions. *Natural Language and Linguistic Theory* 23: 1–51.

Kuroda, S.-Y. 1968. English relativization and certain related problems. *Language* 44: 244–266. [Reprinted in Reibel and Schane (1969).]

Lasnik, H., and M. Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15: 235–289.

Leu, T. 2008. *The internal syntax of determiners*. Doctoral dissertation, New York University.

May, R. 1977. *The grammar of quantification*. Doctoral dissertation, MIT.

Nishigauchi, T. 1986. *Quantification in syntax*. Doctoral dissertation, University of Massachusetts, Amherst.

Nishigauchi, T. 1990. *Quantification in the Theory of Grammar*. Dordrecht: Kluwer.

Pesetsky, D. 1987. Wh-in situ: Movement and unselective binding. In E. J. Reuland and A. G. B. ter Meulen (eds), *The Representation of (In)definiteness*, Cambridge, MA: MIT Press, pp. 98–129.

Postma, G. 1994. The indefinite reading of WH. In R. Bok-Bennema and C. Cremers (eds), *Linguistics in the Netherlands*, Amsterdam: John Benjamins. pp.187–198.

Reibel, D. A., and S. A. Schane. 1969. *English Relativization and Certain Other Related Problems*. Englewood Cliffs, NJ: Prentice-Hall.

Richards, N. 1996. Subjacency forever. In V. Samiian (ed.), *Proceedings of the 26th Western Conference on Linguistics (WECOL)*, Fresno, CA: Department of Linguistics, California State University, pp. 243–255.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 281–337.

Ross, J. R. 1967. *Constraints on variables in syntax*. Doctoral dissertation, MIT. [Published 1986 as *Infinite Syntax!* by ABLEX, Norwood, NJ.]

Starke, M. 2001. *Move dissolves into Merge: a theory of locality*. Doctoral dissertation, Université de Genève.

von Stechow, A. 1996. Against LF pied-piping. *Natural Language Semantics* 4: 57–110.

Taraldsen, K. T. 1981. The theoretical interpretation of a class of 'marked' extractions. In A. Belletti, L. Brandi, and L. Rizzi (eds), *Theory of Markedness in Generative Grammar. Proceedings of the 1979 GLOW Conference*, Pisa: Scuola Normale Superiore di Pisa, pp. 475–516.

Tsai, W.-T. D. 2003. Lexical courtesy revisited: Evidence from Tsou and Seediq wh-constructions. *Gengo Kenkyu [Journal of the Linguistic Society of Japan]* 123: 331–361.

Tuller, L. 1992. Postverbal focus constructions in Chadic. *Natural Language and Linguistic Theory* 10: 303–334.

Watanabe, A. 1992. Subjacency and S-Structure movement of Wh-in-situ. *Journal of East Asian Linguistics* 1: 255–291.

11

Negation, *Wh*-Movement, and the Null Subject Parameter

Luigi Rizzi

1982

11.1 Introduction

One of the most challenging puzzles that linguists try to solve concerns the tension between the differences and similarities among languages, in particular as it relates to language acquisition. We can see that languages can be very different from one another, and yet we know that a normally developing child can acquire any language that he or she is exposed to, with relative ease while going through the same stages. Languages must, then, be at the same time very different and very similar. How is that possible? The solution to this puzzle, generative linguists argue, lies in the notion of parameter: a parameter is a choice point in the grammar; depending on the choice the grammar of a particular language makes, several properties of that language follow directly as a consequence of that particular setting of the parameter. The existence of parameters simplifies the task of language acquisition: languages are fundamentally similar, and their differences can be reduced to different settings of a finite number of parameters, which the child learns by generalizing over the linguistic input to which he or she is exposed. (For a particularly lucid discussion and characterization of parameters, cf. Rizzi [2011].)

Rizzi's paper represents one of the first extensive explorations of the notion of parameter and is concerned with what is perhaps its best-known example, the null subject parameter. As Rizzi points out, a cluster of properties distinguishes null subject languages (that is, languages that do not require an overt subject in tensed clauses), like Italian, from non-null subject languages, like English: Italian (i) allows the subject to be null in a tensed clause (1a), (ii) allows the subject to occur in postverbal position (1b), and (iii) allows extraction of the subject of the embedded

clause even when that clause is introduced by an overt complementizer (1c), that is to say, it does not exhibit the so-called *that*-trace effect:

(1) a. Verrà. (Italian)
        come.FUT
        'He or she will come.' (Literally: Will come.)
     b. Verrà   Gianni.
        come.FUT   Gianni
        'Gianni will come.' (Literally: Will come John.)
     c. Chi credi che verrà?
        who think that come.FUT
        'Who do you think will come?'
        (Literally: Who do you think that will come?)

English, in contrast, does not allow any of these options, as we see from the literal translations of the Italian sentences, which are all ungrammatical.

Rizzi links these three differences to a single parametric difference between the two languages, which is, intuitively, the ability of inflection to function like a pronoun, in the sense that it provides information concerning the person and number of the subject. Rizzi recasts this non-old intuition by saying that the functional head that carries verbal inflection, INFL, is specified as [+pronoun] in Italian; in English, in contrast, it is not. The proposal (and the essence of the notion of parameter) is that from this single difference in the grammar of the two languages, (at least) the three properties mentioned above follow. Let us briefly see how.

Rizzi assumes that a null subject sentence like (1a) has a subject, though one that lacks phonetic content (represented as *e*, for empty category). This subject, like all phonetically null elements, must be licensed in a way that makes its content recoverable (in this paper, this licensing requirement is expressed by the EMPTY CATEGORY PRINCIPLE of Chomsky (1981), which states that empty categories must be properly governed). The head INFL can only license an empty category in subject position when it is specified as [+pronoun], as in Italian, but not otherwise.

Pronominal INFL can be referential and license a null subject with a definite interpretation, as in (1a); it may also be nonreferential and license an expletive null subject (i.e., a phonetically null counterpart of English *there* in sentences like *There exists a solution*). Rizzi suggests that, in examples like those in (1b), [+pronoun] INFL licenses an expletive null subject in preverbal position, which is formally connected with the postverbal subject and shares with it case and thematic role.

Finally, according to Rizzi, it is the pronominal nature of INFL that makes (1c) possible in Italian. [+pronoun] INFL licenses an expletive null subject in preverbal position; the *wh*-phrase *chi* moves from postverbal position, leaving a trace that is governed by the verb:

(2) [chi$_i$ [credi [che e$_i$ INFL$_i$ verrà t$_i$]]]

Since all empty categories are licensed, the sentence is grammatical (in contrast with English, where INFL is not [+pronoun], thus not able to license the empty

category in preverbal position.) On this view, then, the three differences exhibited by Italian and English can be reduced to a single parameter.

Rizzi's proposal raised a number of questions that stimulated much subsequent research. For example, what is the precise characterization of a pronominal INFL? In other words, when is INFL "rich" enough to license a null subject and the cluster of properties that go with it? Does it need to carry both person and number features, or only the former? Does it need to make (person and number) distinctions in the entire verbal paradigm, or are partial distinctions sufficient? Can this approach be extended to languages outside the Romance family, or the Indo-European group? (cf. Rizzi 1986; Gilligan 1987; Jaeggli and Safir 1989; Rohrbacher 1999; Kayne 2001; Biberauer et al. 2010; among others) How does this proposal apply to languages that do not mark agreement on INFL, like the East Asian languages? (cf. Huang 1984; Neeleman and Szendröi 2007; among others). What is the relation between the overt morphological marking on the verb and the formal properties of INFL? (Bobaljik 2002).

11.2   From "NEGATION, *WH*-MOVEMENT AND THE NULL SUBJECT PARAMETER"

0 In recent years the deepening of theoretical models for syntactic research has fruitfully coincided with a renewed interest in comparative syntax; in this perspective, the identification and analysis of systematic patterns of variation has become one of the fundamental goals of linguistic theory. Universal Grammar has been conceived of as a parametrized system, in which a specific core grammar can be derived by fixing a finite number of parameters: different values assigned to the parameters define different grammars and, in the optimal cases, complex patterns of variation are reduced to minimal differences in the parametric choices (see Chomsky [1981] for detailed discussion).

The present study is devoted to the analysis of one of these cases. It is well known that some of the attested natural languages allow phonetically null subjects in tensed clauses, while others do not; the two types are instantiated by Italian and English in (1). It has been shown in recent work that other properties systematically correlate with the null subject property: first of all, null subject languages (henceforth NSL's) generally have a free process of subject inversion, while non-NSL's generally do not (cf. (2)): secondly, non-NSL's generally show COMP-trace effects, while NSL's generally do not (cf. (3)):

(1) a.   *e* verrà
     b.   *e will come
(2) a.   *e* verrà Gianni
     b.   *e will come Gianni
(3) a.   Chi$_i$ credi che e_i verrà?
     b.   *Who$_i$ do you think that e_i will come?

In the perspective of the research program mentioned above, this systematic pattern of variation will be provisionally called the "Null Subject Parameter". In the last few years several attempts have been made to give (at least a partial) theoretical account of this parameter, first by means of specific stipulations, then in terms of general principles (Perlmutter 1971; Bresnan 1972; Chomsky & Lasnik 1977; Taraldsen 1978; Kayne [1980]; Pesetsky 1979). In this article I would like to propose a partial reinterpretation of the parameter based on Italian data. The two main points of the reinterpretation are the following:

(A) in spite of prima facie evidence, Italian has a COMP-t effect similar to the one found in non-NSL's, which does not show up overtly in cases of *wh*-extraction of the subject, but which is manifested in structures of a different type;

(B) the contrast in (3) can be essentially reduced to the contrast in (2); that is to say, Italian can avoid the COMP-t effect by *wh*-extracting the subject from postverbal position.

[. . .]

1   Some recent analyses

1.1 In the system of Chomsky ([1980]) the general principle which is a natural candidate for a theoretical account of the Null Subject Parameter is the Nominative Island Constraint (NIC), and this possibility is developed in slightly different ways in Taraldsen (1978), Kayne ([1980]), Pesetsky (1979). Chomsky ([1981]) has convincingly shown that the NIC, conceived of as a binding principle for anaphors, was a spurious generalisation, and that its effects should be factored out differently. In the system of Chomsky ([1981]), the general principle which subsumes the empirical coverage of NIC as far as the "null subject" effects are concerned is the Empty Category Principle (ECP), a principle which constrains the possible occurrence of empty categories:

(3)   ECP: "*e* must be properly governed."

The common core of the analyses proposed by Taraldsen, Kayne and Pesetsky can be easily rephrased within a theory which contains ECP instead of NIC. Such a transposition is given in Chomsky ([1981]), and will be sketched out in what follows.

Given very reasonable supplementary assumptions, the "English side" of the parameter follows directly from ECP. Structure (1)b is generated if the subject NP is not further rewritten by the base rules, which can be assumed to be optional. But the subject position of a tensed sentence in English is not a properly governed position (no lexical or coindexed category governs the subject), so that (1)b is ruled out by ECP. The same account holds when the subject position is transformationally vacated via right-ward NP movement (structure (2)b); or via *wh*-movement of the subject across an overt complementizer: assuming for (3)b the following structure, derived via successive cyclic application of *wh*-movement:

(5)   Who$_i$ . . . [$_{S'}$[$_{COMP}$ e_i that] [e_i will come]]

the trace in COMP does not govern the trace in subject position since the c-command requirement is not fulfilled, due to the presence of *that*.

Why does Italian allow apparent ECP violations in (1), (2), (3)? A natural answer is the one suggested by Taraldsen (1978), which I will adopt in the revised form of Chomsky ([1981]): languages may vary with respect to the governing properties of the verbal inflection. In languages like Italian the verbal inflection properly governs the subject NP, so that in the *a* sentences of (1)–(3) there is no ECP violation:

(6)   [$_{S'}$ . . . e_i . . . INFL$_i$ . . .]

In languages like English the verbal inflection is not a proper governer, so that the ECP violation in the *b* sentences of (1)–(3) cannot be rescued in the same way.

This hypothesis is appealing, among other ways, in that it gives precise theoretical status to the common observation that in NSL's minimal feature specifications of missing subjects are "recovered" via the verbal inflection. Whether or not there is a strict (or statistical) correlation between richness of inflectional systems and "null subject" properties can only be decided through an extensive cross–linguistic scrutiny, which is beyond the aims of the present work. Aspects of Italian syntax which bear on this question will be discussed in par. 3.6.

1.2 Kayne (1981a) makes the very important observation that the proposed account of (1)b–(3)b in terms of ECP violation carries over to structures created by rules of the syntax of Logical Form. Kayne's argument goes as follows: In French, the negative quantifier-like NP *personne* (= nobody), as well as other elements belonging to the same class, must be construed with the negative particle *ne*, in some sense to be made precise. If *personne* is in object position of an embedded clause, *ne* can be cliticized onto the embedded verb (as in (7)a) or, somewhat more marginally, onto the main verb (as in (7)b):

(7)   a.   J'ai exigé qu'ils *n*'arrêtent *personne*.
               "I have required that they *neg* arrest *nobody*."
         b.   ?Je *n*'ai exigé qu'ils arrêtent *personne*.
               "I *neg* have required that they arrest *nobody*."

But if *personne* is the subject of the embedded clause, only the sentence with embedded *ne* is acceptable:

(8)   a.   J'ai exigé que *personne ne* soit arrêté.
               "I have required that *nobody neg* be arrested."
         b.   *Je *n*'ai exigé que *personne* soit arrêté.
               "I *neg* have required that *nobody* be arrested."

In short, there is a curious subject–object asymmetry with respect to the possibility of construing *personne* with an occurrence of *ne* in a higher clause.

Kayne notices that (7)a and (7)b clearly differ in interpretation: the quantifier-like negative element *personne* is assigned narrow scope in (7)a (i.e. its scope is the embedded clause only) and wide scope in (7)b: the respective (partial) logical forms, assuming May's (1977) format, would be something like the following:

(9)   a. $[_{S'}[_S$ j'ai exigé$[_{S'}$ que *personne*$_i[_S$ ils arrêtent $e_i]]]$
       b. $[_{S'}$ *personne*$_i[_S$ j'ai exigé$[_{S'}$ qui$[_S$ ils arrêtent $e_i]]]]$

The negative particle *ne* thus seems to play the role of an overt scope marker: the scope of a negative quantifier is the S which immediately contains an occurrence of *ne* construed with the quantifier. Given this analysis, the contrast (8)a–(8)b is automatically explained: the respective LF's would be:

(10)   a. $[_{S'}$ $[_S$ j'ai exigé$[_{S'}$ que *personne*$_i$ $[_S$ e_i soit arrêté]]]]$
        b. $[_{S'}$ *personne*$_i$ $[_S$ j'ai exigé $[_{S'}$ que $[_{S'}$ e_i soit arrêté]]]]$

(10)b is ruled out by ECP, since the empty NP in embedded subject position is not properly governed. On the contrary, both (9)b, (10)a are well-formed with respect to ECP, since in both cases the empty NP is properly governed: in (9)b by the verb (a lexical category), in (10)a by the quantifier itself (a coindexed category).

In conclusion, according to Kayne's analysis there is a true generalisation underlying ordinary COMP-t effects like the one shown in (3)b and such scope phenomena as the contrast just discussed. Although the relevant evidence is not uncontroversial, if the proposed interpretation is correct, these facts constitute a very strong argument for a "syntactic" representation of quantifier scope, as in the theory of LF adopted in several recent works: May's (1977) Quantifier Rule is the LF analogue of *wh*-movement, in that both rules move quantifier-like elements to presentential position, thus creating operator-variable structures. The well-formedness of these structures is then uniformly evaluated by such principles as ECP.

Putting now together Taraldsen's account of the Null Subject Parameter and Kayne's analysis of quantifier scope, we would be led to the conclusion that, ceteris paribus, in a null subject language sentences corresponding to (8)b should be acceptable: in the logical form corresponding to (10)b the empty subject of the embedded sentence would be properly governed by the verbal inflection, and no ECP violation would be produced.

[. . .]

2.3 We can now address the central question formulated at the outset: is Italian different from French w.r.t. Kayne's examples? The answer is "no": exactly the same facts are found, modulo the independent differences between the systems of negation in the two languages. Take for instance the following examples, which correspond in structure to the French examples (7)b–(8)b:

(22)   a.   *Non* pretendo che tu arresti *nessuno*.
              "(I) *neg* require that you arrest *nobody*."

   b.   *Non* pretendo che *nessuno* ti arresti.
        "(I) *neg* require that *nobody* arrest you."

(22)a is acceptable (marginally for some speakers) with *nessuno* construed with the higher occurrence of *non*, i.e., given our previous assumptions, with LF (23):

(23)   [*non* + *nessuno*$_i$ [pretendo[che tu arresti e_i]]
       "There is no x such that I require that you arrest x"

Consider now (22)b: it is a well-formed form (but see fn. 12), and in this respect Italian seems to differ crucially from French (cf. (8)b); but upon more careful consideration it turns out that (22)b is only acceptable in the totally irrelevant interpretation explicitly represented in LF (24): *non* and *nessuno* are construed independently of each other, *non* being the negation of the main clause, and *nessuno* receiving narrow scope interpretation as a negated existential via rule (20):

(24)   [*non* pretendo [che *nessuno*$_i$ [e_i ti arresti]]]
                          [+ neg]
       "I do not require that there be no x such that x arrest you"

The relevant interpretation, with *nessuno* construed with the higher *non*, is impossible:

(25)   *[*non* + *nessuno*$_i$ [pretendo[che e_i ti arresti]]]
         "There is no x such that I require that x arrest you"

Consider also the following examples with *niente* (= nothing):

(26)   a.   Piero *non* crede che Gianni possa fare *niente*.
            "Piero *neg* thinks that Gianni can do *nothing*."
       b.   Piero *non* crede che *niente* possa spaventare Gianni.
            "Piero *neg* thinks that *nothing* can frighten Gianni."

(26)a is interpreted as indicated in (27)a, but (26)b cannot have the corresponding interpretation (27)b: it can only be interpreted with *non* and *niente* construed independently, as indicated in (27)c:

(27)   a.   There is no x such that P thinks that G can do x.
       b.   *There is no x such that P thinks that x can frighten G.
       c.   P doesn't think that there is no x such that x can frighten G.

   The same subject-object asymmetry can be detected (even if the contrast is somewhat less sharp) when *nessuno* is an interrogative polarity existential quantifier. Discussing such cases as (15), we have seen that when the interrogative marker is in the same simple clause, no subject-object asymmetry is found. But if the interrogative marker is one clause up, the asymmetry shows up again:

(28)   a.   Gianni mi ha chiesto *se* pensavo che tu avessi contattato *nessuno*.
          "Gianni asked me *whether* I thought that you had contacted *anybody*."
     b.   Gianni mi ha chiesto *se* pensavo che *nessuno* ti avesse contattato.
          "Gianni asked me *whether* I thought that *nobody* had contacted you."

As the glosses indicate, (28)a can be interpreted with *nessuno* as a wide scope existential, but this interpretation is not allowed for (28)b, where *nessuno* can only be interpreted as narrow scope negated existential (via rule (20)).

In conclusion, putting aside the irrelevant interpretations (24), (27)c, etc., Italian turns out to be just like French in not allowing that a variable in preverbal subject position be bound by a "remote" quantifier of the *nessuno* class. It seems very plausible that this exactly parallel behaviour constitutes a true generalization. But if the solution in terms of ECP proposed for the French cases is to be extended to Italian, then the analysis of the null Subject Parameter presented at the outset must be revised, for the reasons already discussed.

2.4 The next observation is that the impossibility of a "remote" binding solely concerns variables in *preverbal* subject position: if a subject containing elements like *nessuno, niente*, etc. is placed in postverbal position at S-structure, then the wide scope interpretation becomes available again; compare the following pairs:

(29)   a.   *Non* pretendo che *nessuno* sia arrestato.
          "(I) do *not* require that *nobody* be arrested."
     b.   *Non* pretendo che _____ sia arrestato *nessuno*.
          "(I) do *not* require that _____ be arrested *anybody*."

(30)   a.   Piero *non* crede che $\left\{ \begin{array}{l} niente \\ nessuno \end{array} \right\}$ mi possa spaventare.

          "Piero does *not* believe that $\left\{ \begin{array}{l} nothing \\ nobody \end{array} \right\}$ can frighten me."

     b.   Piero *non* crede che _____ mi possa spaventare $\left\{ \begin{array}{l} niente \\ nessuno \end{array} \right\}$.

          "Piero does *not* believe that _____ can frighten me $\left\{ \begin{array}{l} anything \\ anybody \end{array} \right\}$."

(31)   a.   Pensi che *nessuno* si presenterà?
          "(Do you) think that *nobody* will show up?"
     b.   Pensi che si presenterà *nessuno*?
          "(Do you) think that _____ will show up anybody?"

(32)   a.   Mi chiedo *se* Gianni pensi che *nessuno* lo stimi.
          "I wonder *whether* Gianni thinks that *nobody* esteems him."
     b.   Mi chiedo se Gianni pensi che _____ lo stimi *nessuno*.
          "I wonder whether Gianni thinks that _____ esteems him *anybody*."

Sentences *a* and *b* only differ in form (essentially) w.r.t. the pre- *vs* post-verbal position of the subject, but they sharply differ in interpretation: as the English glosses indicate, the *a* sentences can only be interpreted with *nessuno* = narrow scope negated existential; on the contrary, the *b* sentences are interpreted with *nessuno* construed with the negative or interrogative marker in the higher clause. Notice that intuitions concerning the different interpretations of such pairs are sharper and easier to detect than ordinary scope intuitions, since they involve a different number of negations: for each pair, the interpretation of the *a* sentence involves one more negation than the interpretation of the *b* sentence (two *vs* one in (29), (30); one *vs* none in (31), (32)).

2.5 In conclusion, the result achieved so far can be summarized in the following descriptive statement: a variable in preverbal subject position cannot be bound by a "remote" operator of the *nessuno* type; a variable in postverbal (subject or object) position can:

(33) a.   *$Op_i \ldots [_{S'} COMP\ e_i\ V \ldots]$
      b. $Op_i \ldots [_{S'} COMP \ldots V\ e_i \ldots]$

The well-formedness of configuration (33)b is expected within an ECP approach: a base generated postverbal NP position (direct object, etc.) is properly governed by the verb; moreover, Kayne (1981b), Belletti & Rizzi (1981) provide evidence showing that a subject moved to the right ends up in a position governed by the verb (hence, properly governed). Therefore, further leftward movement of the postverbal subject in the syntax of LF does not yield ill-formedness, since ECP is not violated.

The problematic case is configuration (33)a. Its ill-formedness would also follow, within an ECP approach, given the assumption that the preverbal position is ungoverned (or, at least, non-properly governed) in Italian too. But, again, this conclusion conflicts with the analysis of the Null Subject Parameter discussed before, which therefore requires some modification.

[. . .]

3.3.1 Consider now subject inversion. It has been argued elsewhere that the subject moved to postverbal position is adjoined to the VP (Kayne 1981b; Belletti & Rizzi 1981); an instance of derived structure would be the following (I assume that the subject NP position and INFL are coindexed by a general agreement convention):

(48) e_i INFL$_i$ [$_{VP}$[vpha telefonato] Gianni$_i$]
       "e_i INFL$_i$ has telephoned Gianni$_i$"

Several questions arise at this point:

1)   Why is this structure well-formed w.r.t. ECP?
2)   Why is it well-formed w.r.t. the binding principle?

3)   How is it interpreted?

4)   How does the postverbal subject receive case?

Concerning question 1), the answer will simply consist in extending to this case the analysis proposed in the preceding paragraph for the pronominal interpretation: INFL can be specified [+ pronoun]; if it is, it becomes a subject clitic in all relevant respects, and qualifies as a proper governer for the empty subject position.

Question 3) requires an elaboration of the previous analysis: in (48) INFL is not interpreted as a definite pronoun; rather, its role is comparable to "presentational" *there* in English (Guéron 1980; Stowell 1979[)]; how can it perform this function? It seems to me that the phenomenon can be viewed as an instance of a fully general process of formation of "dummy" pronouns: in fact, the possibility of using (subject) definite pronominal forms as "dummies" is a well-attested option for natural languages (consider for instance French *il* or German *es*). Following in essence the proposal of Kayne (1981b) for French, I will assume that this option is formally defined by the operation of assigning a specific feature, say [+ dummy], to a (subject) definite pronoun. Since we have suggested that a pronominal INFL in Italian is assimilated in all relevant respects to a subject clitic, it is natural to expect that assignment of [+ dummy] will be available as well. If this process takes place, the pronominal INFL is interpreted on a par with presentational *there* is English, "dummy" *il* in French, etc., as desired.

Now considering question 4), we notice that this problem does not seem to differentiate NSL's and non–NSL's; in fact, it arises in an exactly parallel fashon [sic] in structures with postverbal subjects in non–NSL's:

(49)   There came a man

(50)   Il est venu un garçon

These structures seem to require a "transmission convention" like the following:

(51)   in the structure
       . . . dummy$_i$ . . . NP$_i$ . . .

       where NP$_i$ is coindexed with and in the domain of
       dummy$_i$, copy the Case of dummy$_i$ on NP$_i$.

This convention can be assumed to be operative in the derivation of (48): the dummy pronominal inflection absorbs nominative Case, and transmits it to the postverbal subject via convention (51). That the transmitted Case is in fact nominative is shown in structures in which the postverbal subject is a first or second person pronoun, in which the opposition nominative/non-nominative is morphologically indicated:

(52)   Ho telefonato $\begin{Bmatrix} \text{io} \\ \text{* me} \end{Bmatrix}$

       "Telephoned I/* me"

       [. . .]

3.4 As anticipated in 3.2 [in the full article], the proposed treatment of the Null Subject Parameter directly solves the problem noticed in section 2, i.e., the fact that quantifiers in the preverbal subject position cannot have wide scope in Italian, exactly as in French, while quantifiers in postverbal subject position can have wide scope. In particular, it follows directly from the preceding analysis that the behavior of quantifiers in preverbal subject position does not differ in Null Subject and non-Null Subject Languages.

Consider again a case like

(63)   . . . pretendo [$_{S'}$ che *nessuno* INFL sia arrestato]
     ". . . require [$_{S'}$ that nobody INFL be arrested]"

At S structure the embedded clause has a lexical preverbal subject (*nessuno*); then, INFL cannot be pronominal: otherwise, it would absorb Case, and no Case would be assigned to the lexical subject, in violation of the Case Filter. Therefore, INFL can never be a proper governer in these structures. It follows then that if *nessuno* is construed with a higher negation, thus receiving wide scope, the derived LF

(64)   *nessuno*$_i$ . . . [$_{S'}$ che e_i INFL sia arrestato]

is ruled out by ECP, exactly on a par with the French equivalent. If narrow scope is assigned to *nessuno* via rule (20), no ECP violation is produced, and the derived LF is well-formed:

(65)   . . . [$_{S'}$ che *nessuno*$_i$ [e_i INFL sia arrestato]]
               [+ neg]

In conclusion, in this system the generalization concerning Italian and French quantifiers is directly captured, and reconciled to the idea that INFL counts as a proper governer in Italian for the "null subject" property.

Take now a sentence in which, at S structure, *nessuno* is placed inpostverbal subject position:

(66)   . . . [$_{S'}$ che e_i INFL sia arrestato *nessuno*$_i$]

[H]ere, independently of properties of INFL, *nessuno* is in the VP, i.e., in a governed position. Therefore, if it is construed with a negation in the higher clause and receives wide scope, no ECP violation is produced. The asymmetry described in (33) is thus given a satisfactory theoretical interpretation with the proposed treatment of the parameter.

  [. . .]

3.6 Let us now turn to the second question raised before, which can be rephrased in the following way: why is the distribution of null definite pronouns more restricted than the distribution of "null dummies", i.e., why are both options available in tensed clauses while only one is available in Aux-to-COMP clauses?

The most obvious difference between the two cases is that only in the first is the inflection specified with grammatical features of agreement (person and number). Intuitively speaking, the situation is rather clear: a pronominal INFL can be referential (= function as a definite pronoun) only if its grammatical specification is structured beyond a certain threshold. We may tentatively assume that one of the discriminating factors is the person specification: i.e., a pronominal inflection (perhaps, a pronominal form in general) can be referential only if specified $[\alpha$ person] (α ranging over I, II, III). Then, the pronominal INFL in gerundival and infinitival Aux-to-COMP clauses, which does not have such a specification, can never be referential, and is only available as a dummy (i.e., is appropriately interpretable only if the optional assignment of [+ dummy] applies).

We are thus led to distinguish two separate properties within the "null subject" phenomenology, which can (somewhat redundantly) be phrased in the following way:

(75)   a.   INFL can be specified [+ pronoun].
       b.   INFL   can be referential.[30]

$$\begin{bmatrix} + \text{ pron} \\ F_i \cdots F_n \end{bmatrix}$$

[. . .] Property (75)b is directly related to the actual morphological richness of the inflection, while property (75)a is not: i.e., the actual morphological richness of INFL turns out to be directly relevant to the interpretation, not to the well formedness of a null subject. For intrinsic reasons property (75)b can only be found in a subset of the cases in which property (75)a holds. According to our proposal, property (75)a can be viewed as the theoretical statement of the Null Subject Parameter: languages may vary in having it or not. As for (75)b we may now wonder how it is related to (75)a across languages: should we expect these two properties to occur always jointly? Or rather are they to be considered two related but autonomous parameters? If the latter is true, the two parameters would a priori define four possible language types:

	1	2	3	4
(75) a.	+	+	−	−
b.	+	−	+	−

We immediately notice that column 3 is excluded for intrinsic reasons (if an inflection cannot be pronominal, it cannot be referential either), that column 1 represents the Italian type, and that column 4 represents the English type. The problem then reduces to the following: is column 2 attested in natural languages? i.e., should we expect to find a language which has phonetically null dummy subjects, but no phonetically null definite pronominal subjects? The answer seems to be "yes". For instance, the fundamental properties of such a case are represented by some northern Italian dialects, as the dialect of Padua, described in Benincà (1980). Restricting our attention to the 3rd person singular, and omitting several details, we find the following situation:

(76) a. Piove
          "Rains"
      b. Vien Giorgio
          "Comes Giorgio"
      c. *Vien
          "Comes"
      d. El vien
          "He comes"

In this dialect the subject of a weather verb is null ((76)a), and when a lexical sub-
ject has been moved to the right, the preverbal subject position can be left without
phonetic realization, as in Italian ((76)b); within our account, this shows that this
grammar has property (75)a. But the phonetically null subject cannot have the
interpretation of a definite pronoun ((76)c): in this case, an overt pronominal sub-
ject must appear ((76)d). Then, the grammar characterizing (76) lacks property
(75)b (at least, as far as III person INFL's are concerned).

Other potential instances of the language type represented in column 3 of the
preceding matrix seem to be so-called Dutch A and Icelandic according to the
description of Maling and Zaenen (1978), following the interpretation of Pesetsky
(1979), who reaches conclusions essentially equivalent to ours. These cases con-
sistently support the view that what has been considered so far as a single param-
eter of UG is to be decomposed into the two subparameters (75)a–b, which are
authonomous [sic] but related by an obvious intrinsic connection: we may assume
that specific grammars freely assign values "yes" or "no" to property (75)a; if the
value is "yes", then a value is assigned to property (75)b; otherwise, the configura-
tion referred to in (75)b simply does not arise.

[. . .]

4 Wh-movement

4.1 We will now turn to the third property identified at the outset as characteristic
of NSL's, i.e., the apparent absence of the COMP-trace effect with *wh*
constructions:

(80) Chi credi che verrà?
      "Who do you believe that will come?"

At first glance, this fact now appears problematic on both the two grounds of the
descriptive generalizations identified, and the proposed theoretical account. From
the descriptive viewpoint we find the following unsatisfactory situation: Kayne has
suggested that *wh* constructions and certain quantified constructions in French
uniformly show COMP-trace effects, and we have seen that the corresponding
quantified constructions in Italian seem to fall under the same generalization; but
such examples as (80) seem to indicate that in Italian *wh* constructions do not show
COMP-trace effects: we would therefore be led to the conclusion that quantified

constructions in French and Italian, and *wh* constructions in French are to be incorporated within the same descriptive generalization (the COMP-trace effect), while *wh* constructions in Italian pattern differently. Such a picture cannot, of course, be dismissed on a priori grounds, but it certainly looks rather implausible.

The problem is even worse if considered from the viewpoint of the theory we have proposed in section 3: as a matter of fact, this theory predicts that the most straightforward derivation of (80), via successive cyclic movement of the embedded subject to the main COMP is ill-formed:

(81) $[_{COMP} \text{ chi}_i]$ credi $[_{S'} \text{ che } e_i \text{ INFL}_i \text{ verrà}]$

There are two cases, according to whether the embedded INFL is pronominal or not. If it is non-pronominal, the structure is ruled out by ECP, as its English counterpart. If INFL is pronominal, then ECP is not violated, since the trace is properly governed; but INFL also binds the empty subject position which, being a *wh* variable, should be free because of clause (C) of the binding principle. Hence, the structure is ruled out by the theory of binding. But, if structure (81) is ill-formed, how can the acceptable sentence (80) be derived? As the reader will easily realize, this technical problem is simply the framework-specific characterization of the general descriptive problem mentioned before, i.e., the apparent asymmetry between quantified constructions and *wh* constructions w.r.t. the COMP-trace effect in Italian.

How can these problems be solved? In fact, a very simple solution is suggested by the asymmetry between preverbal and postverbal subject position w.r.t. wide-scope interpretation: consider again the descriptive statement which summarized the results of our analysis of negative quantifiers:

(82)   a. $^*Op_i \ldots [_{S'} COMP \, e_i \, V \ldots]$
       b. $Op_i \ldots [_{S'} COMP \ldots V \, e_i \ldots]$

The idea I would like to explore now is that (82) also adequately describes the situation concerning *wh* extraction of the subject: i.e., taking seriously the parallelism between scope assignment and *wh* movement, since postverbal subjects only can be extracted from clauses by the scope assignment procedure, it is reasonable to consider the possibility that, in a parallel fashon [sic], postverbal subjects only can be *wh* extracted. This would mean that a sentence like (80) does not have the derived structure (81), but only the structure indicated in (83), derived via movement of the embedded subject to the right in the embedded cycle, and then *wh* extraction (whether successive cyclic or not is irrelevant here) from postverbal position:

(83) $[_{COMP} \text{chi}_i]$ credi $[_{S'} \text{ che } e_i \text{ INFL}_i \text{ verrà } e_i]$

Given this analysis, the descriptive problem of the asymmetry between scope assignment and *wh* movement vanishes: there is no asymmetry, and the descriptive

statement (82) holds true for both cases. Also the framework-specific theoretical problem is solved: structure (83) is derivable and well-formed within the framework proposed in section 3. INFL must be pronominal and [+ dummy], as in all other cases of subject inversion; the preverbal trace in the embedded clause is thus properly governed by INFL, and the postverbal trace is properly governed by the verb, so that ECP is not violated. The preverbal trace (an anaphor) is bound by INFL, and the postverbal trace (a variable) is free, for reasons discussed in detail in par 3.4.2; the requirements of the binding principle are thus fulfilled.

In conclusion, a well-formed derivation is correctly associated to sentence (80) within the proposed system. Adopting this approach, the third distinctive property of null subject languages turns out to be a byproduct of the second: in these languages COMP-trace effects can be circumvented by first moving the subject to postverbal position and then *wh* extracting it. Therefore, the contrast between (3) a and (3)b (repeated below in (84)) would be traced back to the contract between (2)a and (2)b (repeated in (85)):

(84) a.   Chi credi che verrà?
     b.   *Who do you believe that will come?

(85) a.   Credo che verrà qualcuno
     b.   *I believe that will come somebody

This analysis straightforwardly solves the problems formulated at the outset of this section.

   [. . .]

Note

30   Where $F_1 \ldots \ldots F_n$ are morphologically realized features of grammatical specification, including [α person].

11.3   Questions pertaining to Rizzi (1982)

1   Brandi and Cordin (1989) discuss two Italian dialects in which pronominal subjects (so-called subject clitics) co-occur with preverbal lexical subjects. To what extent does this lend support to Rizzi's analysis? What exactly might Rizzi say about how Case is assigned in sentences with both a preverbal lexical subject and a preverbal subject clitic (in that order)?
2   Brandi and Cordin (1989) give an example from the Florence dialect that has the form *Chi ha-egli telefonato?* ('who has-it telephoned?' = 'who telephoned?') To what extent do such sentences support Rizzi's proposal?
3   (Extra credit) Rizzi himself discusses somewhat parallel cases from French, such as *Combien de personnes est-il arrivé?* ('how-many of persons is-it arrived?' = 'how many people arrived?'), containing the expletive subject

clitic *il*. All such cases in French involve intransitive verbs; with transitive verbs all such sentences are impossible. How might this fact about French be relevant to Rizzi's discussion of Italian sentences like *Quante ne sono arrivate?* ('how-many of-them are arrived?' = 'how many have arrived?')?

4   Italian transitive sentences can readily have the subject postverbal, in particular when the object is cliticized, for example in *L'ha fatto Paolo* ('it has done Paul' = 'Paul has done it'). But these are, strictly speaking, possible only if the postverbal subject is focalized. What does this imply for Rizzi's analysis of Italian interrogatives such as *Chi l'ha fatto?* ('who it has done?' = 'who did it?')?

5   If we embed the example from the previous question, we reach a sentence like *Credo che l'abbia fatto Paolo* ('I-think that it has done Paolo'); the corresponding interrogative is then *Chi credi che l'abbia fatto?* ('who do-you-think that it has done' = 'who do you think did it?'). To what extent is this kind of Italian sentence plausibly related to (very) colloquial English sentences like *Which guy do you think that he did it?*?

6   Ordóñez (1998) argues that postverbal subjects in Spanish come about as the result of moving the verb leftward past the subject, rather than, as Rizzi was assuming, coming about as the result of moving the subject itself rightward. If Ordóñez's proposal is right for Italian, too, would Rizzi's analysis of the *that*-trace effect have to be changed?

7   Hebrew (Shlonsky 1997) and Finnish (Vainikka and Levy 1999) both have paradigms in which a first or second person subject pronoun can be left unpronounced, but in which a third person (nonexpletive) pronoun must be pronounced. One language other than Hebrew that is mentioned by Rizzi has a similar property. Which one? (Alternatively: Find another such language not mentioned by Rizzi.)

8   Rizzi makes use at various points of a property common to Italian and French having to do with sentences that would in English be, for example, *John has three* (possible in an appropriate context). The word-for-word counterpart in either Italian or French is not possible, e.g., French **Jean a trois*; rather, a clitic *en* that can be thought of as 'of them' is required – *Jean en a trois*. Is Spanish more like French/Italian or more like English? Whatever the answer, what might be the reason?

9   Thinking of the previous question, are there comparable differences among these languages in the realm of comparatives? If so, what are they? (Extra credit: What might they tell us about the derivation of comparatives?)

10 How strong do you think Rizzi's argument is against the possibility of having short *wh*-movement from preverbal subject position in Italian? Give your reasons.

11 (Extra credit) Rizzi, in part based on Kayne's (1981a) proposal for French, analyzes Italian as having LF-movement playing a central role in the syntax of negation. Kayne (1998) disagrees with both his earlier self and with Rizzi, and proposes an approach to negation using overt (non-LF) movement. To what extent is the difference between LF-movement and overt movement important to Rizzi's analysis of subject–object asymmetries found with negation?

12 Are the asymmetries concerning negation discussed by Rizzi found in languages other than Italian and French? Give relevant examples from at least two other languages.

13 Rizzi takes the null subject property of Italian to be located in its INFL, which in Italian can (optionally) be pronominal [+pronoun]. Find, however, at least one discrepancy between INFL in Italian and pronouns in Italian. (Alternatively: Use another null subject language). (Hint: Follow the consonants.)

14 (Extra credit) If INFL can optionally be [+pronoun] in Italian, might clitics also optionally be [+pronoun]? Is that a plausible way of thinking about clitic doubling? Give your reasons.

15 Is Rizzi's Case transmission operation rightward or leftward? Could it be unspecified for "direction"? Discuss why, or why not?

16 Rizzi proposes that INFL in nonfinite contexts can never be referential. What feature does he use for that?

17 Does the feature from the previous question have morphological realization in Italian? Give some details. (Extra credit: Would we expect it to have morphological realization? Why, or why not?) (Additional extra credit: How do the dialects mentioned in Brandi and Cordin 1989 fit in here?)

References

Belletti, A., and L. Rizzi. 1981. The syntax of *ne*: Some theoretical implications. *The Linguistic Review* 1: 117–154.

Benincà, P. 1980. Il (soggetto) clitico a. Ms., University of Padua.

Biberauer, T., A. Holmberg, I. Roberts, and M. Sheehan (eds). 2010. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press.

Bobaljik, J. D. 2002. Realizing Germanic inflection: Why morphology does not drive syntax. *Journal of Comparative Germanic Linguistics* 6: 129–167.

Brandi, L., and P. Cordin. 1989. Two Italian dialects and the null subject parameter. In O. Jaeggli and K. J. Safir (eds), *The Null Subject Parameter*, Dordrecht: Kluwer, pp. 111–142.

Bresnan, J. 1972. *The theory of complementation in English syntax*. Doctoral dissertation, MIT. [Published 1979 by Garland Publishers, New York.]

Chomsky, N. 1980. On binding. *Linguistic Inquiry* 11: 1–46.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N., and H. Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8: 425–504.

Gilligan, G. 1987. *A cross-linguistic approach to the pro-drop parameter*. Doctoral dissertation, University of Southern California.

Guéron, J. 1980. On the syntax and semantics of PP extraposition. *Linguistic Inquiry* 11: 637–678.

Huang, J. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15: 531–574.

Jaeggli, O., and K. Safir. 1989. *The Null Subject Parameter and Parametric Theory*. Dordrecht: Kluwer.

Kayne, R. S. 1980. Extensions of binding and Case-marking. *Linguistic Inquiry* 11: 75–96. [Revised version of Kayne (1979). Reprinted in Kayne (1984, Ch. 1).]

Kayne, R. S. 1981a. Two notes on the NIC. In A. Belletti, L. Brandi, and L. Rizzi (eds), *Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW conference*, Pisa: Scuola Normale Superiore di Pisa, pp. 317–346. [Reprinted in Kayne (1984, Ch. 2).]

Kayne, R. S. 1981b. ECP extensions. *Linguistic Inquiry* 12: 93–133.

Kayne, R. S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, R. S. 1998. Overt vs. covert movement. *Syntax* 1: 128–191.

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2001. A note on clitic doubling in French. In G. Cinque and G. Salvi (eds), *Current Studies in Italian Syntax: Essays Offered to Lorenzo Renzi*, vol. 59 of *North-Holland Linguistic Series: Linguistic Variations*, Amsterdam: North-Holland, pp. 189–212. [Reprinted in Kayne (2000, Ch. 9).]

Maling, J., and A. Zaenen. 1978. The non-universality of a surface filter. *Linguistic Inquiry* 9: 475–497.

May, R. 1977. *The grammar of quantification*. Doctoral dissertation, MIT.

Neeleman, A., and K. Szendröi. 2007. Radical pro drop and the morphology of pronouns. *Linguistic Inquiry* 38: 671–714.

Ordóñez, F. 1998. Postverbal asymmetries in Spanish. *Natural Language and Linguistic Theory* 16: 313–346.

Perlmutter, D. 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart, and Winston.

Pesetsky, D. 1979. Complementizer-trace phenomena and the Nominative Island Condition. *The Linguistic Review.* 1: 297–343.

Rizzi, L. 1986. Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17: 501–557.

Rizzi, L. 2011. On the elements of syntactic variation. In *Working Papers of Centro Interdipartimentale di Studi Cognitivi sul Linguaggio (CISCL)*, Università di Siena. Distributed by MIT Working Papers in Linguistics.

Rohrbacher, B. 1999. *Morphology-driven Syntax: A Theory of V to I Raising and Pro-drop. Linguistik Aktuell/Linguistics Today* 15. Amsterdam and Philadelphia: John Benjamins.

Shlonsky, U. 1997. *Clause Structure and Word Order in Hebrew and Arabic: An Essay in Comparative Semitic Syntax*. New York and Oxford: Oxford University Press.

Stowell, T. 1979. What was there before there was there. In D. Farkas, W. M. Jacobsen, and K. W. Todrys (eds), *Papers from the Fourteenth Regional Meeting of the Chicago Linguistic Society*, Chicago: University of Chicago, pp. 458–471.

Taraldsen, K. T. 1978. On the NIC, vacuous application and the *that-trace* filter. Unpublished paper, MIT.

Vainikka, A., and Y. Levy. 1999. Empty subjects in Finnish and Hebrew. *Natural Language and Linguistic Theory* 17: 613–671.

12

The Mirror Principle and Morphosyntactic Explanation

Mark Baker

1985

12.1   Introduction

This article addresses the relationship between syntax and morphology, asking whether they are independent or related in any way (an issue also extensively pursued in Baker 1988). For example, if we think of a syntactic operation that changes the grammatical function of a given noun phrase (like passivization) and adds a corresponding morpheme, would we expect the syntactic process (passive formation) and the morphological process (adding a passive morpheme to the verb) to proceed in parallel, or to be independent?

Since there is no a priori reason to think that they should go together, one might logically expect them to be independent. Yet Baker's investigation of agglutinative languages shows that they go hand-in-hand. *The Mirror Principle and Morphosyntactic Explanation* argues that the relative order of morphemes reflects the order in which the syntactic processes take place: the morpheme closest to the verbal root corresponds to the syntactic operation that applies first, and the morpheme farthest from the verbal root to the syntactic operation that applies last. The sequence of the syntactic operations can be read off the morphemes' relative distance from the verbal root.

Why should the operations of syntax and morphology go hand-in-hand? Baker's article answers this question with a simple yet radical claim: they are one and the same process. This means that morphological processes that combine (grammatical function changing) morphemes with a verb do not take place prior to syntactic processes (e.g., in the lexicon), or independently of them, but jointly with them. In other words, the morpheme is added to the verbal root via a syntactic process. For example, if the first syntactic operation that takes place is passivization, a movement operation

will bring together the verbal root and the passive morpheme. If agreement with an NP in subject position occurs next, another movement operation will bring together the complex form consisting of the verbal root and the passive morpheme with the agreement morpheme. As a result, the agreement morpheme will be farther from the verbal root than the passive morpheme. As Baker shows, processes like this one are reflected in a robust body of evidence from agglutinative languages.

Particularly strong confirmation of the parallelism between morphology and syntax comes from Cinque's work, in particular his 1999 book *Adverbs and Functional Heads*. Discussing a large and typologically diverse set of languages, Cinque uses data from the position of adverbs to provide what he suggests to be a universal hierarchy of functional categories that make up the structure of the clause. He also examines the relative positions of affixes that express the same semantic distinctions as the adverbs (namely tense, mood, aspect, etc.) and observes that the order of these elements mirrors the order of the adverbs. Cinque argues that both the adverb order and the affix order reflect the same underlying hierarchical structure, with the linear order of the adverbs reflecting it directly and the linear order of the affixes resulting from local syntactic movement, precisely as predicted by the Mirror Principle.

12.2   From "THE MIRROR PRINCIPLE AND MORPHOSYNTACTIC EXPLANATION"

Pretheoretically, there are processes in languages of the world that have both a syntactic component and a morphological component. An example is the English passive, illustrated in (1):

(1)   a.   The cats chase the mouse every day.
      b.   The mouse is chased by the cats every day.

(1b) differs from (1a) in two ways. First, the NP that bears the patient or "logical object" semantic role appears as the surface direct object in (1a) but as the surface subject in (1b). Second, the main verb in (1b) is morphologically derived from the (stem of the) verb in (1a) by suffixing the *-ed* morpheme. Any complete account of the passive construction will have to encompass both of these aspects, the syntactic and the morphological. On this, all are agreed. How to integrate the two components into a unified account is another matter, however, and differing viewpoints abound regarding which component is primary and which is derived, at what level(s) of representation the two are explicitly related, and so on (for a cross section, see Chomsky (1981), Bresnan ([1982]), Perlmutter and Postal (1977), Marantz [1984]). Part of the reason for this diversity is that the phenomena in and of themselves do not supply a wide enough range of evidence to guide theoretical decisions in this area. This article will shed new light on these issues by considering interactions of these processes in morphologically complex languages. In particular, it will argue that the morphology and the syntax in this class of cases must be two

aspects of a single process. This result in turn will be shown to place strong, substantive constraints on the kind of syntactic framework that should be adopted.

1 The Mirror Principle introduced

Consider the pattern of verbal agreement in the Austronesian language Chamorro (data from Gibson (1980)):[1]

(2)   a.   Man–dikiki'.
           pl–small
           'They are small.'
      b.   Para#u#fan-s-in-aolak i   famagu'un gi as tata-n-niha.
           irr-3pS-pl-pass-spank the children   obl father-their
           'The children are going to be spanked by their father.'
      c.   Hu#na'-fan-otchu siha.
           1sS-caus-pl-eat   them
           'I made them eat.'

The focus of attention here is on the prefix *man-*/*fan-*. Gibson states that this morpheme appears in a simple clause if and only if the clause is intransitive and has a plural subject. (2a) gives a typical example of this situation. The passive structure in (2b) fits with this generalization as well – as long as we take the generalization to refer to surface representation and not to an "underlying" or "semantic" representation. Thus, *fan-* agrees with the plural NP 'children', which is the derived subject, but not the singular NP 'their father', which is the underlying subject. Furthermore, the underlying clause would be transitive, not intransitive as is required for *fan-* to appear. The morphological causative in (2c), on the other hand, leads in exactly the opposite direction. Here *fan-* agrees not with the surface subject of the sentence 'I', which is singular (as shown by the other agreement morpheme *hu*), but rather with 'them'. This nominal is the underlying, semantic subject of the root 'eat', but on the surface it is a direct object. Similarly, the sentence is transitive on the surface, which should disallow *fan-*, whereas the root verb 'eat' is intransitive in this usage. Therefore, we can keep our generalization about the distribution of *fan-*, but this time the generalization must crucially refer to an underlying representation, rather than the surface one.

How do we understand this behavior of *fan-* verbs? As a preliminary step, notice that another factor exactly correlates with these differing syntactic characterizations of the verbal agreement: the differing position of the agreement morpheme in the verb's morphological structure. In (2b), where agreement is with the surface subject, the agreement morpheme occurs outside the passive morpheme, which is between it and the verb root; in (2c) agreement is with the underlying subject and the agreement morpheme occurs inside the causative morpheme, between it and the verb root.

Or consider the following sentences from Quechua, a South American Indian language (data from Muysken (1981)):

(3) a. Maqa–naku–ya–chi–n.
         beat-recip-dur-caus-3S
         'He$_j$ is causing them$_i$ to beat each other$_i$.'
     b. Maqa–chi–naku–rka–n.
         beat-caus-recip-pl-3S
         'They$_i$ let someone$_j$ beat each other$_i$.'

Even though these two sentences contain essentially the same morphemes, they have very different interpretations: in (3a) the semantic subject of the verb root 'beat' and its direct object are understood as being in a reciprocal relationship, whereas in (3b) the causer and the direct object are understood in this way. How are we to explain this difference? Why aren't the interpretations the other way around?

Once again, the key is the morphological structure of the two verbs involved – in particular, the relative order of the causative and reciprocal morphemes. In (3a), where the reciprocal binds the object to the underlying subject, the reciprocal morpheme is inside the causative morpheme, that is, closer to the verb stem. On the other hand, in (3b), where the reciprocal binds the object to the causer, which is the surface subject (as confirmed by the plural agreement in (3b)), the reciprocal morpheme is outside the causative morpheme, farther from the verb stem.
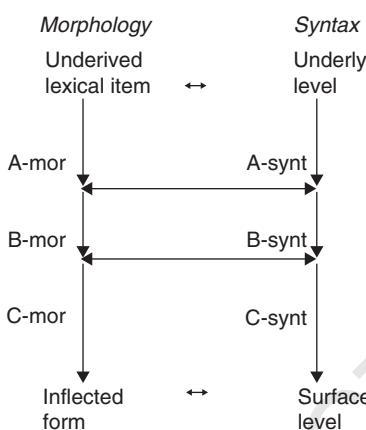
Based on these observations, it seems that these two very different sets of facts can be explained and conceptually unified in terms of a theory of how the morphological and syntactic components are related. Indeed, they are explained by the simple statement that the processes involved – passive, agreement, causative, and reciprocal – simultaneously have morphological effects (such as adding an affix to the verb) and syntactic effects (such as changing grammatical functions). This is not necessary a priori; it is certainly imaginable that Universal Grammar would allow a dissociation of the two, such that each happens independently and the results must be consistent with one another. In fact, two currently influential frameworks, Government–Binding Theory and Relational Grammar, have this property, at least in some cases (see the discussion in section 6 [in the full article]). However, I argue on the basis of facts like those shown above that any framework that does not start by unifying the morphological and syntactic aspects of these processes must in effect do so by stipulating a principle of Universal Grammar that might be stated informally as follows:

(4) *The Mirror Principle*
     Morphological derivations must directly reflect syntactic derivations (and vice versa).
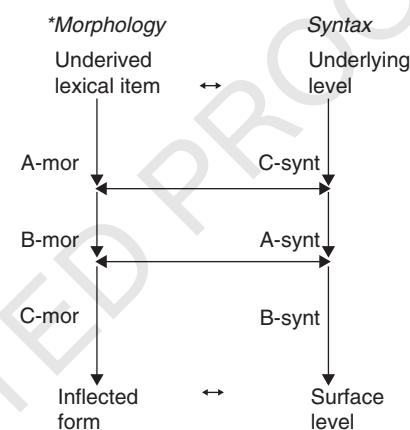
Suppose for illustration that the analysis of a given structure involves three processes, A, B, and C, and that all of these processes have both morphological and syntactic components. Then by (4), the morphological and syntactic derivations must match, as shown in (5a). If they do not, as in (5b), then the structure is ruled out by this principle. (The issues represented here will be developed more fully below.) The form of my argument will be as follows: I will show that,

given independently motivated facts about morphology and syntax taken in isolation, the Mirror Principle explains the observed patterns in Chamorro, Quechua, and many other languages. More than that, it limits the class of possible morphological structures and how they may be related to syntactic structures in a way that seems to be correct universally. Thus, the Mirror Principle is needed to fill a gap in the program of explanatory generative grammar. This will then be interpreted as evidence for a syntactic framework in which morphology and syntax can be directly related to the same processes, because only in this case will the generalizations about language encoded in the Mirror Principle follow naturally.

(5)   a.   b.

Morphology		*Syntax*		*Morphology*		*Syntax*
Underived lexical item	↔	Underlying level		Underived lexical item	↔	Underlying level
A-mor		A-synt		A-mor		C-synt
B-mor		B-synt		B-mor		A-synt
C-mor		C-synt		C-mor		B-synt
Inflected form	↔	Surface level		Inflected form	↔	Surface level

[. . .]

Given a portion of a word of the form (7a),

(7)   a.   . . . verb–affixA–affixB . . .
      b.   . . . [. . . [[verb] affixA] affixB] . . .

then from the ordered, cyclic nature of morphology, we conclude that part of the structure of the word is as shown in (7b) – that is, that affixA is attached before affixB.[6] Furthermore, given the unity of inflection and derivation from a strictly morphological viewpoint, this conclusion is valid for inflectional morphology as well as for derivational morphology, where it is more familiar. Thus, the order in which morphemes appear on the verb reflects the order in which the morphological processes that add those morphemes apply. This then gives the independent content to the notion of a morphological derivation that is needed to make the Mirror Principle meaningful. Specifically, the Mirror Principle claims that the morphological ordering known via the morpheme order must match the syntax (and vice versa). Thus, in example (7) it would claim that the syntactic process associated with affixA must occur before the syntactic process associated with affixB. This is one source of the empirical content of the principle.

   [. . .]

Now consider the general case, where more than one GF-rule or agreement process must be appealed to in the analysis of a given sentence. Since we are purposely focusing on processes that crucially involve grammatical functions, the output of any one process will depend on the GF-structure that it gets as input, which in turn will depend on which (if any) processes have happened before it. To put this another way, these processes stand in potential "feeding" and "bleeding" relationships to one another. Whether we actually observe "feeding" or "bleeding" between the two rules will give us syntactic evidence concerning the order in which they must have applied. For example, consider again Passive and Object Agreement. If a given structure in some language shows no object agreement with the surface subject of a passive sentence, then Passive "bleeds" Object Agreement. Hence, Passive must apply first. On the other hand, if the structure does show object agreement with this nominal, Object Agreement must apply first, because of the uniformity constraint on agreement. In this way, we can establish a syntactic derivation for a given structure in which independently characterizable processes apply in a particular order to account for the properties of that structure. This gives the independent content to the notion of "syntactic derivation" that is needed to make the Mirror Principle meaningful. Specifically, the Mirror Principle now claims that the syntactic ordering known via examination of these feeding and bleeding relationships must match the morphological ordering known independently by examining morpheme orders. Thus, the Mirror Principle will have strong empirical consequences.

3 Interactions between GF-Rules and Agreement

In this section I will show first how the agreement facts from Chamorro introduced in section 1 can be explained using the Mirror Principle and then how the results of that discussion can be generalized to predict a restrictive universal typology of agreement, which is correct over a range of languages.

3.1 Chamorro and *fan-* Agreement

Consider once again the pattern of Chamorro verbal agreement given in (2) (repeated here):

(15)   a.   Man–dikiki'.
            pl–small
            'They are small.'
       b.   Para#u#fan–s–in–aolak i   famagu'un gi as tata–n–niha.
            irr–3pS–pl–pass–spank the children obl father–their
            'The children are going to be spanked by their father.'
       c.   Hu#na'–fan–otchu siha.
            1sS–caus–pl–eat   them
            'I made them eat.'

We have seen that *fan-* normally shows the plurality of the subject in an intransitive clause, as in (15a). However, in passive sentences like (15b) the relevant sense of "subject" is crucially "surface subject," whereas in causative sentences like

(15c) it is crucially "semantic subject." Furthermore, this difference correlates with a difference in morphological structure: in (15b) *fan-* precedes the passive marker *-in-*, whereas in (15c) it follows the causative marker *na'-*. This correspondence led us to posit a direct link between morphological structure and syntactic structure, encoded by the Mirror Principle. We now return to the task of showing that this principle plays an important role in explaining the interactions of Chamorro's agreements and GF-rules, given an understanding of how these processes work individually.

The Chamorro agreements and GF-rules are clearly described by Gibson (1980). I follow her exposition here, translating her generalizations into the notation presented in section 2.

   (i)   *man-/fan-* Agreement. Morphologically, the proper morpheme is simply prefixed to the verb. Syntactically, Gibson states the following generalization: "The prefix *man-/fan-* is attached to the predicate of a finally intransitive clause if and only if the final 1 [= subject] of the clause is plural" (p. 25). In our terms, this can be represented as follows:

   (16)   *Number Agreement (Chamorro)*
          NP1 VERB . . . ®   NP1^i VERB^i . . .
          subj   subj
          Condition: Nothing fills the object slot.

Here, the cosuperscripting relation expresses number agreement, the plural form being *man-/fan-* and the singular form .

   (ii)   Passive. The passive has two morphological shapes, *ma-* and *-in-*, the choice between the two depending roughly on the number of the semantic subject and to some extent on the animacy of the semantic object. *Ma-* is a normal prefix. *-in-*, on the other hand, can appear two ways. Usually it occurs infixed into the stem, placed immediately after the stem's first consonant. If, however, the stem begins with a liquid or a nasal, the affix is metathesized to *ni-* and is prefixed to the verb root. Abstracting away from the details, in all these cases the passive is attached in a position definable only in terms of the beginning of the stem, making it a prefix in a slightly generalized sense. On the syntactic side, the Chamorro passive is essentially identical to its English counterpart, the differences following from independent differences in how the two languages express their surface subjects and objects (Chamorro has VSO word order, with optional fronting of the subject; English is SVO, etc.). [. . .] A typical example of an active-passive pair in Chamorro is given in (17):

   (17)   a.   Si Juan ha#dulalak si   Jose.
               PN Juan 3sS-follow SN Jose
               'Juan followed Jose.'

        b.   D–in–ilalak   si   Jose as   Juan.
             pass-follow PN Jose obl Juan
             'Jose was followed by Juan.'

(iii)   Causative. Morphologically, the Chamorro causative is derived simply by pre-fixing *na'*-. Syntactically, Gibson (1980) provides extensive evidence that the causative produces three changes in grammatical functions: a new NP argument (the "causer") is added as the subject; the (old) subject becomes the object; and, if the verb was transitive, the object becomes oblique. In our notation:

(18)   *Causative (Chamorro)*
       NP1 V (NP2) . . . ®NP3 V NP1 (NP2) . . .
       subj   obj   subj   obj   obl

Typical examples of causative sentences are given in (19) with an intransitive verb root, and in (20) with a transitive verb root:

(19)   Ha#na'-maipi si   Maria i   hanum.
       3sS-caus-hot   PN Maria the water
       'Maria heated the water.'
(20)   Ha#na'-taitai ham   i ma'estru ni esti na lebblu.
       3sS-caus-read 1pex-obj   teacher   obl this   book
       'The teacher made/let/had us read this book.'

Turning now to more complex examples, let us consider the passive sentence (15b). Given the reasoning from section 2.1, the morphological structure of the verb must be as in (21), which implies that the passive morpheme is added before fan-:[16]

(21)   [. . .[fan [in   [saolak]]]]
          pl   pass spank

Now the Mirror Principle comes into play, requiring in this situation that the syntactic effects of Passive must happen before the syntactic coindexing associated with *fan-* is done. In other words, the syntactic derivation must be (22a) and not (22b):

(22)   a.   b.

Hence, Passive must "feed" Number Agreement, which can only register derived subjects in this case.

The same considerations apply to the causative example (15c), but the conclusion is the opposite. This time *fan-* occurs closer to the verb than the GF-rule morpheme does, so it must be added first in the morphological derivation:

(23) [. . .[na' [fan [otchu]]]]
       caus pl   eat

Therefore, by the Mirror Principle, the coindexing associated with Number Agreement must be effected before the GF-changes associated with Causative. Thus, the syntactic derivation must be (24a) and not (24b):

(24) a.   b.

      NP1(pl) VERB   NP1   VERB
        subj   subj
              │ *fan-*   │ Causative
              ▼   ▼
      NP1(pl)^i VERB^i NP3(pl) VERB NP1
        subj   subj   obj
              │ Causative   │ *fan-*
              ▼   ▼
      NP3 VERB^i NP1(pl)^i NP3(pl)^i VERB^i NP1
       subj   obj   subj   obj
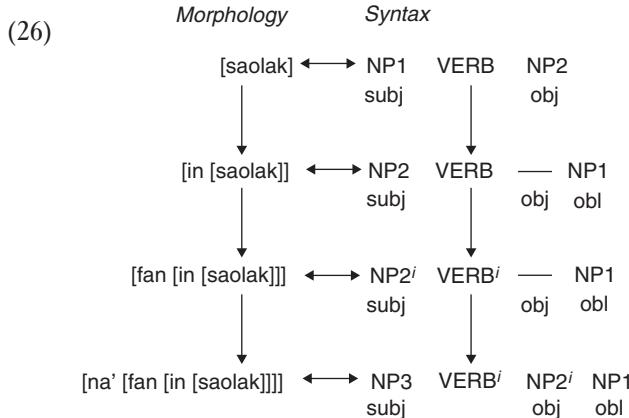
Hence, this time Number Agreement can only register the plurality of the underlying subject and not the plurality of the derived subject.

In fact, these analyses can be combined to explain a still more complicated example. Consider (25), which is the causative of a passive:

(25)   Hu#na'-fan-s-in-aolak i famagu'un gi as tata-n-niha.
       1sS-caus-pl-pass-spank children   obl father-their
       'I had the children spanked by their father.'

Here *fan-* can only be registering the plurality of 'the children', since this is the only plural nominal in the sentence. Yet this nominal is neither the semantic subject nor the surface subject of the verb. It would, however, be an intermediate subject: the subject after Passive has applied but before Causative has. And, not surprisingly, *fan-* appears between the two GF-rule morphemes in the morphological structure, exactly as required by the Mirror Principle. The parallelism between the syntactic and morphological derivations is diagrammed in (26). If the relative order of some process on one side of this diagram were changed while holding the other side constant, the structure would be ungrammatical by the Mirror Principle.

(26)   *Morphology*   *Syntax*

In conclusion, we have been able to account for complex interactions of processes in terms of independent properties of these processes plus a simple constraint on how they are combined: namely, we have explained the apparent idiosyncrasies of *fan-* agreement in Chamorro in terms of the very general Mirror Principle, plus the independently observable facts about where the *fan-* morpheme itself appears relative to other verbal affixes.

3.2   A universal restriction on Agreement processes

Given this explanation for language-particular facts about agreement in Chamorro, the next question is: How general is this account? Is the Mirror Principle a peculiarity of Chamorro, or is it a principle of Universal Grammar? This subsection will explore these questions by examining agreement behaviors in other morphologically complex languages.

A striking thing about Chamorro's *fan-* agreement is its rarity; clear examples of agreement morphemes that can appear intermixed with GF-rule morphemes seem quite unusual. This does not mean, however, that the Mirror Principle has nothing to say about agreement in other cases; only that the range of examples from any one language will be less compelling. Nevertheless, predicting overall trends drawn from a range of languages has a deeper compulsion of its own.

Suppose that the Mirror Principle is part of Universal Grammar. Then the reasoning in section 3.1 concerning the syntactic behavior of *fan-* as a function of its position relative to GF-rule morphemes will hold true in the general case. Thus, if any agreement morpheme X is closer to the verb root V than a GF-rule morpheme Y, then X must be attached to V before Y is (by cyclicity), and by the Mirror Principle the agreement relation between V and its NP will be established before the grammatical functions are changed. Therefore, X will uniformly express agreement with underlying (semantic) grammatical functions. Similarly, if X is farther from V than Y is, then by the Mirror Principle the grammatical functions will be changed before the agreement relation is established, and X will uniformly express agreement with surface grammatical functions. Thus, there are two aspects to any structure that involves both agreement and a GF-rule: the

relative position of the two affixes involved, and whether the agreement references underlying or surface grammatical relations. A priori, these two aspects could vary independently of each other, but the Mirror Principle claims that they cannot. Of the four logical possibilities for combining these two aspects in a single structure, it predicts that only two will be attested in natural language. This is represented in (27), where the permitted combinations are marked + and the prohibited ones *.

(27)

		Syntax	*Morphology*
a.	+	agreement with semantic GFs	agreement is closer to V
b.	*	agreement with semantic GFs	GF-morpheme is closer to V
c.	*	agreement with surface GFs	agreement is closer to V
d.	+	agreement with surface GFs	GF-morpheme is closer to V

In testing whether this restriction induced by the Mirror Principle is correct crosslinguistically, we find first that pattern (27d – surface agreement morphemes farther out from the verb root – is by far the most common among languages of the world. For example, in addition to the rather limited number agreement discussed in section 3.1, Chamorro has a fuller paradigm of person-number agreement, which is related to the tense-aspect system of the language. Gibson (1980) records the following forms:

(28)

	Realis		*Irrealis*	
	sing.	pl.	sing.	pl.
1	hu	ta (incl)	(bai)u	(u)ta (incl)
		in (excl)		(bai)in (excl)
2	un	in	un	in
3	ha	ma	u	u(ma)

These morphemes attach to the beginning of the verb. They always come before (outside) the GF-rule prefixes and refer exclusively to surface grammatical relations, as in the forms cited in section 3.1 and in (29):

(29) a. I famagu'un ma#dulak si Jose.
          the children   3pS-follow PN Jose
          'The children followed Jose.'
      b. Para#u#fan-s-in-aolak i   famagu'un gi as tata-n-niha.
          irr-3pS-pl-pass-spank   the children obl father-their
          'The children will be spanked by their father.'
      c. Hu#na'-podding i   bola.
          1sS-caus-fall   the ball
          'I let the ball fall.'

(29a) is a simple example of this kind of subject agreement; (29b) and (29c) show these agreement markers occurring outside the passive morpheme and the causative morpheme, respectively. In each case, agreement is with the surface subject. In this connection, consider again example (15c) (repeated here):

(30)   Hu#na'-fan-otchu siha.
        1sS-caus-pl-eat   them
        'I made them eat.'

This sentence contains two agreement morphemes (*fan-* and *hu-*), both of which must generally be analyzed as subject agreement morphemes (see for instance (29b)). Yet here the two disagree in features: *fan-* is plural, whereas *hu-* is singular. The reason for this mismatch is by now familiar: *fan-* agrees with the underlying subject, whereas *hu-* agrees with the surface subject, and these two are not the same. The only construction in which the two morphemes can conflict in this way is the causative construction, and the causative morpheme is the only GF-rule morpheme that can appear between the two. Once again, this follows from the Mirror Principle. Similar agreement patterns are found in Turkish, Sanskrit, and Quechua, except that in these languages the relevant morphemes are suffixes rather than prefixes.

   Not all natural language agreements can be interpreted as instances of pattern (27d), however. One that cannot is found in Achenese, an Austronesian language described in part by Lawler (1977). Consider the following active-passive pairs (where (*o*) = *older* and (*y*) = *younger*):

(31)   a.   i.   Gɔpnyan ka gɨ-cɔm   lon.
                    she(o)   perf 3(o)-kiss me
                    'She (older) already kissed me.'
              ii.   Lon ka   gɨ-cɔm le-gɔpnyan.
                    me   perf 3(o)-kiss by-her(o)
                    'I have been kissed by her.'
        b.   i.   Drɔn   nɨ-pajoh   bɔh-mamplam.
                    you(o) 2(o)-eat   fruit-mango
                    'You ate the mangoes.'
              ii.   Bɔh-mamplam nɨ-pajoh le-drɔn.
                    fruit-mango   2(o)-eat by-you(o)
                    'The mangoes were eaten by you.'

Lawler argues that the (ii) sentences are produced by a GF-changing rule of passive, rather than by a "pure" topicalization rule that does not change grammatical functions. In particular, he shows that the preverbal semantic object in these sentences behaves like a surface subject in that it can be raised or controlled if the sentence is embedded under an appropriate matrix predicate:

(32)   Urɨng agam nyan ji-utaha gɨ-pɨreta   le-dɔʔto.
        person male that 3(y)-try 3(o)-examine by-doctor
        'That man (younger) tried to be examined by the doctor (older).'

Furthermore, Lawler shows that the semantic subject cannot appear postverbally in a *le*-phrase when oblique nominals are fronted under topicalization, but only when there is a semantic object present to undergo passivization:

(33) a. Ngən-moto urɨng agam nyan ji-jaʔ tikula.
with–car person male that 3(y)-go school
'With the car, that man went to school.'

b. *Ngən-moto ji-jaʔ tikula le-urɨng agam nyan.
with–car 3(y)-go school by-person male that

Now we observe that in the active sentences (31ai–bi), the verbal agreement morpheme references the NP in the subject position. In their passive counterparts, however, the agreement morpheme agrees not with the subject NP, but with the NP in the postverbal *le*-phrase. Thus, Achenese verbal agreement is unaffected by passivization. Given our assumption that an agreement relation must be uniform at the point where it is established, we conclude that Achenese shows agreement with underlying (i.e. semantic level) subjects. Now, if Achenese had a passive morpheme that appeared closer to the verb than this agreement morpheme does, it would be an instance of pattern (27b), falsifying the Mirror Principle. In fact, the Mirror Principle is not falsified, because Achenese has another unusual property – it has no (overt) passive morphology.

[. . .]

To summarize: As represented in (27), the Mirror Principle predicts that, of four conceivable morphosyntactic relationships between agreement and GF-rules, the languages of the world will attest only two. Of the two allowed patterns, one (namely, (27d)) is quite common and is manifested in languages as diverse as Chamorro, Turkish, Sanskrit, and Quechua, while data from Achenese agreement and Huichol verb suppletion are consistent with the other (namely, (27a)). Chamorro *fan*-agreement shows both of these patterns in different constructions. Finally, I have been able to find no clear case of the disallowed patterns in (27b) and (27c). This supports the claim that the Mirror Principle, or more precisely the generalization described by the Mirror Principle, must be a consequence of Universal Grammar.

[. . .]

Notes

1 Glosses follow these conventions: Person-number agreement morphemes are indicated by Arabic numerals for person (1, 2, or 3), *s* or *p* for singular or plural, and *S* or *O* for subject or object. Simple number agreement is marked *sg* or *pl*. Grammatical function changing rules (GF-changing rules) are marked *pass* (passive), *caus* (causative), *refl* (reflexive), and so on. Other glossed morphemes are generally aspect/tense/mood markers and roughly follow the conventions of their sources. Transcriptions also follow the sources, except that certain diacritics, tone markers, etc., are occasionally suppressed.

6 Again, the sense of "before" intended here is not necessarily a temporal one, but that is by far the easiest way to talk, and I will continue to do so, both in morphology and in syntax.

16 In fact, the derivations in (22b) and (24b) are ruled out for an independent reason – the condition that fan- appears only on intransitive verbs. Thus,

perhaps it is more insightful to start from the two possible syntactic derivations and use the Mirror Principle to derive the morpheme orders instead of the other way around. The situation is symmetrical.

12.3 Questions pertaining to Baker (1985)

1   To what extent is it reasonable to call a language "agglutinative"? Is English an agglutinative language?

2   How do English gerunds and English past participles differ with respect to agglutination? Think about the morpheme *-s* that we see in the English present tense: it is perfectly agglutinative, except when it attaches to *be*, *have*, *do*, and *say*. Why might that be?

3   To what extent is Baker's conclusion that syntax and morphology are strongly parallel reinforced or not by Chomsky's (1995, Ch. 4) bare phrase structure proposal? How does the answer depend on whether you take Merge to apply to words or to morphemes?

4   To what extent is Baker's position consonant with Halle and Marantz's (1993) Distributed Morphology, and with more recent work in that framework?

5   (Extra credit) The "inner agreement" that Baker discusses for Chamorro as occurring between the causative morpheme and the verb root is agreement in number. A plausible conjecture is that such inner agreement for person is never found, in any language. Can you find any counterexamples? If not, why might that be, that is to say, why might the conjecture be valid?

6   To what extent is Huichol suppletive agreement like Chamorro inner agreement? Are they more like French or Italian finite verb agreement or more like French or Italian past participle agreement (cf. Burzio 1986 and Kayne 1985, 1989), and in what respects?

7   Does Chamorro inner agreement (cf. *fan-* p. 000) distinguish between unergatives and unaccusatives? In that respect, how is it similar to or different from Romance past participle agreement (cf. Burzio 1986; Kayne 2000, Ch. 2 and 3)?

8   How does Spanish clitic doubling differ from Huichol suppletive agreement and Chamorro inner agreement? How is it similar to them?

9   In showing (partial) dissociation between person agreement and number agreement, Chamorro and Huichol have something in common with the agreement found in Spanish possessives. What? (Cf. Bosque and Demonte 1999.)

10 (Extra credit) Discuss similarities and differences between Baker's derivations and those given in Kayne (1975, Ch. 6) for French causatives interacting with reflexives and reciprocals.

11 Julien (2002) makes the proposal that the notion "word" is not relevant to syntax. Discuss whether or not that is compatible with Baker's proposal. You may also bring in Myers (1987) and Koopman and Szabolcsi (2000). If these authors are on the right track, what would the implications be for how seriously syntacticians should take conventional orthography, and what, precisely, would remain of the classical notion of morphology?

12 French has both *Cette poupée est petite comme tout* ('this doll is small like anything') and *Cette poupée est toute petite*. In the latter, *tout* seems to be agreeing

with *petite* – both have the *-e* suffix that indicates feminine gender (as in *cette*, agreeing with feminine *poupée*). How might the relation between these two sentences be expressed derivationally? Would *toute* necessarily correspond to a constituent/word?

13   What exactly is essential to the formation of passive sentences? Discuss the relevance of derived nominals (Chomsky 1970) and of French causatives of the *faire . . . par* type (Kayne 1975, Ch. 3).

14   Find kinds of causative sentences in English. In what way(s) are they similar to / different from causatives in Chamorro and Bemba?

15   To what extent, if any, does Collins's (2005) approach to passives require changing Baker's discussion?

16   Cinque (1999) proposes a universal hierarchy of functional heads, interpreting crosslinguistic variation as the result of differences in movement. Discuss how the Mirror Principle fits in with Cinque's proposal.

17   To what extent is Baker assuming that morphemes of the sort he discusses can or cannot be moved around in the course of the syntactic derivation? (Hint: See Baker 1988: 19) Does the Head Movement Constraint (Travis 1984: 131) help or hinder? Would Baker be happy if the language faculty disallowed head movement entirely?

18   What does Greenberg (1966) say about prefixes and suffixes? Extra credit: To what extent does Greenberg (1966) bear on Baker's discussion?

19   (Extra credit) Some languages have agreement on adpositions. In some of these, the agreement morpheme precedes the adposition, in others it follows the adposition. What other property or properties of these languages might this difference between prefixal and suffixal agreement correlate/cluster with?

20   Some verbs have both a prefix and a suffix. Does the Mirror Principle tell us anything about such cases? More generally, does it make a distinction between prefixes and suffixes? Extra credit: What kinds of asymmetries might there be between prefixes and suffixes? (Hint: Look at Kayne 1994, Ch. 4, and Cinque 1999.)

21   How might one proceed if one wanted to show that infixes are invariably reducible to prefixes?

References

Baker, M. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Bosque, I., and V. Demonte. 1999. *Gramática descriptiva de la lengua española* (3 volumes). Madrid: Espasa.

Bresnan, J. 1982. The passive in lexical theory. In J. Bresnan (ed.), *The Mental Representation of Grammatical Relations*, Cambridge, MA: MIT Press.

Burzio, L. 1986. *Italian Syntax: A Government-Binding Approach*. Dordrecht: Reidel.

Chomsky, N. 1970. Remarks on nominalization. In R. Jacobs and P. Rosenbaum (eds), *Readings in English Transformational Grammar*, Waltham, MA: Ginn, pp. 184–221.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York and Oxford: Oxford University Press.

Collins, C. 2005. A smuggling approach to the passive in English. *Syntax* 8: 81–120.

Gibson, J. 1980. *Clause union in Chamorro and in Universal Grammar*. Doctoral dissertation, University of California, San Diego.

Greenberg, J. H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (ed.), *Universals of Language*, Cambridge, MA: MIT Press, pp. 73–113.

Halle, M., and A. Marantz. 1993. Distributed morphology and the pieces of inflection. In K. Hale and S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press, pp. 111–176.

Julien, M. 2002. *Syntactic Heads and Word Formation*. New York and Oxford: Oxford University Press.

Kayne, R. S. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

Kayne, R. S. 1985. Principles of particle constructions. In J. Guéron, H.-G. Obenauer, and J.-Y. Pollock (eds), *Grammatical Representation*, Dordrecht: Foris, pp. 101–140.

Kayne, R. S. 1989. Facets of Romance past participle agreement. In P. Benincà (ed.), *Dialect Variation and the Theory of Grammar*, Dordrecht: Foris, pp. 85– 103. [Reprinted in Kayne (2000, Ch. 3).]

Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Koopman, H., and A. Szabolcsi. 2000. *Verbal Complexes*. Cambridge, MA: MIT Press.

Lawler, J. 1977. A agrees with B in Achenese: A problem for Relational Grammar. In P. Cole and J. Sadock (eds), *Grammatical Relations*. *Syntax and Semantics* 8. New York: Academic Press.

Marantz, A. 1984. *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.

Muysken, P. 1981. Quechua word structure. In F. Heny (ed.), *Binding and Filtering*, Cambridge, MA: MIT Press.

Myers, S. P. 1987. *Tone and the structure of words in Shona*. Doctoral dissertation, University of Massachusetts, Amherst.

Perlmutter, D. (ed.). 1983. *Studies in Relational Grammar* 1. Chicago: University of Chicago Press.

Perlmutter, D., and P. Postal. 1977. Toward a universal characterization of passivization. In K. Whistler, C. Chiarelloet, and R. van Vahn (eds), *Proceedings of the 3rd Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, pp. 394–417. [Reprinted in D. Perlmutter (1983).]

Travis, L. 1984. *Parameters and effects of word order variation*. Doctoral dissertation, MIT.

13

On the Double Object Construction

Richard K. Larson

1988

13.1 Introduction

Predicates that denote an event of transfer typically take three arguments: an agent (the initiator of the event), a theme (the entity undergoing the transfer), and a goal (the end point of the transfer). In some languages, like English, these predicates may realize the theme and the goal arguments in more than one way. The verb can be followed by a direct object expressing the theme and a prepositional phrase expressing the goal (V NP PP), as in (1a), or by two noun phrases (V NP NP), expressing the goal and the theme, respectively, as in (1b):

(1) a. John sent a letter to Mary. ('dative construction')
    b. John sent Mary a letter. ('double object construction')

The question naturally arises of why these two possibilities are available for predicates of this class (which includes *give, tell, show*, among others) and what exactly the relation is between the dative and the double object construction. These issues have been debated, within generative grammar, from the early years (cf. Chomsky 1955/1975, 1981; Oehrle 1976; among others) until the present (cf. Basilico 2008; Michelioudakis 2012; among others). In this paper, Larson puts forth an analysis that views the double object construction as related to the dative construction in a way akin to that in which a passive sentence is related to its active counterpart. In so doing, he suggests a novel way of looking at the internal structure of the VP, and at the structural relation between the verb and its arguments.

    On the Double Object Construction takes as its starting point an important observation made in Barss and Lasnik (1986), namely that, in double object constructions,

the two noun phrases do not symmetrically c-command one another; rather, the NP that is first in linear order c-commands the second one, and not vice versa. Evidence for this claim comes from examples that involve the binding of an anaphor, like those in (2), as well as other contexts that involve a c-command relation between the two noun phrases (such as pronominal binding by a quantifier, weak cross-over, and NPI licensing).

(2) a.   I showed Mary *herself*.
     b.   *I showed *herself* Mary.

This pattern of grammaticality suggests that, in double object constructions, the first of the two objects NPs (*Mary* in (2a)) c-commands the second one (the anaphor *herself* in (2a)), but not vice versa.

   Larson notes that the lack of symmetry shown by the two objects in the double object construction is also exhibited by the NP and PP in the dative construction:

(3) a.   I presented/showed Mary to *herself*.
     b.   *I presented/showed *herself* to Mary.

This leads him to assume that, in both cases, the first NP immediately following the verb in linear order asymmetrically c-commands the second one. Assuming Baker's (1985) *Uniformity of θ-assignment Hypothesis* ("Identical thematic relations are represented by identical structural relations between the item at D-Structure"), Larson suggests that the two constructions start out with the same syntactic configuration, and are related by a movement operation, DATIVE SHIFT.

   The structure that Larson suggests is novel in that it involves a layered VP:

(4)   [$_{VP}$ John V [$_{VP}$ a letter V to Mary]]

The outer layer, or "shell," consists of a V that has the agent argument in its specifier and a VP as its complement; this second VP has a V head, the theme argument as its specifier and the goal as its complement. In the dative construction, this configuration yields the liner order (subject) V NP PP, because the lexical verb raises from the head of the inner VP to the head of the higher VP. In the double object construction, in addition to verb movement, the goal argument raises to the specifier of the inner VP, and the theme argument is demoted to a V' adjunct – similarly to cases of passivization, where the object raises to the specifier position usually occupied by the subject, and the subject occurs in a lower structural position.

   The novel view of the VP put forth in *On the Double Object Construction* has led to new ways of thinking about a number of issues and given rise to fruitful debate. For example, having the V and the PP form a constituent allows Larson to analyze "Heavy NP Shift" not as rightward movement of the heavy NP, as was standard at the time, but rather as the result of the constituent sister of the object (V PP)

raising to its left. This style of analysis has been extended to a wider range of rightward dependencies in subsequent work (Kayne 1994). Larson's proposal that the subject is introduced into syntactic structure in a separate projection from the object resonated with proposals about the realization of thematic structure in syntax (cf. Marantz 1984). It was further developed in Chomsky (1995) and Kratzer (1996), who connect this higher verbal head to object case assignment, and label the higher layer vP and VoiceP, respectively. Pylkkänen (2008) extends this further, by using a variety of heads to introduce verbal arguments. Larson's VP shell proposal leads to a view of constituency within the VP that conflicts with some of the traditional tests of constituency. These constituency paradoxes are explored in Pesetsky (1995), who advocates multiple parallel representations, and Phillips (2003), who suggests that constituency tests can reflect intermediate stages produced within a top-down view of syntactic derivation. Finally, the existence of a transformation relation between the dative and double object constructions that Larson adopts from earlier work continues to be debated: Pesetsky (1995) and Harley (2003), for example, build on the insights of Larson's paper, keeping the hierarchical structure that captures the asymmetry between the first and second NPs, but suggest that the two differ in the composition of the VP and are not derivationally related.

13.2   From "ON THE DOUBLE OBJECT CONSTRUCTION"

Barss and Lasnik (1986) discuss certain asymmetries in double object constructions, such as (1a,b):

(1)   a.   John sent Mary a letter.
      b.   I promised Felix a new set of golf clubs.

They observe facts suggesting that the second NP (*a letter, a new set of golf clubs*) is in each case in the domain of the first, but not vice versa. These results conflict with standard views about the syntax of double object sentences and appear to raise problems for the view that "α is in the domain of β" should be explicated in terms of the structural notion of c-command.

   In this article I present an analysis of the double object construction that implements a proposal about dative structure first suggested by Chomsky (1955/1975). According to this view, a simple dative like *John sent a letter to Mary* derives from an underlying form in which the verb and its indirect object make up a constituent that excludes the direct object. The specific proposal adopted here is that dative complement constructions like *John sent a letter to Mary* involve an underlying clauselike VP whose "subject" is *a letter* and whose "object" is *(to) Mary* (2a); this inner constituent is obscured at S-Structure by an operation of V Raising (2b):

(2)   a.   John [$_{VP}$ a letter [$_{V'}$ send to Mary]]
      b.   John send [$_{VP}$ a letter [$_{V'}$ t to Mary]]

With this view of dative complementation, double objects can be syntactically derived by a modern form of Dative Shift. In particular, they can be produced by applying the familiar operations responsible for passive sentences within VP. The former indirect object (*Mary*) becomes a derived VP "subject," and the former direct object (*a letter*) assumes adjunct status within V′. As I show, the resulting structure accounts for Barss and Lasnik's facts straightforwardly in terms of c-command and provides insight into various other properties of the double object structure as well.

After briefly reviewing Barss and Lasnik's observations in section 1, I introduce the account of dative complementation adopted here in section 2. In section 3 I present a derivational account of double objects that identifies Dative Shift as Passive, and I show that apparent surface differences between the two operations (morphological marking, Case assignment, and so on) are independently explainable. In section 4 I argue for the connection between Passive and Dative Shift with data from indirect passives and psych-verb constructions, and in sections 5 and 6 I examine English-internal and crosslinguistic constraints on Dative Shift. Finally, I conclude, in section 7, with a discussion of the VP complementation structures that play a central role in this account.

1   Asymmetries of syntactic domain

Barss and Lasnik (1986) point out a number of important asymmetries in the behavior of the two objects in double object constructions. All involve phenomena in which constituent structure relations – specifically, c-command – have been assumed to play a central role. Thus, reflexives and reciprocals (anaphors) must be c-commanded by their antecedents. Double object structures show an asymmetry with respect to the licensing of anaphors:

(3)   a.   I showed Mary herself.
          *I showed herself Mary.

A quantifier must c-command a pronoun at S-Structure if it is to bind it. Double objects show asymmetries regarding quantifier-pronoun binding possibilities:

(3)   b.   I gave every worker$_i$ his$_i$ paycheck.
          *I gave its$_i$ owner every paycheck$_i$.

A *wh*-phrase c-commanded at D-Structure by an NP containing a pronoun cannot be moved over that NP if *wh*- and the pronoun are coreferential. This is the so-called weak crossover effect. Double objects show weak crossover asymmetries:

(3)   c.   Which man$_i$ did you send his$_i$ paycheck?
          *Whose$_i$ pay did you send his$_i$ mother?

A *wh*-phrase cannot in general be moved over another *wh*-phrase that c-commands it (in other words, is "superior" to it) in underlying representation. Double objects show an asymmetry in superiority effects:

(3) d. Who did you give which paycheck?
         *Which paycheck did you give who?
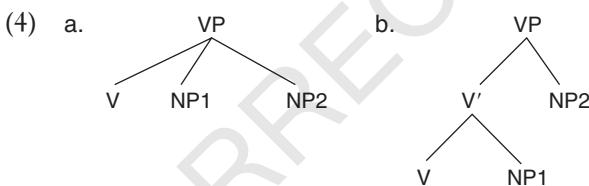
Constructions of the form *each . . . the other*, as in *Each man saw the other* or *Each man saw the other's friend*, may have a reciprocal reading when and only when the *each*-phrase c-commands the *other*-phrase. Double objects show asymmetries with respect to the *each . . . the other* construction on its reciprocal reading:

(3) e. I showed each man the other's socks.
         *I showed the other's friend each man.

Finally, negative polarity items must occur in the c-command domain of an "affective element" such as negation or a negative quantifier. Double objects show asymmetries with respect to a negative polarity item such as *any* and a licensing affective element:

(3) f. I showed no one anything.
         *I showed anyone nothing.

If it is assumed that these phenomena do indeed involve c-command, then (3a–f) all point to the same conclusion: in constructions involving a verb phrase of the form V–NP–NP, the first NP c-commands the second, but not vice versa. As Barss and Lasnik observe, this immediately casts doubt upon the two most frequently assumed structures for double objects:

(4) a.
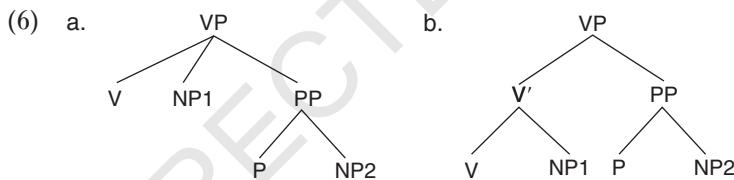
(4a) is the structure for double objects proposed by Oehrle (1976); (4b) is the one proposed by Chomsky (1981). Under a definition of c-command based on first branching nodes (Reinhart (1979)), NP1 and NP2 mutually c-command each other in (4a); hence, this structure predicts no asymmetries in relations based solely on hierarchical structure. In (4b) NP2 asymmetrically c-commands NP1, predicting that the latter is in the domain of the former but not conversely. Both sets of predictions are strongly contradicted by the facts in (3). Under a definition of c-command based on containment in maximal projections (Aoun and Sportiche (1983)), NP1 and NP2 will mutually c-command each other in both (4a) and (4b), predicting no asymmetries of syntactic domain. Again, this prediction is falsified by the data in (3). Evidently one of two conclusions is possible: (a) the syntactic data noted above are not in fact to be explicated by c-command alone; some other notions (such as linear precedence) must be invoked; or (b) these facts are indeed structural and some configuration other than (4a) or (4b) is involved.

    The situation with double objects contrasts with that of standard oblique dative structures. The asymmetries observed with V–NP–NP structures occur with V–NP–PP structures as well:

(5)   a.   I presented/showed Mary to herself.   (anaphor binding)
             *I presented/showed herself to Mary.
       b.   I gave/sent every check$_i$ to its$_i$ owner.   (quantifier binding)
             ??I gave/sent his$_i$ paycheck to every worker$_i$.
       c.   Which check$_i$ did you send to its$_i$ owner?   (weak crossover)
             *Which worker$_i$ did you send his$_i$ check to?
       d.   Which check did you send to who?   (superiority)
             *Whom did you send which check to?
             (*To whom did you send which check?)
       e.   I sent each boy to the other's parents.   (*each . . . the other*)
             *I sent the other's check to each boy.
       f.   I sent no presents to any of the children.   (negative polarity items)
       I.   *sent any of the packages to none of the children.

In the case of oblique datives, however, these results do not appear to raise any special problems for c-command. The facts are accommodated smoothly, it seems, by appealing to the structure introduced by PP. Suppose the VPs in (5) are as in (6a) or (6b):

(6)   a.   b.

     [. . .]
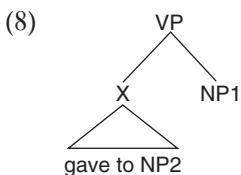
Then in (6a) NP1 asymmetrically c-commands NP2 under the definition of c-command proposed by Reinhart (1979): NP2 is dominated by a branching node (PP) not dominating NP1. Similarly, in (6b) NP1 asymmetrically c-commands NP2 under the definition of c-command given by Aoun and Sportiche (1983): NP2 is contained in a maximal projection (PP) that fails to contain NP1. This illustrates quite clearly why double objects present such a puzzle for syntactic analysis: if complement asymmetry in standard datives is simply a matter of the structure introduced by PP, then why, in double object constructions, where such structure is absent, do we not find symmetric behavior? This is what we expect, but it is not what we see.

2   The structure of datives revisited

Chomsky (1955/1975) proposes an interesting alternative analysis of dative structures, one that attributes asymmetry among complements to a source other than PP structure. According to Chomsky, a sentence like (7a) is actually derived from a structure of the form (7b) (by extraposition of the PP *to him*):

(7) a. The teacher gave several books to him.
     b. The teacher [gave to him] several books.

In (7b) the indirect object is in fact an "inner object" forming a constituent with the verb that excludes the surface direct object. Here, as in (6a,b), there is an underlying asymmetry between dative verb complements. The indirect object (NP2) is in the structural domain of the direct object (NP1), but not conversely:

(8)

However, in (8) this asymmetry is *not* a matter of PP structure. Rather, it is introduced by the branching node labeled X, presumably some projection of V.

The structure in (8) departs quite sharply from those in (6) under plausible assumptions about the relation between thematic and hierarchical structure. In (6a) the two complements are structurally on a par: both are sister to V. We may take this to correspond to the view that theme and goal are both assigned by V (the latter perhaps with some contribution by the preposition). In (6b) the direct object alone is sister to V, whereas the indirect object phrase is sister to the small verbal constituent V′. This we can take to correspond to the view that the direct object alone receives a θ-role directly from V, whereas the indirect object receives its role "compositionally" from V′. Structure (8) asserts what is in effect the inverse of (6b): here it is the indirect object that is the direct argument of the verb, the object NP1 receiving a θ-role from the "phrasal verb" *give to him*.

The position taken in Chomsky (1955/1975) can be supported, I believe, by arguments parallel to those given in Marantz (1984) for the claim that it is VP that assigns a θ-role to the matrix subject, and not simply V. Marantz observes that the predicate expressed by a transitive verb+object regularly depends on the contribution of the object, as shown by VPs like *throw a baseball, throw support behind a candidate, throw a boxing match*. A similar phenomenon can be observed with datives. For example, consider the pair (9a,b):

(9) a. Beethoven gave the Fifth Symphony to the world.
     b. Beethoven gave the Fifth Symphony to his patron.

Giving an object to the world (to posterity, mankind, etc.) has a rather different character from giving an object to an individual. In the first case we understand the given object to be the Fifth Symphony qua composition; the transfer of possession is metaphorical, so that (9a) is roughly synonymous with 'Beethoven created the Fifth Symphony'. In the second case we understand a physical object to be transferred – perhaps a sheaf of papers on which the composition is transcribed.

The exact semantic role assigned to the direct object thus depends on the nature of the recipient appearing in the goal phrase.

The idea that a verb and its outer complements can form a single thematic complex is also supported by the existence of "discontinuous idioms" of the following kind (noted in Emonds (1972)):

(10) a. Lasorda *sent* his starting pitcher *to the showers*.

b. Mary *took* Felix $\begin{Bmatrix} to\ the\ cleaners \\ to\ task \\ into\ consideration \end{Bmatrix}$.

c. Felix *threw* Oscar *to the wolves*.
d. Max *carries* such behavior *to extremes*.

Evidently in (10a) the dative verb *send* assigns a thematic role to the object *his starting pitcher* in concert with the complement phrase *to the showers*; similarly for (10b–d). The possibility of such idioms is straightforward under the structure in (8), where the indicated elements form an underlying constituent. It is quite unexpected under the structures in (6), however, where V and the outer complement form no thematic complex.

The argument from idiom data appears at first to be compromised by examples like (11a–d) (pointed out to me by D. Pesetsky), which seem to involve verb + object idioms (*give x's all, give hell, give the boot, give the creeps, show x's cards*) that assign a compositional role to the indirect object:

(11) a. Max gave his all to linguistics.
      b. Alice gives hell to anyone who uses her training wheels.
      c. Oscar will give the boot to any employee that shows up late.
      d. The Count gives the creeps to anyone he's around long enough.
      e. Phyllis should show her cards to other group participants.

However, on closer inspection it is not clear that such examples really raise a problem. Note that the standard entailment X-*give*-Y-*to*-$Z \Rightarrow Z$-*get*-Y is preserved with the examples (11a–d), and note further that under this entailment the original idiomatic force is preserved:

(12) a. Linguistics gets [my all].
      b. I caught/got [hell] from Alice.
      c. Peter got [the boot].
      d. Geez, you get [the creeps] just looking at him.

Similarly for (11e); alongside this example we find (12e), where again the idiomatic force is preserved:
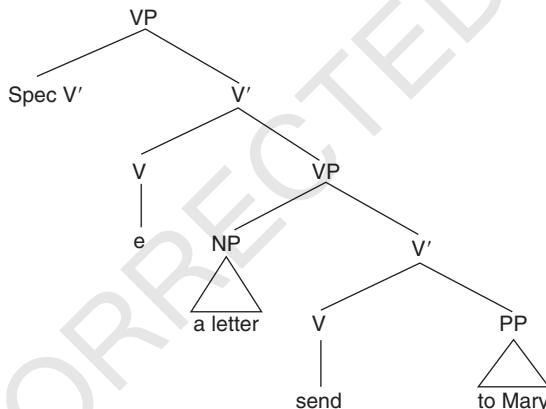
(12) e. Unwittingly, Alice tipped [her cards].

These results are unexpected on the view that verb + object is an idiomatic complex; for example, if V were being understood idiomatically in (11a–d), there would surely be no expectation that the entailment *X-give-Y-to-Z ⇒ Z-get-Y* would hold, as it clearly does. What these facts suggest, then, is that contrary to initial impressions, the idiomaticity in (11a–e) lies not in the verb + object combination but rather in the object alone. That is, (11) and (12) suggest that *one's all, hell, the boot*, and so on, are being treated by the grammar as rather strange sorts of objects that, because they can be given, can be gotten as well. On this view, *give* and *show* do not in fact form idiom complexes in (11) or (12); rather, they simply interact compositionally with a semantically opaque NP.

2.1   V Raising

In analyzing the structure of double objects, I will adopt a version of Chomsky's (1955/1975) proposal, one deriving from work by Bach (1979), Dowty (1979), and Jacobson (1983; 1987). The basic assumption is that the VP in a dative is as illustrated in (13):
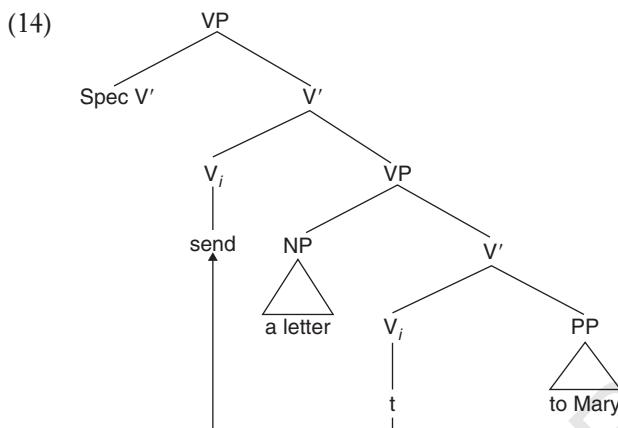
(13)   a.   John sent a letter to Mary.
        b.

According to (13b), the verb phrase underlying *send a letter to Mary* is a strictly binary branching structure. The VP consists of an empty V taking a VP complement whose specifier is *a letter*, whose head is *send*, and whose sole complement is the PP *to Mary*. This structure may be understood intuitively as follows: *send* takes the complement *to Mary*, forming a small predicate send-to-Mary as in Chomsky (1955/1975). The latter is predicated of an "inner subject" *a letter*, forming a VP with clauselike structure: *a letter send to Mary*. This VP is then in turn predicated of a subject like *John* to yield the full sentence (13a).

Of course, *John a letter send to Mary* is not a well-formed sentence of English: the verb must appear to the left of *a letter*. The central assumption here is that the correct surface form arises by movement of the verb *send* to the empty V position – that is, head-to-head movement along lines discussed by Baker (1985) and

Chomsky ([1986]). This movement leaves a trace in the original site and creates a sequence of coindexed V positions:

(14)

V Raising may be taken to follow from certain Case and agreement requirements holding of Infl, V, and NP. Suppose, following the general proposals of Roberts (1985), that V must ultimately head a projection governed by Infl in order to receive tense and agreement information. Furthermore, suppose (following Stowell (1981), Travis (1985), and Koopman (1985)) that Case is assigned under government, where the direction of government is rightward in English. In (13b) V is not the head of a projection governed by I. Moreover, the NP *a letter* in the lower SpecV′ is not governed by the verb and so cannot receive Case. V may be seen as raising in (14) to meet these joint requirements. In the resulting configuration the VP headed by *send* is governed by Infl. Furthermore, V may be plausibly analyzed as governing *a letter*: V is to the left of NP and NP is the specifier of a maximal projection sister to it; hence, *send* can assign Objective Case to *a letter* in (14), as required.

[. . .]

3   The structure of double object constructions

With the account of dative constructions developed above we now return to double object structures. I will argue that domain asymmetries and various other properties of this construction can be explained under a derivational approach to double object structures.

[. . .] In particular, in languages with so-called applicative constructions (see Marantz (1984), Baker (1985) for discussion) oblique and double object structures show a highly productive relation strongly suggestive of derivational relatedness. This argues that transformational operations similar to "Dative Shift" must be available in principle. Second, a derivational approach to the dative–double object relation is clearly desirable under any strong theses about the relation between structure and assignment of thematic roles. For example, Baker (1985) advances the following hypothesis:
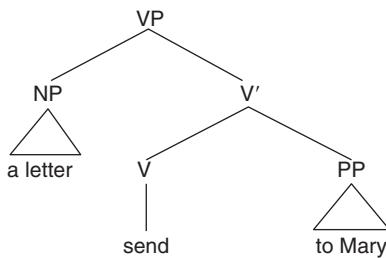
Uniformity of θ*–Assignment Hypothesis*
Identical thematic relationships are represented by identical structural relations between the items at the level of D-Structure.
[. . .]

3.1 "Dative Shift" as passive

These considerations establish a prima facie case for attempting to construe the dative–double object relation transformationally. The challenge, then, is to bring this derivation within the scope of established theoretical principles and to constrain it in appropriate ways. Recall the underlying VP of a typical dative like *John sent a letter to Mary*:
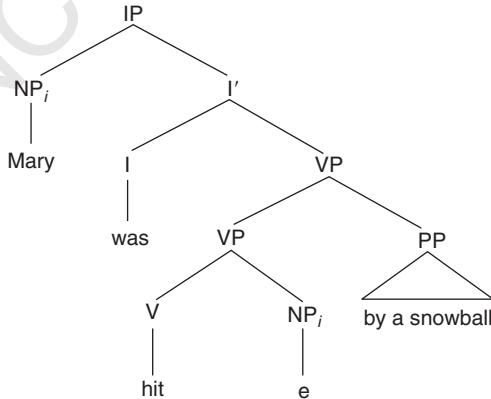
(23)

The deep VP is clauselike, with the NPs *a letter* and *Mary* standing roughly in the relation of subject and object. Suppose we strengthen this parallel by assuming that the governed preposition *to* appearing in (23) has the status of (dative) Case marking, analogous to that appearing on indirect objects in more highly inflected languages.

Consider now the possibility of extending operations generally held to apply between subjects and objects to structures like (23). In particular, consider the possibility of passive formation in the inner VP. Under familiar proposals, the derivation of passives involves two central effects: withdrawal of Case from an object position, and suppression of thematic role assignment to a subject position (see Burzio (1986), Chomsky (1981)). This triggers NP Movement to subject position. The suppressed subject θ-role is (optionally) realized by an adjunct phrase:

(24)

Suppose we amend this account slightly in the following way: rather than assuming that a subject θ-role is suppressed in passives, we will assume that it is assigned in a special way – specifically, in an adjunct configuration:
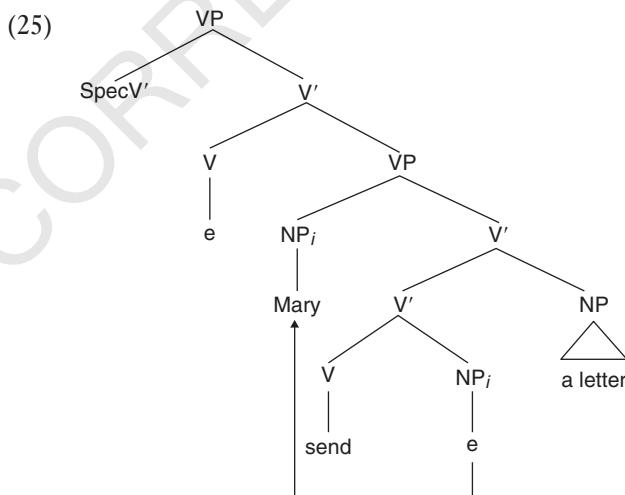
Argument Demotion
If α is a θ-role assigned by X^i, then α may be assigned (up to optionality) to an adjunct of X^i.
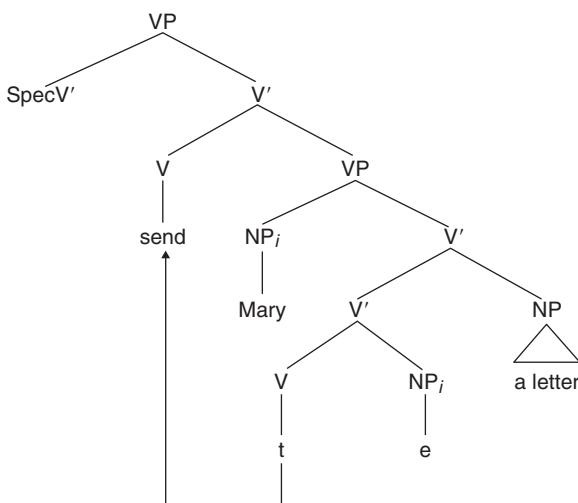
This modification leaves the analysis of (24) unchanged. The IP subject receives its thematic role compositionally from VP; hence, when the subject θ-role is demoted in a passive and is assigned to the *by*-phrase, the latter appears adjoined to VP.

Let us apply this amended view of Passive to *send* as it occurs in the inner VP in (23). First, Passive absorbs the Case assigned to the indirect object. Assuming that we can regard the preposition *to* governed by *send* as pure Case marking, this amounts to saying that *to* is absorbed. Second, the θ-role assigned to the subject of VP (the direct object role) undergoes demotion, reducing this position to nonthematic status. Since the direct object receives its θ-role from V′, under Argument Demotion this θ-role must be assigned to a V′ adjunct. Accordingly, the direct object is realized as a V′ adjunct. The situation is thus as follows: the indirect object is Caseless in its deep position, and the VP subject position is nonthematic (and hence empty). In the usual way, then, the indirect object undergoes NP Movement to the VP subject position (25).

[. . .]

(25)

(26)

[. . .]

13.3 Questions pertaining to Larson (1988)

1 Discuss the implications, for Larson and in general, of the similarity that holds between *Mary always shows to Susan her papers (vs. *Mary always shows her papers to Susan*) and *Of course Mary with care revises her papers (vs. *Of course Mary revises her papers with care*). Bring in Belletti and Rizzi (2012).

2 What are some of the key issues that distinguish Larson's movement approach to double object sentences from Pylkkänen's (2008) nonmovement approach? Bring into your discussion den Dikken (1995).

3 What are some of the key issues that distinguish Larson's movement approach to double object sentences from the alternative movement approach suggested in Kayne (2008), based on Szabolcsi (1983/84, 1994)?

4 Larson does not share the following judgments reported by O'Grady (1980), who finds passive sentences like? *The Red Cross has been donated tons of money this year* to be appreciably less bad than the corresponding actives such as *People have donated the Red Cross tons of money this year*. Discuss the implications for Larson's analysis of the existence of a subset of English speakers who share O'Grady's judgments.

5 In what respects would Larson's analysis have to be modified if it were to turn out that head movement is not allowed by the language faculty? (Extra credit: What might, in that case, underlie the absence of a head movement option?)

6 In what respects would Larson's analysis have to be modified if Kayne (1999, 2004) were to turn out to be right in proposing that prepositions are merged, not with what we think of as their objects, but rather outside VP?

7 Discuss the tension that appears to hold between Larson's analysis of psych-verbs, based on Belletti and Rizzi (1988), and Baker's (1993) analysis of the widespread absence of unaccusative double object verbs.

8   Psych-verb sentences like *That has been worrying John a lot* feel close to sentences like *That has been causing John a lot of worry*, as well as to *That has been giving John reason to worry*. How might such links be expressed precisely? Bring in Pesetsky (1995) on nonagentive causers. (Extra credit: What is the relation in general between the verbs *cause* and *give*?)

9   A double object sentence like *Staying out in the rain has given John the flu* is appreciably more natural than its PP counterpart *?*Staying out in the rain has given the flu to John*, as opposed to the fully natural *Mary has shown her papers to John*. Larson discusses this kind of contrast in primarily interpretive terms. How, thinking of Collins (2008), might one recast Larson's discussion in terms of a difference between *to* by itself (as in the *flu* example) and *to* accompanied by a second, silent preposition?

10 The marginal *flu* example of the previous question improves if the indirect object is made heavy, as in ?*Staying out in the rain has given the flu to most of the children who had been playing outside*. How might this fact be related to the contrast between *The cars of John are very expensive* and ?*The cars of most of the students who attend this school are very expensive*?

11 Discuss the extent to which Larson's account of double object sentences that have a close paraphrase with *to* can be extended to double object sentences that have a close paraphrase with *for*.

12 Although English has both *Mary took the books to Susan* and *Mary took the books from Susan*, the sentence *Mary took Susan the books* can correspond only to the former, i.e., to the one with *to*. Why might that be? (Bring in Landau 2002 and Baltin 2009.) (Extra credit: Compare the difference between *to* and *from* to the difference between *with* and *of* seen in *They provided us with lots of money* and *They deprived us of lots of money*.)

13 How would Larson's analysis have to be revised if Collins (2005) is right in his analysis of passives? Bring in the question of passive-like *-able* sentences such as *This book would even be readable by a ten-year-old*.

14 Larson discusses the fact that passives such as *That book was given him for his birthday*, with a pronominal indirect object, are more widely accepted than corresponding sentences with a lexical indirect object such as *That book was given John for his birthday*. Larson does not, however, discuss the fact that there is an anti-focus effect in the latter, insofar as *That book was given John, not Bill* is appreciably less acceptable. Why might that be? (Bring in Rizzi 1997 and Jayaseelan 2001.)

15 In discussing the greater acceptability of *That book was given him for his birthday*, with a pronominal indirect object, Larson suggests that this might be due to clitic pronouns not being subject to the Case filter. (A similar idea is found in Baker 1993.) Evaluate this idea, using examples from as many languages as possible.

16 In discussing languages like French, which disallow double object sentences of the English sort, Larson adopts an idea from Kayne (1981), according to whom double object sentences are possible only in languages that allow preposition stranding. Zhang (1990) has, however, argued, in particular on the basis of Chinese and Indonesian, that double object sentences do not imply

preposition stranding; all that holds, according to Zhang, is that preposition stranding languages always allow double object sentences. Evaluate the implications of Zhang (1990) for Larson's analysis.

References

Aoun, J., and D. Sportiche. 1983. The formal theory of government. *The Linguistic Review* 2: 211–236.

Bach, E. 1979. Control in Montague Grammar. *Linguistic Inquiry* 10: 515–531.

Baker, M. 1985. *Incorporation: A theory of grammatical function changing*. Doctoral dissertation, MIT. [Also 1988, Chicago: University of Chicago Press.]

Baker, M. 1993. Why unaccusative verbs cannot dative shift. In A. Schafer (ed.), *Proceedings of NELS 23*, Amherst: GLSA, University of Massachusetts, vol. 1, pp. 33–47.

Baltin, M. 2009. The properties of negative non-finite complements. Ms., New York University.

Barss, A., and H. Lasnik. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17: 347–354.

Basilico, D. 2008. Particle verbs and benefactive double objects in English: High and low attachments. *Natural Language and Linguistic Theory* 26: 731–773.

Belletti, A., and L. Rizzi. 1988. Psych verbs and θ-theory. *Natural Language and Linguistic Theory* 6: 291–352.

Belletti, A., and L. Rizzi. 2012. Moving verbal chunks in the low functional field. In L. Brugè, A. Cardinaletti, G. Giusti, N. Munaro, and C. Poletto (eds), *Functional Heads: The Cartography of Syntactic Structures 7*, New York and Oxford: Oxford University Press, pp. 129–137.

Burzio, L. 1986. *Italian Syntax: A Government-Binding Approach*. Dordrecht: Reidel.

Chomsky, N. 1955/1975. *The Logical Structure of Linguistic Theory*. New York and London: Plenum Press. [Also 1985, University of Chicago Press.]

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Collins, C. 2005. A smuggling approach to the passive in English. *Syntax* 8: 81–120.

Collins, C. 2008. Home sweet home. In L. Levinson and O. Săvescu-Ciucivara (eds), *NYU Working Papers in Linguistics* 1, New York: Dept. of Linguistics, New York University, pp. 1–34.

den Dikken, M. 1995. *Particles: On the Syntax of Verb-Particle, Triadic, and Causative Constructions*. New York: Oxford University Press.

Dowty, D. 1979. Dative 'Movement' and Thomason's Extensions of Montague Grammar. In S. Davis and M. Mithun (eds), *Linguistics, Philosophy and Montague Grammar*, Austin: University of Texas Press, pp. 153–222.

Emonds, J. 1972. Evidence that indirect object movement is a structure-preserving rule. *Foundations of Language* 8: 546–561.

Harley, H. 2003. Possession and the double object construction. In P. Pica and J. Rooryck (eds), *Linguistic Variation Yearbook* 2, Amsterdam and Philadelphia: John Benjamins, pp. 31–70.

Jacobson, P. 1983. *Grammatical relations*. Paper presented at the 1983 Winter meeting of the Linguistic Society of America, Minneapolis, Minnesota.

Jacobson, P. 1987. Phrase structure, grammatical relations and discontinuous constituents. In G. Huck and A. Ojeda (eds), *Discontinuous Constituency. Syntax and Semantics* 20, New York: Academic Press, pp. 27–69.

Jayaseelan, K. A. 2001. IP-internal topic and focus phrases. *Studia Linguistica* 55: 39–75.

Kayne, R. S. 1981. On certain differences between French and English. *Linguistic Inquiry* 12: 349–371.

Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. S. 1999. Prepositional complementizers as attractors. *Probus* 11: 39–73. [Reprinted in Kayne (2000, Ch. 14).]

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2004. Prepositions as probes. In A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures* 3, New York and Oxford: Oxford University Press, pp. 192–212.

Kayne, R. S. 2008. Expletives, datives, and the tension between morphology and syntax. In T. Biberauer (ed.), *The Limits of Syntactic Variation*, Amsterdam and Philadelphia: John Benjamins, pp. 175–217. [Reprinted in Kayne (2010, Ch. 6).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Koopman, H. 1985. On deriving deep and surface order. In C. Jones and P. Sells (eds), *Proceedings of NELS 14*, GLSA, University of Massachusetts, Amherst, pp. 220–235.

Kratzer, A. 1996. Severing the external argument from the verb. In J. Rooryck and L. Zaring (eds), *Phrase Structure and the Lexicon*, Dordrecht: Kluwer, pp. 109–137.

Landau, I. 2002. (Un)interpretable Neg in Comp. *Linguistic Inquiry* 33: 465–492.

Marantz, A. 1984. *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.

Michelioudakis, D. 2012. *Dative arguments and abstract Case in Greek*. Doctoral dissertation, University of Cambridge.

Oehrle, R. 1976. *The grammatical status of the English dative alternation*. Doctoral dissertation, MIT.

O'Grady, W. 1980. The universal characterization of passivization. *Linguistic Analysis* 6: 393–405.

Pesetsky, D. 1995. *Zero Syntax: Experiencers and Cascades*. Cambridge, MA: MIT Press.

Phillips, C. 2003. Linear order and constituency. *Linguistic Inquiry* 34: 37–90.

Pylkkänen, L. 2008. *Introducing Arguments*. Cambridge, MA: MIT Press.

Reinhart, T. 1979. Syntactic domains for semantic rules. In F. Guenthner and S. Schmidt (eds), *Formal Semantics and Pragmatics for Natural Language*, Dordrecht: Reidel, pp. 107–130.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 281–337.

Roberts, I. 1985. Agreement parameters and the development of English modal auxiliaries. *Natural Language and Linguistic Theory* 3: 21–58.

Stowell, T. 1981. *Origins of phrase structure*. Doctoral dissertation, MIT.

Szabolcsi, A. 1983/84. The possessor that ran away from home. *The Linguistic Review* 3: 89–102.

Szabolcsi, A. 1994. The noun phrase. In F. Kiefer and K. É. Kiss (eds), *The Syntactic Structure of Hungarian. Syntax and Semantics 27*, San Diego: Academic Press, pp. 179–274.

Travis, L. 1985. *Parameters and effects of word order variation*. Doctoral dissertation, MIT.

Zhang, S. 1990. Correlations between the double object construction and preposition stranding. *Linguistic Inquiry* 21: 312–316.

14

Facets of Romance Past Participle Agreement

Richard F. Kayne

1989

14.1   Introduction

A striking feature of natural language is the phenomenon often labeled agreement: the presence of a certain feature on multiple elements in a given local domain. For example, many languages exhibit agreement between the subject and the finite verb. In a Romance language, like Italian, this means that both express the same value for the person and number features, as exemplified in (1):

(1)   Mentre voi   riman*ete* qui,   noi   torn*iamo* a casa.   (Italian)
       while   you(2pl) stay(2pl) here, we(1pl)   return(1pl) home
       'While you stay here, we're going back home.'

Within generative syntax, subject–verb agreement is often accounted for by making use of two structural positions, the head and the specifier of a projection, appealing to the special nature of the specifier–head configuration. In the case of (1), for example, the subject, which bears person and number features, is said to be in the specifier of a functional head (Infl, or T) that also bears person and number features, and this configuration is said to trigger agreement in those features (the details of how exactly they come to share the same value vary across proposals).

The question arises whether this geometric configuration is specific to subject–verb agreement or more generally a configuration that gives rise to agreement in language. In *Facets of Romance Past Participle Agreement*, Kayne addresses this question by examining data from French and Italian, and proposes an analysis that makes past-participle agreement analogous to subject–verb agreement, in the sense that it also depends on a Spec–head configuration.

The central set of examples is the following:

(2) a. Paul a repeint(*-es) les chaises. (French)
     Paul has repainted-AGR the chairs
   b. Paul les a repeint-es.
     Paul them has repainted-AGR
     'Paul has repainted them.'
   c. Combien de chaises Paul a repeint-es?
     how.many of chairs Paul has repainted-AGR
     'How many chairs has Paul repainted?'

In (2a) the object follows the past participle in linear order, and the participle cannot agree with it. In (2b,c) the object occurs to the left of the participle, and the two agree in number and gender. Notice, however, that the object is in different positions in the two examples: in (2b) the object clitic pronoun immediately precedes the auxiliary, whereas in (2c) the object is a *wh*-phrase at the front of the sentence, to the left of the subject. If agreement were established on the basis of linear adjacency, we would expect the object to prevent agreement between the subject and the finite auxiliary in (2b), and the subject to prevent agreement between the object and the participle in (2c), contrary to fact.

Kayne proposes that agreement does not depend on linear adjacency, nor is it the result of a direct relation between the participle and its complement. Instead, as in Kayne (1985), he postulates the existence of a functional head, AGR, sister of VP, and argues that objects that occur to the left of the verb move through the specifier of AGR, as indicated below for the case of the clitic pronoun:

(3) Paul les$_k$ a [$_{AgrP}$ e$_k$ AGR [$_{VP}$ repeint-es e$_k$]]

Past participle agreement, Kayne argues, results from the object instantiating a spec–head configuration with the AGR head. Since the object can only instantiate this configuration when it moves out of the VP, this analysis straightforwardly captures the correlation seen in (2) between the preverbal position of the object and the presence of past participle agreement.

The AGR head postulated in this paper later came to be called AgrO and viewed as both the locus of agreement and of accusative Case assignment (cf. Chomsky 1993). In subsequent work, Chomsky (2000, 2001) questioned the privileged status of the Spec–head configuration and proposed that agreement be seen as the result of the operation Agree, which is established between a head and the closest phrase in its c-command domain that meets certain requirements. This move successfully captures cases of long-distance agreement (e.g. in Icelandic), as well as agreement in English existential constructions, but does not adequately capture the pattern of past participle agreement discussed in this paper. The correct structural representation of agreement relationships and the constraints on their morphological manifestations have been extensively explored

in much subsequent work on a range of different languages (cf. among others, Chung 1998; Polinsky and Potsdam 2001; Bhatt 2005; Koopman 2003, 2005; Boeckx 2006; Baker 2008; Miyagawa 2009).

14.2   From "FACETS OF ROMANCE PAST PARTICIPLE AGREEMENT"

This article will address itself primarily to the agreement found in past participle constructions containing an auxiliary comparable to English 'have'. As seen in (1)-(3), the past participle in French does not agree with a following NP, but may agree in the corresponding accusative clitic construction or Wh-construction:

(1) a.   Paul a repeint les chaises. ('P has repainted the chairs')
    b.   *Paul a repeintes les chaises. (same, with agreement added)
(2) Paul les a repeintes. ('P them-has repainted')
(3) les chaises que Paul a repeintes ('the chairs that. . .')

It is usual to say that in (2) and (3), the past participle agrees with the clitic or Wh-phrase, much as the auxiliary 'a' in (1)-(3) is said to agree with the subject NP 'Paul'. We shall argue, however, that this parallelism is only partially valid, and that there is a significant distinction to be drawn between the two types of agreement, namely that although the finite auxiliary does agree directly with its subject, the past participle in (2) and (3) does not agree directly in the same sense with either clitic or Wh-phrase. Rather, the past participle agreement in (2) and (3) must be mediated by an empty category that intervenes between the clitic or Wh-phrase and the past participle.

What this means for (2), for example, is that the S-structure representation 'Paul les$_i$ a repeintes [e]$_i$', with one empty category corresponding to the D-structure position of the clitic, is not adequate, and that we should adopt instead a representation like (4):

(4)   Paul les$_i$ a [e]$_i$ repeintes [e]$_i$

In (4), a direct agreement relation will hold between the participle and the empty category preceding it, rather than between participle and clitic. (Given the ungrammaticality of (1b), we do not want to say that there is a direct agreement relation between the participle and the empty category in object position).

The reasons for adopting (4) are of two kinds: On the one hand, it will give us a more restrictive characterization of agreement than would be available if (4) were not adopted. On the other, there are certain empirical properties of past participle agreement that (4), but not an analysis lacking the first empty category, will allow us to elucidate. To mention one before closing this introductory section, the agreement in (3) is impossible if the subject of the auxiliary is an expletive; we will show below how (the Wh counterpart of) (4) leads to an account.

Section 1

We start from the position that, all other things being equal, it is desirable to have a maximally unified theory of past participle agreement and finite verb agreement. The latter is generally analyzed as involving a node AGR: 'NP AGR $[_{VP} V. . .]$', and we will do the same for the former:

(5)   Paul les$_i$ a [e]$_i$ AGR$_i$ repeintes [e]$_i$

(At some point in the derivation AGR will lower to V or V raise to AGR – the choice does not affect what follows.) The presence of AGR in (5) will allow us to distinguish French and Italian straightforwardly from Spanish, which does not have (2) (or (3)), by saying that Spanish active past participles are incompatible with AGR (a possible reason is mentioned below), and occur only in a structure like 'NP CL$_i$ $V_{aux} V_{pp}$ [e]$_i$'.

The agreement relation will now be expressed by coindexing (or linking, in Higginbotham's (1983) sense) AGR with some NP. We know from the case of finite verb agreement that this relation is subject to locality conditions. Thus, the finite verb will agree with its own subject, but not with the subject of the next highest verb up. Having an empty category preceding AGR in (5) yields a more strongly local agreement relation than would be the case if AGR were linked directly to the clitic. This consideration has particular force in the Wh counterpart of (5), e.g. in (7), corresponding to the sentence (6):

(6)   Je me demande combien de tables Paul a repeintes. ('I wonder how-many of tables P has repainted')
(7)   . . .combien de tables$_i$ Paul a [e]$_i$ AGR$_i$ repeintes [e]$_i$

Were the first empty category absent, AGR$_i$ would have to be linked directly to the Wh-phrase, across the IP headed by that AGR$_i$, the VP headed by the auxiliary, and the IP headed by the (unlike indexed) agreement associated with the auxiliary. If, on the other hand, we adopt (7), we can maintain the characterization of AGR as linking only with a NP that it governs (or, perhaps, that governs it), given that a government relation does hold between the first empty category and AGR, although not between the Wh-phrase and AGR.

Thinking of the agreement relation in terms of government, as in Chomsky (1981, 211), gives us a straightforward account of the ungrammaticality of (1b) in French, since in '. . .AGR $[_{VP}$ repeint(*es) les chaises]', no government relation can hold between AGR and the NP governed by V (although AGR m-commands that NP, the intervening V creates a minimality barrier, in Chomsky's ([1986], 42) sense).

The postulation of an extra empty category in (5) and (7) leads to the question of why that position could not be filled by a lexical NP:

(8)   *Paul a ces tables repeint(es). ('P has these tables repainted')

(8) is sharply ungrammatical with 'a' taken as tense auxiliary 'have', with or without past participle agreement. The contrast between (8) and (6)/(7) is not, however, unfamiliar, in that a rather similar one occurs elsewhere in French, in the case of infinitival complements embedded under verbs like 'croire' ('believe'):

(9)   la personne que je croyais avoir disparu ('the person that I believed
        to-have disappeared')
(10)   *Je croyais cette personne avoir disparu.

In Kayne (1984, 1.1.3, 5.3), it is argued that this can be accounted for if 'cette personne' in (10) is not in a Case-marked position (unlike its English counterpart), and if Wh-movement in (9) has the effect of allowing Case to be assigned to the relevant chain via another position (Comp). The reason that 'cette personne' in (10) is not in a Case-marked position is not that 'croire' is not a Case-assigner (it is), but rather that the infinitival subject position is too far away from 'croire' (in present terms, it is separated from it by both CP and IP).

Now the ungrammaticality of (10) and that of (8), although similar, must not be fully identified, since their clitic counterparts diverge sharply, that of (8) being fully grammatical, as in (2)/(5), while that of (10) is quite deviant:

(11)   *?Je la croyais avoir disparu. ('I her-believed to-have disappeared')

Our proposal is that (8) and (10) both violate Case theory requirements, that in both the offending lexical NP is in a non-Case marked position (whence the similarity), but that the reason for Case not reaching that position is different in the two constructions: Whereas in (10), it is that a Case-assigning verb is too far away, in (8) the higher verb (the auxiliary), though near enough, is incapable of assigning Case at all.

That tense auxiliary 'avoir' is not a Case-assigner is further supported by the observation that 'avoir' is incompatible with the accusative clitic 'le' that can stand for various predicate elements, and in particular for (passive) past participles:

(12)   Paul sera photographié par Marie. ('P will-be ph. by M')
(13)   Paul le sera par Marie.

If we take (the trace of) 'le' in (13) to receive Case from copula 'être' ('be'), and if tense auxiliary 'avoir' cannot assign Case, then (14) vs. (15) falls into place:

(14)   Paul a téléphoné (à Marie). ('P has tel. to M')
(15)   *Paul l'a (à Marie).

Against the background of the ungrammaticality of (8), which crucially involves the inability of auxiliary 'avoir' to assign Case, consider again (2), repeated here as (16):

(16)   Paul les a repeintes.

If the accusative Case of the clitic cannot be due to 'avoir', as now must be true, then that Case must come from the participle 'repeintes'. That is, in (17) 'repeintes' assigns Case to the NP position that it governs:

(17)   Paul les a [e] AGR repeintes [e]$_{Case}$

That an active past participle can assign Case is clearly supported by an Italian construction studied by Belletti (1981), in which one finds an active past participle unaccompanied by any auxiliary, yet taking an object and assigning accusative Case to it:

(18)   Conosciuto me, Maria. . . ('known/met me, M. . .')

Although (18) is absent from French, that is arguably due to independent considerations (cf. section 6 below), so that we can plausibly take (18) as support for French (17), as well as for the equivalent of (17) in Italian.

  If past participles in French (and Italian) can assign Case, then more must be said about the ungrammaticality of (8), i.e. about the ill-formedness of (19):

(19)   *Paul a ces tables AGR repeintes [e]$_{Case}$

Consider a principle such as (20):

(20)   If a Case-marked chain is headed by an A-position, then that A-position
        must be assigned Case.

Since nothing about (19) would allow one to say that 'ces tables' is in an A-bar-position, (19) is ruled out by (20). (17), on the other hand, contains a chain headed by the position of the clitic, which is not an A-position.

  We are now in a position to return to the Wh case of past participle agreement illustrated in (7), repeated here as (21), with Case added:

(21)   . . .combien de tables$_i$ Paul a [e]$_i$ AGR$_i$ repeintes [e]$_i$, Case

Since the Wh-phrase 'combien de tables' is, being an operator, not part of the relevant chain, it follows from (20) that the first empty category must not be in an A-position. Let us take it, then, to be adjoined to the IP (strictly speaking AGRP) headed by AGR$_i$ as in (22):

(22)   . . .combien de tables$_i$ Paul a [$_{IP}$ [e]$_i$ [$_{IP}$ AGR$_i$ repeintes [e]$_{i, Case}$]]

(If there is in addition an [e]$_i$ in the Spec position of AGR$_i$, then the adjoined [e]$_i$, which is not in an A-position, must be taken to be the head of the chain, so that (20) does not come into play; if there is no such [e]$_i$ in Spec position, then [e]$_i$ in object position can be taken as head.) In the framework of Chomsky ([1986]), the intervening lower IP segment in (22) will not block the government relation between the first [e]$_i$ and AGR$_i$, as desired.

Summing up this section, we see that a restrictive theory of agreement imposes the presence of an extra $[\,e\,]_i$ in (22) (and (17)), and that, given the inability of auxiliary 'avoir' to assign Case, this $[\,e\,]_i$ must, for Case-theoretic reasons, be taken in (22) to be in an A-bar-position. In the next section, we attempt to show that the presence of this $[\,e\,]_i$ in A-bar-position in (22) has distinctly favorable empirical consequences.

[. . .]

Section 3

Ruwet (1982, 150) has observed an interesting contrast between the following two constructions:

(37)   une femme qu'on a dit belle ('a woman that one has said beautiful')
(38)   une femme qu'on a dit ne pas être belle ('. . .neg. not to-be. . .')

(37) is an instance of Wh-movement applied to the subject of a small clause embedded under 'dire', and (38) an example of the same except that the small clause is replaced by an infinitive. Ruwet's observation is that past participle agreement, although possible in (37), yields a deviant result in (38):

(39)   une femme qu'on a dite belle
(40)   *?une femme qu'on a dite ne pas être belle

The relevant substructure of (39) is 'Wh$_i$ NP a $[_{IP}\,[\,e\,]_i\,[_{IP}$ AGR$_i$ dite $[[\,e\,]_i$ belle]]]' in which the second $[\,e\,]_i$ is governed by the participle across the small clause boundary. The corresponding representation for (40) must involve more than simply replacing the small clause by an infinitival IP, however, since as discussed earlier for (9) and (10), French does not permit the subject of an embedded infinitive to be governed across a boundary by a verb like 'dire'. Our proposal for (9), and similarly (38)/(40), has attributed an essential role to the trace in Comp left by Wh-movement. Thus, (38) will have a representation such as '. . .Wh$_i$ NP a dit $[_{CP}$ $[\,e\,]_i\,[_{IP}\,[\,e\,]_i$ ne pas être belle]]'. That of (40) will be the same, except that the theory of agreement adopted here forces the postulation of an extra empty category adjoined to the IP headed by the participial AGR:

(41)   . . .Wh$_i$ NP a $[_{IP}\,[\,e\,]_i\,[_{IP}$ AGR$_i$ dite $[_{CP}\,[\,e\,]_i\,[_{IP}\,[\,e\,]_i\ldots]]]]$

Now in discussing (22) above, we noted that it might be the case that we should also have an $[\,e\,]_i$ in the Spec of AGR$_i$. If so, then we will have a clear violation of the 'improper movement' constraint seen to play a role in the preceding section [in the full article], in that in (41) the $[\,e\,]_i$ in Spec of CP, an A-bar-position, will be bound by this $[\,e\,]_i$ in Spec of AGR$_i$, an A-position. If, on the other hand, there were no such $[\,e\,]_i$ that must be added to (41), then we would apparently have to interpret 'improper movement' broadly enough that the binding of $[\,e\,]_i$ in Spec of CP by AGR$_i$ itself triggers the violation. This might not be necessary, however, if

we took (41) to violate a stronger form of improper movement than is currently assumed, namely one that would go back to Chomsky (1973, section 3) and prohibit a category in Spec of CP from being bound (within the scope of the maximal operator) by any phrase not itself in the Spec of some CP. In this way the [e]$_i$ adjoined to IP in (41) would become the illicit binder.

Thus (40) is accounted for, and correctly distinguished from (39). (39) differs from (40) precisely in that the subject position of the small clause is accessible to government by the matrix verb. Hence the presence of an [e]$_i$ Spec of CP is not required, so that the [e]$_i$ adjoined to IP in (39) yields no illicit binding relation.

Section 4

We have postulated the existence of an IP-adjoined empty category in the Wh past participle agreement construction. For the comparable construction in which the participle agrees with a clitic, we have not been led to suppose that there is such an IP-adjoined category – cf. (17) above, with the empty category in Spec of AGR. We take this difference in representation to be supported by the fact that, contrary to the impression perhaps given by French, these two subcases of past participle agreement are not always found together. This is seen most readily in nonformal spoken Italian, which has the past participle agreeing with a clitic, but not with a Wh-phrase:

(42) Paolo ha visto le ragazze. ('P has seen the girls')
(43) Paolo le ha viste.
(44) le ragazze che Paolo ha visto/*viste

As it turns out, neither French nor Italian constitutes an isolated case. In fact, French itself is not homogeneous, in that there are many speakers who have past participle agreement neither with object clitics nor with object Wh-phrases (i.e. these speakers share the Spanish paradigm). Those who do have it always, as far as we know, have it both with clitics and with Wh-phrases. The French situation is further complicated by the fact that past participle agreement is one of the areas most discussed by normative grammarians. So it is in a sense welcome to find that the combination of agreement with clitic and agreement with Wh-phrases is attested in various nearby languages/dialects, for example those varieties of Occitan described by Camproux (1958, 323), Kelly (1973, 196, 200), as well as that referred to by Seguy [sic] (1950, 53); also the Vaudois of Reymond and Bossard (1979, 93, 147), the Normandy dialect of Lepelley (1974, 107, 113) and apparently the Brittany dialect of Hervé (1973, 84).

On the other hand, the Italian paradigm of (43)-(44), with clitic agreement but without Wh-agreement, is also robust, shared as it is by Catalan according to Fabra (1981, para. 73), by the varieties of Occitan described in Rohlfs (1977, 223), Bonnaud (1974, 39, 57), Lamouche (1902, 106), Marshall (1984, 66) and Miremont (1976, 54, 55), by the Beuil dialect described by Blinkenberg (1948, 118), and by the Corsican, Milanese, Cremonese and Bolognese dialects of Italian, to judge by Yvia Croce (1979, 137), Beretta (1984, 124, 145), Rossini (1975, 126) and Mainoldi (1950, 63) respectively.

What is to the best of our knowledge notably lacking, though, is a language having the Wh-case of object agreement, but not the clitic case. This means that adjunction of NP to the IP complement of the auxiliary is not automatically available, even to a language that otherwise uses the '$[_{IP}\ [\ e\ ]_i\ AGR_i\ V_{pp}\ [\ e\ ]_{i,\ Case}]$' structure that underlies clitic agreement. Conversely, it means that if a language allows '$[_{IP}\ [\ e\ ]_i\ [_{IP}\ AGR_i\ V_{PP}\ [\ e\ ]_{i,\ Case}]]$', then the corresponding structure with $[\ e\ ]_i$ in Spec of AGR but with no IP adjunction is straightforward.

[. . .]

Conclusion

Various Romance languages, to varying degrees, have constructions in which it appears that a past participle is agreeing with its object NP. We have argued that all such cases should be interpreted otherwise: The past participle never agrees directly with an NP in object position. Rather, when there is agreement, that agreement is due to the NP having moved to (or through) a position governed by an abstract element AGR generated as sister to the VP headed by the participle.

The question arises as to whether this characterization of agreement between a verbal form and a NP as always being mediated by such an AGR and by a government relation with that AGR should be taken to be valid within Romance or more widely. The widest possible interpretation, the study of the consequence of which is well beyond the scope of this article, would require that many apparent cases of agreement between a verbal form and a NP object (more generally: NP governed by V) be analyzed as involving clitics (i.e. clitic doubling of some sort) rather than AGR.

14.3 Questions pertaining to Kayne (1989)

1   Pronouns in Romance languages vary in form depending on features of person, number, gender, and case. Which of these play a role in verbal agreement? Do all types of verbs (finite verbs, infinitives, gerunds, past participles) show the same agreement behavior?

2   (Extra credit) Find a Romance language/dialect in which finite verbs show agreement in gender.

3   Which is more commonly found within Romance, past participle agreement with a *wh*-phrase or past participle agreement with an object clitic? Why would there be such an asymmetry?

4   Some Romance languages have past participle agreement with a following lexical object. The theory of agreement in this paper says that that object cannot be in canonical object position. What kind of evidence might one look for in trying to test this theoretical prediction?

5   Past participle constructions in Romance typically contain an auxiliary preceding the participle. That auxiliary is sometimes *have*, sometimes *be*. On the basis of this paper, to what extent does the identity of the auxiliary matter to past participle agreement? Are there important data bearing on this that are not

included in this paper? If so, give some examples. (Hint: Look at Kayne 1993 and D'Alessandro and Roberts 2010.)

6 An independent AGR head of the sort made use of here (and in Kayne 1985) is commonly found in syntactic work of the early nineties. Where in Chomsky (1995) is a transition made to a theory having agreement features appear only as part of another head such as Tense?

7 Discuss the similarities and differences between Chomsky's (1957) approach to agreement and the two approaches of the previous question (AGR; agreement features).

8 (Long answer required) To what extent would it be straightforward to transpose the proposals in the present paper to such a theory as that of Chomsky (1995)?

9 Where in Chomsky's work of the early 21st century is a further transition made to a theory based on an operation called Agree that has the property that a tensed verb can agree with a phrase that it asymmetrically c-commands? What central property of Romance past participle agreement does Agree seem to be incompatible with?

10 D'Alessandro and Roberts (2010) develop an approach to past participle agreement using Agree that tries to preserve many of the results of the present paper. Find properties of past participle agreement that they do not discuss. (Extra credit: To what extent does the incompleteness of their discussion reduce the plausibility of their proposal?)

11 Clitic doubling of the sort found in Spanish in sentences like *Yo le dí un libro a Juan* ('I him(dative) gave a book to J') has something in common with past participle agreement (and with finite verb agreement), in that the clitic (here *le*) and the lexical argument (here *Juan*) have to agree in certain features. Yet there are differences between clitic doubling and verb agreement. Find as many as you can.

12 Are there cases in which clitic doubling between a clitic and a lexical DP co-occurs with past participle agreement involving that same clitic, in one sentence? Why exactly is Spanish not a likely candidate? Find at least one language/dialect that does seem to have the co-occurence sought. (Hint: Look at the footnotes in Kayne 2003a.)

13 There are speakers of French for whom past participle agreement holds optionally, that is, both the sentence with it and the corresponding sentence without it are possible. Would you expect there to be a difference in interpretation between two such sentences? Why, or why not? (For relevant discussion, see Déprez 1998 and references cited there.)

14 In Italian, past participle agreement with a third-person accusative clitic is obligatory. Yet with a first- or second-person accusative clitic it is, for some speakers, optional (and for some, actually impossible). Why would there be such a difference between third-person and first-/second-person? What is the relevance of the fact that in passives past participle agreement with the subject of the sentence is obligatory for all speakers, for all persons?

15 Justify a minimum number of different types of languages/dialects that one can distinguish from one another using just past participle agreement.

16 Which aspects of Taraldsen (1991) are compatible with the present paper, and which are not?

17 What does Romance past participle agreement have in common (and what does it not have in common) with the nonstandard English agreement discussed by Kimball and Aissen (1971) and Kayne (2003b)? Bring in den Dikken (2001).

18 Standard Italian allows a past participle to agree with a following direct object in the so-called "absolute construction," e.g. *Conosciuta me, Gianni. . .* ('known/met(fem.) me, John . . .'). What might be the reason for the difference between this and the impossibility of such agreement in **Gianni ha conosciuta me* ('G has known/met(fem.) me')? Bring in Pollock (1989).

19 The "absolute" construction mentioned in the preceding question is absent from French. Why exactly might that be?

20 (Extra credit) Is so-called object agreement in Basque (cf. Laka 1993) closer to Spanish clitic doubling or to French and Italian past participle agreement? Give your reasons. To what extent do you agree with Preminger's (2009) approach?

References

Baker, M. C. 2008. *The Syntax of Agreement and Concord.* Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.

Belletti, A. 1981. Frasi ridotte assolute. *Rivista di Grammatica Generativa* 6: 3–32.

Beretta, C. 1984. *Contributo per una grammatica del milanese contemporaneo,* II Edizione, Edizioni del Circolo Filologico Milanese. Milan: Virgilio.

Bhatt, R. 2005. Long distance agreement in Hindi–Urdu. *Natural Language and Linguistic Theory* 23: 757–807.

Blinkenberg, A. 1948. *Le patois de Beuil* (Acta Jutlandica. Aarsskrift for Aarhus Universitet XX, 3 (H34)). Copenhagen: Ejnar Munksgaard.

Boeckx, C. 2006. *Agreement Systems*. Amsterdam: John Benjamins.

Bonnaud, P. 1974. *Nouvelle grammaire auvergnate*. Clermont-Ferrand: Cercle d'Auvergne.

Camproux, C. 1958. *Étude syntaxique des parlers gévaudanais* (Publications de la Faculté des lettres de l'Université de Montpellier, XX). Paris: Presses Universitaires de France.

Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. 1973. Conditions on transformations. In S. Anderson and P. Kiparsky (eds), *A Festschrift for Morris Halle*, New York: Holt, Rinehart and Winston, pp. 232–286.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 1993. A Minimalist program for linguistic theory. In K. Hale and S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press, pp. 1–52.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2000. Minimalist inquiries: The framework. In R. Martin, D. Michaels, and J. Uriagereka (eds), *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, Cambridge, MA: MIT Press, pp. 89–155.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, pp. 1–52.

Chung, S. 1998. *The Design of Agreement: Evidence from Chamorro*. Chicago and London: University of Chicago Press.

D'Alessandro, R., and I. Roberts. 2010. Past participle agreement in Abruzzese: Split auxiliary selection and the null-subject parameter. *Natural Language and Linguistic Theory* 28: 41–72.

Déprez, V. 1998. Semantic effects of agreement: The case of French past participle agreement. *Probus* 10: 1–65.

den Dikken, M. 2001. "Pluringulars", pronouns and quirky agreement. *The Linguistic Review* 18: 19–41.

Fabra, P. 1981. *Gramatica catalana* (tenth edition). Barcelona: Teide.

Hervé, B. 1973. *Le parler de Plouguenast*. Master's thesis, Université de Haute-Bretagne.

Higginbotham, J. 1983. Logical form, binding, and nominals. *Linguistic Inquiry* 14: 395–420.

Kayne, R. S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, R. S. 1985. L'accord du participe passé en français et en italien. *Modèles Linguistiques* 7: 73–90. [English version in Kayne (2000, Ch. 2).]

Kayne, R. S. 1993. Toward a modular theory of auxiliary selection. *Studia Linguistica* 47: 3–31. [Reprinted in Kayne (2000, Ch. 7).]

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2003a. Person morphemes and reflexives in Italian, French and related languages. In C. Tortora (ed.), *The Syntax of Italian Dialects*, New York and Oxford: Oxford University Press, pp. 102–136.

Kayne, R. S. 2003b. Some remarks on agreement and on Heavy-NP-Shift. In M. Ukaji, M. Ike-Uchi, Y. Nishimura, and H. Nakajima (eds), *Current Issues in English Linguistics*. Vol. 2 of *Special Publications of the English Linguistic Society of Japan*, Tokyo: Kaitakusha, pp. 67–86. [Reprinted in Kayne (2005, Ch. 11).]

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kelly, R. C. 1973. *A Descriptive Analysis of Gascon*. The Hague: Mouton.

Kimball, J., and J. Aissen. 1971. I think, you think, he think. *Linguistic Inquiry* 2: 242–246.

Koopman, H. 2003. The locality of agreement and the structure of DP in Maasai. In W. E. Griffin (ed.), *The Role of Agreement in Natural Language*. Vol. 53 of *Texas Linguistics Forum*, Austin: Department of Linguistics, University of Texas at Austin, pp. 206–227.

Koopman, H. 2005. On the parallelism of DPs and clauses: Evidence from Kisongo Maasai. In A. Carnie, H. Harley, and S. A. Dooley (eds), *Verb First: On the Syntax of Verb-initial Languages. Linguistik Aktuell/Linguistics Today* 73, Amsterdam and Philadelphia: John Benjamins, pp. 281–302.

Laka, I. 1993. The structure of inflection: A case study in X^0 syntax. In J. I. Hualde and J. Ortiz de Urbina (eds), *Generative Studies in Basque Linguistics*, Amsterdam: John Benjamins, pp. 21–70.

Lamouche, L. 1902. *Essai de grammaire languedocienne (Dialectes de Montpellier et de Lodève)*. Paris: H. Welter.

Lepelley, R. 1974. *Le parler normand du Val de Saire (Manche)* (Cahier des Annales de Normandie N° 7). Caen: Musée de Normandie.

Mainoldi, P. 1950. *Manuale dell'odierno dialetto Bolognese*. Bologna: Società Tipografica Mareggiani.

Marshall, M. M. 1984. *The Dialect of Notre-Dame-de-Sanilhac: A Natural Generative Phonology* (Stanford French and Italian Studies 31). Saratoga, CA: Anma Libri.

Miremont, P. 1976. *La syntaxe occitane du Périgord*. Cuers: P. Miremont.

Miyagawa, S. 2009. *Why Agree? Why Move? Unifying Agreement-based and Discourse-configurational Languages*. Linguistic Inquiry Monographs, Cambridge, MA: MIT Press.

Polinsky, M., and E. Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19: 583–646.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Preminger, O. 2009. Breaking agreements: Distinguishing agreement and clitic-doubling by their failures. *Linguistic Inquiry* 40: 619–666.

Reymond, J., and M. Bossard. 1979. *Le patois vaudois* (second edition). Lausanne: Payot.

Rohlfs, G. 1977. *Le gascon. Études de philologie pyrénéenne* (third edition). Tübingen: Niemeyer and Pau: Marrimpouey Jeune.

Rossini, G. 1975. *Capitoli di morfologia e sintassi del dialetto cremonese*. Pubblicazioni della Facoltà di Lettere e Filosofia dell'Università di Milano, LXXVI – Sezione a cura dell'istituto di glottologia, 2. Firenze: La Nuova Italia.

Ruwet, N. 1982. *Grammaire des insultes et autres études*. Paris: Éditions du Seuil.

Séguy, J. 1950. *Le français parlé à Toulouse*. Toulouse: Édouard Privat.

Taraldsen, K. T. 1991. A directionality parameter for subject-object linking. In R. Freidin (ed.), *Principles and Parameters in Comparative Grammar*, Cambridge, MA: MIT Press, pp. 219–268.

Yvia Croce, H. 1979. *Grammaire corse* (second edition). Ajaccio: Cyrnos et Méditerranée.

15

Verb Movement, Universal Grammar, and the Structure of IP

Jean–Yves Pollock

1989

15.1 Introduction

Word order differences across languages might be explained by postulating that phrases combine in different ways, or that they combine in the same way but movement alters the order of the words they contain. For example, English and French differ in the relative order of the finite verb with respect to an adverb like *often*, as in (1):

(1) a. Jean *embrasse souvent* Marie. (French)
        John kisses   often   Mary
     b. John *often kisses* Mary.

Emonds (1978) argued that these differences should be accounted for by assuming that the adverb is structurally higher than the VP in both languages and that finite verbs obligatorily move to a position higher than the adverb in French, but not in English (where verb movement is restricted to finite auxiliaries; cf. Jackendoff 1972 and Emonds 1976). In this paper, Pollock shows the explanatory power of this approach, carrying out a detailed comparison of the relative position of verbal forms (auxiliaries, lexical verbs, modals, finite, and nonfinite forms) with respect to VP-adverbs and negative markers in French and English. This careful comparison leads him to a deeper understanding of the syntactic behavior of verbs, the properties of movement, and the structure of the clause.

   Pollock makes the novel empirical observation that French exhibits the same restrictions on verb movement as English, but only in infinitival clauses. If we look

at the relative position of a lexical verb in its infinitival form with respect to an adverb like *often*, we see it may occur to the right or to the left of the adverb:

(2) a. *Souvent paraître* triste pendant son voyage de noce, c'est rare.
      often   to.seem sad   during   one's honeymoon   that's rare
      'To often look sad during one's honeymoon is rare.'
    b. *Paraître souvent* triste pendant son voyage de noce, c'est rare.
      to.seem   often   sad   during   one's honeymoon   that's rare

But if we look at infinitival forms with respect to the negative marker *pas*, we see a contrast: an infinitival auxiliary may (optionally) occur to the left of *pas*, but an infinitival form of a lexical verb cannot:

(3) a. N'*être pas* heureux est une condition pour écrire des romans.
      neg-to-be neg happy   is a   condition for   writing some novels
      'Not to be happy is a prerequisite for writing novels.'
    b. *Ne *sembler pas* heureux est une condition pour . . .
      ne to.seem NEG happy   is a   condition for . . .

If we were to see this as the result of the adverb phrase and the verb phrase combining in different ways, we would have to say that their combination is sensitive to properties of the verb. Pollock instead provides a simple and elegant analysis by appealing to verb movement. Since a lexical verb in its infinitival form can occur to the left of *souvent* but not to the left of *pas*, he proposes that a VP-adverb like *souvent* is structurally lower than the negative marker *pas*. He then suggests that verb movement takes place in two steps:

(4) V_i pas e_i [$_{VP}$ souvent t_i ...]

The first step, labeled SHORT VERB MOVEMENT, takes the verb to a position higher than VP-adverbs like *souvent*; the final step takes it to a position (usually identified as Infl, or I) higher than the negative marker *pas*. Now the pattern can be accounted for by appealing to properties of the verb, revealing that the same properties are relevant in both languages. In English, lexical verbs do not move at all, infinitival auxiliaries optionally undergo Short Verb Movement, and finite auxiliaries move to I. In French, infinitival lexical verbs are restricted to Short Verb Movement, infinitival auxiliaries move to I optionally, and all finite forms move to I obligatorily.

This analysis constitutes the basis for a number of questions and proposals, of which we will mention only two. What are the positions to which verbs move? Pollock suggests that we view what was considered to be a single node, I, as consisting of distinct projections: a head Agr, host of agreement features, and a head I, host of tense features. This has come to be known as the SPLIT-IP hypothesis. Why do auxiliaries move more than lexical verbs, in both French and English? Pollock suggests that this is because auxiliaries do not assign theta-roles, and therefore can freely raise to the head Agr; lexical verbs, in contrast, assign theta-roles, and cannot raise to Agr unless Agr is morphologically "rich."

Many of the proposals in Pollock's article have inspired much subsequent work. The idea that the "richness" of agreement drives syntactic movement led to attempts to better understand when an inflectional paradigm counts as "rich" (cf. Rohrbacher 1999, among others) and more generally to reach a better understanding of the correlation between the morphological manifestation of agreement and the articulation of syntactic structure (cf. Bobaljik 2002). The idea that the functional layer of the clause consists of two functional heads, I and Agr (with Neg in between) was critically examined in Belletti (1990). Belletti observed that, in Romance, tense suffixes are closer to the verbal stem than agreement suffixes and therefore, according to the Mirror Principle (Baker 1985), the hierarchical order should be Agr-I. Chomsky (1991) further developed this view, suggesting the existence of two Agr heads (the lower one associated with object agreement and case, the higher one with subject agreement and case), with a head T in between: AgrST-AgrO. Cinque (1999) pursued these ideas and the evidence provided by a verb movement analysis of word order patterns, and offered an even more articulated view of clausal structure, in which the inflectional domain consists of a number of functional projections expressing not only agreement and tense, but also aspect, mood, and modality.

15.2   From "VERB MOVEMENT, UNIVERSAL GRAMMAR, AND THE STRUCTURE OF IP"

In this article I will attempt to shed some light on a few systematic differences between French and English with respect to the syntax of sentence negation, questions, adverbs, floating quantifiers, and quantification at a distance. In line with much recent work in comparative syntax, I will suggest that the differences between the two languages in these seemingly unrelated areas can, and therefore should, be correlated.

I will in fact show that they can be deduced from the structure of Universal Grammar (UG) and one abstract parameter having to do with what I will call the "opacity" or "transparency" of Agr(eement) in French and (Modern) English.

As I proceed, I will be led to take a fresh look at old problems concerning the structure of the simple sentence in English and French – questions concerning so-called *Do* Support, the syntactic status of auxiliary verbs, and other related problems that have been on the research agenda ever since Chomsky (1955) put them there. This article can therefore be viewed as an attempt to show how recent proposals in the "principles and parameters" framework of generative grammar can be brought to bear on long-standing problems and puzzles and how they in fact provide real explanations for them.

In order to reach this goal, I will adopt an approach to the structure of IP that is more highly articulated than most contemporary work would seem to suggest. I will provide empirical arguments in favor of the view that Infl(ection) should not be considered as one constituent with two different sets of features ([±Tense, ±Agr]) and that instead each of these sets of features is the syntactic head of a maximal projection, AgrP and IP (the latter to be called, more

perspicuously, T(ense)P. In the same spirit, I will suggest that both French and English have a maximal projection NegP. Each such maximal projection will be shown to be a potential barrier for certain types of movements (see Chomsky ([1986])). Having established this, I will demonstrate that the Empty Category Principle (ECP) (specifically the Head Movement Constraint of Chomsky ([1986], sec. 11)), θ-theory, and quantification theory provide all the tools needed to understand the core cases of the syntax of negation, questions, adverbs, and floating quantifiers in the two languages under study. The more idiosyncratic features of Modern English will be shown to follow from the "opacity" of its Agr and from its clause structure. The ECP, quantification theory, and θ-theory, which are not open to parametric variations, would seem to virtually require a language with these properties to develop an auxiliary verb like English *do* with all its specific characteristics.

The article is constructed as follows. Sections 1 and 2 are essentially descriptive and introduce data that any comparative analysis of French and English should cover. Section 1 deals with fairly well known properties of tensed clauses, and section 2 with less well known facts in infinitival clauses. Sections 3, 4, and 5 aim at providing real explanations for the stipulative aspects of the informal analysis suggested on the basis of sections 1 and 2. Section 6 introduces more data, analyses, and speculations concerning further comparative work in the area of Verb Movement and concludes the article.

1 French versus English verb movement in tensed Clauses

1.1 Comparative implications of a standard analysis

Although work in the 1950s and 1960s in generative grammar was seldom, if ever, of a comparative nature, some of it had obvious comparative implications. This is true of Emonds's and Jackendoff's work on the French and English auxiliary systems in the (late) 1960s. Suppose that French has an obligatory rule of Verb Raising to Infl ("Aux" in Emonds's (1978) terminology) but that English has only a limited version of that rule, the so-called *Have/Be* Raising of Emonds (1976) and Jackendoff (1972). Suppose further that French and English share the D-Structure form in (1), where (Adv) is an optional adverbial position that can be occupied by VP adverbs like *often/souvent, seldom/rarement, hardly/à peìne*.

(1) $[_{IP}$ NP I $([_{Neg}$ not/pas$]) [_{VP}$(Adv) V . . .]]

If we take these proposals seriously, as we should, we can account for the minimal pairs in (2)–(5) as the surface reflex of one abstract syntactic difference, the respective scope of Verb Movement in the two languages.

(2)   a.   *John likes not Mary.
       b.   Jean (n') aime pas Marie.
(3)   a.   *Likes he Mary?
       b.   Aime-t-il Marie?

(4) a. *John kisses often Mary.
     b.   Jean embrasse souvent Marie.
     c.   John often kisses Mary.
     d. *Jean souvent embrasse Marie.
(5) a. *My friends love all Mary.
     b.   Mes amis aiment tous Marie.
     c.   My friends all love Mary.
     d. *Mes amis tous aiment Marie.

Clearly, (2a) is excluded because for the verb to end up in prenegative position, it would have to move to Infl, which it cannot since English Verb Movement is restricted to *have* and *be*. (2b) is fine because all lexical verbs undergo Verb Movement in French. (3a) is straightforwardly excluded if we analyze so-called Aux-NP Inversion as movement to the left of Infl (say, (head) movement of Infl to Comp, as in Chomsky ([1986])): for a lexical verb like *kiss* to occur in presubject position, it would first have to move to Infl, which it cannot. Therefore, (3b) is fine for exactly the same reasons as (2b): lexical verbs move to Infl in French. Given the structure in (1), the facts in (4) and (5) also follow straightforwardly. Assuming that neither French nor English allows for Adverb Movement (to the right), the only way for *often* in (4a) to end up between the verb and its object would be for the verb to move to Infl, which it cannot do. The only acceptable English sentence is therefore (4c). Since *embrasser* can, on the contrary, move to Infl, (4b) is accounted for. As for the ungrammaticality of (4d), it can also be dealt with if we assume, as Emonds (1978) did, that French Verb Movement to Infl is obligatory.

The facts in (5) are obviously parallel to those in (4) and will receive the same explanation if we adopt Kayne's (1975) view that floating quantifiers move to adverbial positions.

[. . .]

2 Verb movement in infinitives

2.1 On the structure of infinitives

Before I start examining Verb Movement in infinitives, I must spell out my assumptions about their structure. I will adopt what I take to be the null hypothesis and assume that they differ from finite clauses only in the feature composition of their Infl(ection) (and/or Comp). It will suffice for the time being to assume that infinitives and tensed sentences are distinguished by some feature, say [−finite] and [+finite]. If that is indeed the only difference between them, then structure (1) – $[_{IP}$–NPI $([_{Neg}$ not/ pas]) $[_{VP}$ (Adv)V . . .]] – is the D-Structure form of both types of sentences.

I will also take up Chomsky's (1981) far-reaching hypothesis that there are no construction-specific rules. In particular, I will assume that no specific rules of Negative or Adverb Movement are at work in infinitives (or, for that matter, in any other clause type). If that is indeed true, then we can adopt the descriptive statement (14), upon which the rest of my argumentation crucially depends:

(14)   *Not* and *ne . . . pas* stand in the same structural position in tensed clauses,
        infinitives, and gerunds.

2.2   *Ne . . . pas* and verb movement in French infinitives

It follows from the assumptions made in the previous section that we can investi-
gate the properties of Verb Movement in infinitives by simply looking at the order
of their constituents. Let us first consider French infinitives with *être* and *avoir*:

(15)   a.   Ne pas être heureux est une condition pour écrire des romans.
              '*Ne* to not be happy is a prerequisite for writing novels.'
        b.   N'être pas heureux est une condition pour écrire des romans.
              '*Ne* to be not happy . . .'
        c.   Ne pas avoir eu d'enfance heureuse est une condition pour écrire des
              romans.
              '*Ne* not to have had a happy childhood is a prerequisite for writing
              novels.'
        d.   N'avoir pas eu d'enfance heureuse est une condition pour écrire des
              romans.
              '*Ne* to have not had a happy childhood . . .'

   [. . .]

(15a–b) show that *être* can but need not move to [–finite] Infl. The order of
constituents in (15a) is interesting in itself because in our terms it is the overt
manifestation of the D-Structure form (1) that was postulated on completely dif-
ferent grounds. Without (1) and the Verb Movement analysis, the acceptability of
(15a) would force us to hold the inelegant view that the order of elements in tensed
clauses and infinitives obeys totally different principles. On the theory advocated
here, we need only say that Verb Movement is optional in infinitives to account for
both (15a) and (15b), an otherwise rather mysterious pair. (15c–d) [. . .] will
clearly be accounted for in the same way. The sentences (15b,d [. . .]) are usually
considered somewhat literary and "recherché" (see Gaatone (1971, 51)) but are
perfectly fine. Emonds's (1978) analysis of French infinitives fails here: since it
assumes that infinitives do not have an Infl node (an Aux node in Emonds's termi-
nology), it cannot describe them. [. . .] We can conclude that the obligatoriness of
Verb Movement to Infl is to be correlated with the presence of the feature [+finite].
    Let us now consider infinitives with lexical verbs. The situation here contrasts
sharply with the paradigm in (15), as (16) shows:

(16)   a.   Ne pas sembler heureux est une condition pour écrire des romans.
              '*Ne* not to seem happy is a prerequisite for writing novels.'
        b.   *Ne sembler pas heureux est une condition pour écrire des romans.
              '*Ne* to seem not happy . . .'
        c.   Ne pas posséder de voiture en banlieue rend la vie difficile.
              '*Ne* not to own a car in the suburbs makes life difficult.'

    d.   *Ne posséder pas de voiture en banlieue rend la vie difficile.
         '*Ne* to own not a car . . .'

[. . .]

It appears that although Verb Movement can apply to auxiliaries and lexical *avoir*, it cannot apply to lexical verbs in infinitives in French. This should obviously be looked at in the same light as the lexical restrictions on Verb Movement in tensed clauses in English.

[. . .]

[S]ince Verb Movement to Infl is impossible in infinitives with lexical verbs (recall (16b,d [. . .])), a "naive" extension of our previous reasoning would lead us to expect sentences of the form *Lexical V + Adv + Complements* to be equally impossible. [. . .] This is not true, as shown by the acceptable sentences in (27) [. . .]:

(27)   a.   Parler à peine l'italien après cinq ans d'étude dénote un manque de
           don pour les langues.
           'To speak hardly Italian . . .'
      b.   Paraître souvent triste pendant son voyage de noce, c'est rare.
           'To look often sad . . .'
      c.   Perdre complètement la tête pour les belles étudiantes, c'est
           dangereux!
           'To lose completely one's head . . .'
      d.   Oublier presque son nom, ça n'arrive pas fréquemment.
           'To forget almost one's name . . .'

[. . .]

We obviously do not want to lose our previous generalization. We must therefore impute the grammaticality of (27) [. . .] to the existence of some yet to be described grammatical process. Continuing to assume, as above, that there are no rules of Adverb Movement, the process in question is easily circumscribed. If the adverbs in (27) are generated in the VP-initial position in (1), then it must be a Verb Movement rule, different from Verb Movement to Infl, moving the nonfinite verb to some intermediate position before the negative adverb *pas*.

[. . .] Adverbs like *à peine* and *presque* occur only in VP-initial position, as shown by the ungrammaticality of (32a–b):

(32)   a.   *Jean   comprend   la   question   presque.
           Jean   understands   the   question   almost
      b.   *Jean   lit   les   journaux   à peine.
           Jean   reads   the   papers   hardly

It follows that sentences like (27a) and (27d) can only be derived via Verb Movement.
[. . .]

We can safely conclude, then, that French infinitives like those in (27) can, and sometimes must, be analyzed as involving a Verb Movement rule that moves the verb to some position between the negative adverb *pas* (also *plus*, *guère*, *jamais*, and

so on) and the VP-initial adverb position in (1). Call this "short" Verb Movement. Unlike Verb Movement to [–finite] Infl, it is not lexically restricted: it applies to auxiliaries and to lexical verbs alike. In addition, French, like many other languages (see Koster (1986)), has a rule "scrambling" NPs to the right, adjoining them to VP.

[. . .]

2.4.2   English

[. . .]

If English also had lexically unrestricted short Verb Movement, the sentences in (38) should be well-formed, which they clearly are not. English does seem to have a lexically restricted version of short Verb Movement, however, as shown by the acceptability of (39c,f,g,i).

(38) a.   *To speak hardly Italian after years of hard work means you have no gift for languages.
      b.   *To look often sad during one's honeymoon is rare.
      c.   *To lose completely one's head over pretty students is dangerous!
      d.   *To forget almost one's name doesn't happen frequently.

(39) a.   I believe John to often be sarcastic.
      b.   I believe John to often sound sarcastic.
      c.   (?)I believe John to be often sarcastic.
      d.   *I believe John to sound often sarcastic.
      e.   The English were then said to never have had it so good.
      f.   The English were then said to have never had it so good.

[. . .]
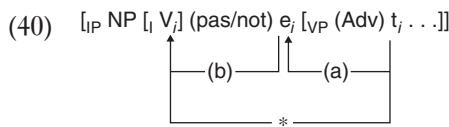
Although the facts are again a bit murky, it does appear that a significant number of speakers make fairly sharp distinctions between sentences like (39c) and (39d) [. . .]. If this is indeed correct, we are led to the conclusion that short Verb Movement in English is fundamentally restricted to *have* and *be*.

     We have now arrived at a fairly striking result. Although the French/English contrast with respect to Verb Movement to Infl ceases to exist in infinitives – in both languages only auxiliaries can undergo the rule – the very same contrast crops up in infinitives in another form: short Verb Movement is free in French but restricted to *be* and *have* in English. We would clearly miss a desirable generaliza-tion if we failed to relate the lexical restrictions on short Verb Movement and those on Verb Movement to Infl in tensed clauses: English shows lexical restrictions on both, French on neither.

     Let us take this correlation seriously and assume that Verb Movement to Infl exhibits lexical restrictions in tensed clauses if and only if short Verb Movement is also lexically restricted. A natural way of expressing this generalization formally consists in assuming that Verb Movement to Infl is not a one-step process but rather the sum of two more "local" processes, the first one consisting of short Verb Movement, the second one moving the verbs to Infl from the intermediate position

they thus reach. If we make the further hypothesis that Verb Movement to Infl can never be a one-step process, we will indeed express the correlation in its strongest possible form. We can represent this as shown in diagram (40). Our next task will be to explain why Verb Movement to Infl has to have these properties.

(40) $[_{IP}$ NP $[_I$ V$_i]$ (pas/not) e$_i$ $[_{VP}$ (Adv) t$_i$ . . .]]

3   Short verb movement as movement to Agr
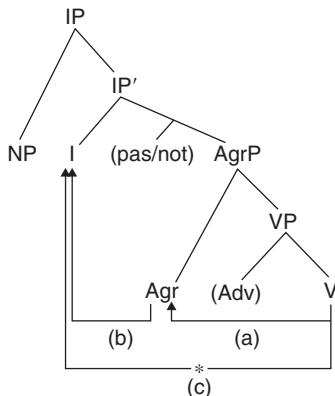
So far we have only seen evidence that there is need for short Verb Movement in the grammar of French and English. We obviously have not yet characterized the rule.

In the *Barriers* framework that I will be presupposing throughout, Verb Movement is an instance of head movement that obeys the Head Movement Constraint (HMC) formulated by Chomsky ([1986], 71):

(41)   Movement of a zero-level category β is restricted to the position of a head α that governs the maximal projection γ of β, where α θ-governs or L-marks γ if α ≠ Comp.

Clearly, if we take the position to which short Verb Movement "hops" verbs to be a head different from Infl and "closer" to VP, we will obtain as an automatic consequence the "step-by-step" derivation represented informally in (40): it will be a consequence of the existence of (41) in UG, since a direct "jump" from the VP position to Infl would violate the ECP (see Chomsky ([1986], sec. 11) and section 5.5 of [the full version of] this article). In order to reach this desirable result, I will henceforth assume that short Verb Movement is in fact Verb Movement to Agr. Agr I will assume is a category in its own right, to be distinguished from Tense, which is the head of what has so far been called Infl. We might more appropriately call the latter T(ense) and its maximal projection TP. Agr is also the head of a maximal projection AgrP, the properties of which we will come back to in sections 4 and 5.

In short, I am suggesting that a somewhat fleshed-out version of (40) should read as follows:

15.3 Questions pertaining to Pollock (1989)

1   English *do*-support distinguishes *be* (and some instances of *have*) from other
    verbs, for example *We know that they don't become angry easily* vs. **We know that
    they don't be angry very often*. *Be* is also special with respect to VP-deletion (cf.
    Williams 1977; Johnson 2001), even in the presence of an auxiliary other than
    do. This is illustrated by the fact that for many speakers *The fact that he doesn't
    deserve to become famous doesn't mean that he won't*, with a deleted "become
    famous," is appreciably better than **The fact that he doesn't deserve to be famous
    doesn't mean that he won't*, with a deleted "be famous."

    Discuss the merits and demerits of a solution to the second pair of facts
    according to which what is deleted in VP-deletion examples necessarily
    contains *do*; that is, rather than having '. . . *he won't* ~~become famous~~, what
    English really has there is '. . . *he won't do* ~~become famous~~.'

2   English readily allows sentences like *John has two sisters, but Bill has three*.
    French (like Italian) does not, as seen in *Jean a deux soeurs, mais Bill *(en) a
    trois*. In such cases, French requires the presence of a pronominal clitic *en* that
    corresponds fairly well to English *of them/of 'em* in examples such as the
    following (due to David Perlmutter): *I need a taxi. That's too bad. Two of 'em
    just went by*. Unification of French and English could be accomplished by
    attributing to the initial English example a silent/deleted instance of "of
    them," as in '. . . *but Bill has three*.'

    Discuss the ways in which this proposal is or is not similar to that of the
    previous question, bringing in Pollock's idea "that *do* is a substitute for . . .
    verbs."

3   How might the proposal in question 1 bear on the relation between British
    English allowing sentences like *Yes, I may do* and American English not
    allowing them? What parameter(s) might now be involved?

4   Pollock's idea "that *do* is a substitute for . . . verbs" seems close to saying that
    do is to V(P) as pronouns are to N(P)/D(P). If so, then one might see a paral-
    lelism between classic Condition C facts (**He$_i$ thinks that John$_i$ is really smart*
    vs. *His$_i$ mother thinks that John$_i$ is really smart*) and the well-known contrast,
    with unstressed *do*, between **Your friends do like chemistry* vs. *Your friends don't
    like chemistry*. Assuming this parallelism to be a linguistically significant one,
    explore some of its implications.

5   A familiar (economy-type) idea, which the suggestion of the previous question
    is a competitor of, is that *do*-support comes into play in English only when
    necessary, that is, only when the result of not having *do* would result in ungram-
    maticality. Yet with unstressed *do* there is a contrast between *John will definitely
    not participate in that race* and **John does definitely not participate in races like
    that*, despite the fact that leaving out *do* does not help: **John definitely not
    participate(s) in races like that*. How might a proponent of the economy
    approach to *do*-support attempt to integrate such facts?

6   Pollock's splitting of Infl into Tense and Agr led directly to Shlonsky's (1989)
    proposal that Agr itself should be split into Person and Number and Gender,

with each being a separate functional head. Going in the other direction, Chomsky (1995) argued that Agr shouldn't be an independent head at all, given its uninterpretable status. Thinking in particular of Person, discuss the relevance to Shlonsky vs. Chomsky of clitic doubling of the Romance type. Bring in Belletti (2005).

7 Discuss the relevance to Pollock vs. Shlonsky vs. Chomsky of Kayne's (2008) argument that (at least) one type of pronominal clitic must be considered to be a (phrasal) remnant, rather than a simple head.

8 French differs from English with respect to verb–adverb vs. adverb–verb order. What are the key differences between Pollock's use of movement to account for those facts and Jayaseelan's (2010)?

9 To what extent are Bantu languages similar to or different from Romance languages when it comes to the relative order of verb and adverb?

10 Pollock, in agreement with many others, has head-movement bring the tensed auxiliary to presubject position in English root interrogatives. Nilsen (2003) later suggested phrasal movement for Scandinavian V-2. Discuss the implications for this debate of Johnson (1988) and Kayne (1997). (Extra credit: Bring in the fact that, for (some) Scots, negation doesn't prepose past the subject along with V.)

11 Like Pollock, Jayaseelan (2010, note 24) takes the position that the French–English difference concerning V(P)-movement has to do with the morphological richness of the verb paradigm in the two languages. What architecture of the language faculty would or would not lead one to expect morphological richness to be able to play such a role?

12 Pollock has the special behavior of auxiliary *have*, relative to verb raising, depend on such *have* not assigning any theta-role. In some varieties of English, to one extent or another, ordinary possessive *have* shows the same special behavior, as in *Have you any money?*. Thinking of the fact that Pollock, in his note 6, takes the *got* of *Have you got any money?* to be a past participle, evaluate the following proposal, from Leonard (2007):

> All instances of English *have* showing the special behavior of *Have you seen the paper?*, *Have you any money?*, *We haven't any money*, etc. are accompanied by a past participle, i.e. they are all instances of auxiliary *have*. In those cases where no past participle is visible, *have* is accompanied by a silent counterpart of *got*.

13 While discussing the fact that infinitival *avoir* ('have') can optionally raise in French in a way not open to lexical verbs, Pollock notes that such raising is somewhat literary. How much difference would it make to his analysis if colloquial French, which disallows such raising, were taken to be a separate language from literary French? Justify your answer.

14 In his note 50, Pollock remains cautious about committing himself to the view that NegP is present in all languages. Discuss the extent to which this view does or does not jibe with Zanuttini (1997) and with Cinque (1999).

15 As Roberts (2000) notes, English *n't* is excluded from nonfinite contexts, e.g., *No, they haven't* vs. **They seem to haven't finished*. For some speakers, unstressed *not* is itself conversely excluded from the context of an overt finite

auxiliary: *It's John who won't/*will not/*'ll not help us*, *It's John who hasn't/*has not/* 's not done his homework*, *It's John who doesn't/*does not want to help*. How might these facts be fit into Pollock's framework? Discuss in addition the fact that such speakers nonetheless accept *I'm not ready, You're not ready, He's not ready*, bringing in Bernstein and Tortora (2005).

16   French allows *pour (ne) pas qu'elle parte* ('for neg not that she leaves' = 'in order that she not leave'), with (both *ne* and) *pas* between *pour* and complementizer *que*. It seems that a Romance language will allow this only if the precomplementizer negation is or contains a *pas*-type negation, that is, the type that in Romance follows the finite verb. Evaluate the following proposal: In such cases *(ne) pas*, or its equivalent in another Romance language, reaches its surface position via movement from within the clause introduced by *que*.

17   Pollock's discussion of imperatives backgrounds the well-known fact that, within English, imperatives are unique in having the subject readily remain silent. Thus imperative *Buy yourself an ice cream* contrasts with *I think that *(you) buy yourself an ice cream too often*. Thinking of Ross (1970), discuss the merits and demerits of taking imperatives (in English or in other languages) to be instances of control with a silent matrix verb. (Extra credit: Bring in (very) colloquial English root sentences like *Bought yourself an ice cream, did you?*)

18   Pollock takes the double possibility of *It's unfortunate not to have finished on time/to have not finished on time* to indicate optional movement of *have*. Evaluate the alternative possibility of taking there to be two distinct types of *not*s, bringing in Kayne (1999) on instances on multiple *not*s within one sentence.

19   Discuss the extent to which Cinque's (1999) section on adverbs and past participles in various Italian dialects requires amending Pollock's analysis.

20   In some English, infinitival *be* cannot readily be followed by sentential *not* (or by sentential adverbials), e.g., *For them not to be/to not be late would be surprising* vs. *?*For them to be not late would be surprising*. Yet even that English perfectly well allows VP-deletion with infinitival *be*, as in *For them not to be would be surprising*. How does this bear on Pollock's proposals?

References

Baker, M. 1985. The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry* 16: 373–415.

Belletti, A. 1990. *Generalized Verb Movement*. Torino: Rosenberg & Sellier.

Belletti, A. 2005. Extended doubling and the VP periphery. *Probus* 17: 1–35.

Bernstein, J., and C. Tortora. 2005. Two types of possessive forms in English. *Lingua* 115: 1221–1242.

Bobaljik, J. 2002. Realizing Germanic inflection: Why morphology does not drive syntax. *Journal of Comparative Germanic Linguistics* 6: 129–167.

Chomsky, N. 1955/1975. *The Logical Structure of Linguistic Theory*. New York and London: Plenum Press. [Also 1985, Chicago: University of Chicago Press.]

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 1991. Some notes on economy of derivation and representation. In R. Freidin (ed.), *Principles and Parameters in Comparative Grammar*, Cambridge, MA: MIT Press, pp. 417–454.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Cinque, G. 1999. *Adverbs and Functional Hheads: A Cross-Linguistic Perspective*. New York and Oxford: Oxford University Press.

Emonds, J. E. 1976. *A Transformational Approach to English Ssyntax*. New York: Academic Press.

Emonds, J. E. 1978. The verbal complex V'-V in French. *Linguistic Inquiry* 9: 151–175.

Gaatone, D. 1971. *Etudes descriptives du système de la négation en français contemporain*. Genève: Droz.

Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jayaseelan, K. A. 2010. Stacking, stranding, and pied-piping: A proposal about word order. *Syntax* 13: 298–330.

Johnson, K. 1988. Verb raising and 'have'. In D. Fekete and Z. Laubitz (eds), *McGill Working Papers in Linguistics* 6. *Special Issue on Comparative Germanic Syntax*, Montreal: Dept. of Linguistics, McGill University, pp. 156–167.

Johnson, K. 2001. What VP ellipsis can do, and what it can't, but not why. In M. Baltin and C. Collins (eds), *The Handbook of Contemporary Syntactic Theory*, Oxford: Blackwell, pp. 439–479.

Kayne, R. S. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

Kayne, R. S. 1997. The English complementizer *of*. *Journal of Comparative Germanic Linguistics* 1: 43–54. [Reprinted in Kayne (2000, Ch. 12).]

Kayne, R. S. 1999. Prepositional complementizers as attractors. *Probus* 11: 39–73. [Reprinted in Kayne (2000, Ch. 14).]

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2008. Expletives, datives, and the tension between morphology and syntax. In T. Biberauer (ed.), *The Limits of Syntactic Variation*, Amsterdam and Philadelphia: John Benjamins, pp. 175–217. [Reprinted in Kayne (2010, Ch. 6).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Koster, J. 1986. The relation between pro-drop, scrambling and verb movements. Ms., Groningen University.

Leonard, A. 2007. *To have got deletion: an analysis of the anomalous verbs of possession in British English*. Master's thesis, University College, London.

Nilsen, Ø. 2003. *Eliminating positions: the syntax and semantics of sentence modification*. Doctoral dissertation, Utrecht University.

Roberts, I. 2000. Auxiliary reduction and negation reduction: A rough sketch. *Snippets* 1: 15–16.

Rohrbacher, B. W. 1999. *Morphology-Driven Syntax: A Theory of V to I Raising and Pro-Drop. Linguistik Aktuell/Linguistics Today* 15. Amsterdam and Philadelphia: John Benjamins.

Ross, J. R. 1970. On declarative sentences. In R. Jacobs and P. Rosenbaum (eds), *Readings in English Transformational Grammar*, Waltham, MA: Ginn, pp. 222–272.

Shlonsky, U. 1989. The hierarchical representation of subject verb agreement. Ms., University of Haifa.

Williams, E. 1977. Discourse and logical form. *Linguistic Inquiry* 8: 101–139.

Zanuttini, R. 1997. *Negation and Clausal Structure: A Comparative Study of Romance Languages*. New York and Oxford: Oxford University Press.

16

Parameters of Phrase Structure

Lisa Travis

1989

16.1   Introduction

As mentioned in the introduction to Rizzi's (1982) article, in Ch. 11 of this volume, the notion of parameter developed as a partial answer to the puzzle of language acquisition: the need to figure out how a child can come to know so many subtle facts about his or her language, given limited exposure to linguistic data, in a relatively small amount of time, while going through similar stages of language acquisition regardless of the language being acquired. Among the first examples of parameters are the NULL SUBJECT PARAMETER (Rizzi 1982) and the HEAD DIRECTIONALITY PARAMETER: while the former holds the promise of deriving a cluster of differences between null subject and non-null subject languages, the latter is intended to account for differences in word order across languages. In particular, the Head Directionality Parameter appeals to the phrase structural notions of head and complement of X-bar theory, and allows two possibilities: in a given language, heads either precede or follow their complements.

This parameter allows us to account for a number of properties by appealing to a single, more abstract property of the grammatical system – exactly as parameters are supposed to do. If it is set as head-initial, we expect every head, whether it is a verb, an adposition, a complementizer, or a determiner, to precede its complement, as in English. Conversely, if it is set as head-final, we expect every head to follow its complement, so that a verb should follow its object (OV), an adposition its complement (hence the term "post-position"), and a complementizer the clause it embeds, as in Japanese. This parameter would lead us to expect that languages are consistently either head-initial or head-final. But many languages are not consistent in the relative order of heads and complements, challenging the usefulness of the Head Directionality Parameter.

Travis's (1989) paper calls attention to languages that are "inconsistent," and more specifically to the fact that there are two kinds of inconsistencies: languages like German, which are inconsistent from one head to the other (being OV and having pre-positions), and languages such as Chinese and Kpelle, which are inconsistent within a single head's maximal projection (having some constituents precede and others follow the verb). Travis views the patterns from Chinese and Kpelle as suggesting that basic word order can be (partly) determined by factors other than head-directionality, namely (i) the direction of theta-role assignment and (ii) the direction of Case assignment. She proposes that in addition to the Head Directionality Parameter, there is a directionality parameter for each of these two subdomains, which is suggested by the following observations. Chinese puts theta-marked constituents in postverbal position and adjuncts in preverbal position. This is accounted for if Chinese is head-final by default, but requires theta-marking to be rightward. In Kpelle, in contrast, object NPs precede the verb, but PPs follow the verb. This is accounted for if Kpelle is head-initial by default, but requires Case marking to be leftward.

Given three directionality parameters, the system is very powerful. In order to disallow nonexistent word orders, Travis restricts the ways in which the parameters can interact, by proposing that only one of the three parameters may be specified in a given language. If, for example, Case-assignment is specified as rightward, then theta-role assignment is unspecified, and the Head Parameter is set for head-final by default (and hence for all categories). If, on the other hand, no subdomain parameter is specified (i.e., neither for Case-assignment nor for theta-role assignment), then the Head Parameter can be specified and the value can vary across categories (as, for example, in German).

The differences in the order of head and complement exhibited across languages have been the topic of much discussion among both typologists and theoretical linguists. While the Head Directionality Parameter seemed very appealing in an X-bar theoretic framework, the points raised by Travis clearly show that a more fine-grained approach to word order is necessary, one that involves movement. If Universal Grammar forces phrase structure to be antisymmetric, as proposed in Kayne (1994) and discussed in the introduction to Kayne (2003) in Ch. 33 of this volume, then there cannot be any Head Directionality Parameter. Kayne (2011) proposes an answer to the question of why that should be so.

16.2   From "PARAMETERS OF PHRASE STRUCTURE"

In the enthusiasm to replace a system of rules with a system of principles and parameters, it is important not to lose sight of the reasons for this shift of formalism. The problem that the principles and parameters framework seeks to solve is: How can a grammatical system be flexible enough to account for language variation while at the same time be, to a large extent, restricted in order to account for the relative ease of language acquisition and the impossibility of certain language types? While rules were descriptively adequate, they were far too powerful. Not only were they capable of describing language structures that, while logically

possible, were not intuitively plausible, they also raised the question of how a child was able to internalize such a system.

In an effort to restrict the power of rule type formalisms, more and more constraints on rules were introduced until, in Government-Binding (GB) theory (e.g., Chomsky 1981), there was nothing left to the rules beyond Affect-α. The grammar became a system of restrictions rather than a system of rules. In the transformational component, Affect-α has the more specific realizations of Move-α and Delete-α (see, e.g., Lasnik and Saito [1992]). In the base component, which is the subject of this paper, one could assume that Affect-α would appear in the form of Generate-α. The grammar of restrictions is broken down into principles and parameters. Restrictions which are common to all languages are encoded in principles of Universal Grammar (UG), and these restrictions (such as Subjacency, Case Filter, Theta Criterion, Projection Principle, etc.) are considered to be part of the innate language faculty. Language variation is allowed through parameters which introduce a limited flexibility to the system. For instance, while Subjacency may prohibit movement across more than one bounding node, language variation in extraction structures is accounted for through parameterization of which nodes are considered bounding nodes by any language (Rizzi 1982; Sportiche 1981). Within the base component, generation is restricted by universals such as X-bar theory, but language specific variation of word order may be captured through parameters such as head-initial/head-final as well as those presented below.

This new framework has appealing consequences both in the field of language typology and in the field of language acquisition. Principles represent both what is common to all languages and what need not be learned by the child. Parameters represent the range of variation that can be found in natural languages as well as what has to be learned by the child.

There is a certain amount of tension in this system, however. While languages appear to vary in a large number of ways, children can acquire diverse languages with relative ease. In other words, if a child had to actually acquire each language-specific fact separately, the task of language acquisition would be much greater than evidence indicates. This tension may be relieved, however, by assuming that parameters account for a clustering of language specific properties. For instance, if a parameter *P* accounts for seven language properties, even though language A may differ from language B in seven ways, the child would only have to hear one bit of evidence to set the value for *P* which would then project six other language differences. Further, other predictive powers are gained by using parameters to account for a group of properties, since the range of possible languages is restricted. If parameter *P* incorporates properties *1* and *2*, there should be no language which exhibits *1* without also exhibiting *2*. It is important to note that this system is restrictive only when a parameter incorporates more than one language difference. As soon as there is a one-to-one mapping between language differences and parameters, we lose the explanations for the ease of language acquisition and the restrictions on language variation.

It is the aim of this paper to present a system of parameters which is powerful enough to account for the diversity of word order in natural language while being

restrictive enough to provide an explanatory account for language acquisition and restrictions on language variation. Since further in-depth research needs to be done in this area and on the specific languages involved, the framework presented here is intended as a working model.

Word order is one of the more obvious ways in which languages differ. In an *Aspects* model of syntactic theory, word order was encoded in Phrase Structure Rules. The problem with this system was that the rules were too powerful and the range of possible word orders was not at all diminished. Phrase structure rules also encoded two disparate types of relationships: dominance relations and precedence relations. The latter may be further divided into the ordering of non-heads with respect to one another and the ordering of non-heads with respect to heads.

These different relations are being teased apart in the GB framework. Dominance relations are restricted by X-bar theory. The ordering of non-heads with respect to one another is restricted by subcomponents of the grammar such as Case theory (Stowell 1981). The order of non-heads with respect to heads is restricted by one of the first parameters proposed, the headedness parameter, although it was not so called. By setting a value for headedness, separate language-specific facts were collapsed. This headedness parameter captured Greenberg's (1963) observation that VO languages tended to be prepositional and OV languages tended to be postpositional. The problem with this system, however, is that it is not powerful enough to describe existing languages. While purely head-final and head-initial languages may be described with such a parameter, difficulties arise in languages such as German which are OV but prepositional, and in languages like Chinese and Kpelle which appear to have head-internal VPs.

In this paper I examine the headedness parameter and the relation of non-heads to heads. First I discuss languages for which the headedness parameter is not sufficient and propose that smaller domains within a maximal projection may be determined by assuming that θ-role assignment and Case assignment create their own domains, and that head-internal constituents may be described by setting a direction for any of these subdomains. Such additions to the inventory of parameters, however, predict a wide variety of word orders that have not been reported. To solve this problem, in the second part of the paper I suggest a way of restricting the use of these parameters. The end result is that these parameters will represent a cluster of properties and thereby have the predictive power needed for explanatory adequacy.

1   The parameters

Two languages which present problems for the headedness parameter are Chinese and Kpelle. Both languages have head-internal VPs which suggests that distinctions finer than those provided by the headedness parameter are needed to explain head/non-head relations. As we see below, Chinese and Kpelle pose different problems for the headedness parameter and provide valuable insights into the ways in which word order may vary crosslinguistically.

1.1 Chinese

Beginning with Chinese, we can see that if one categorizes languages as being VO/OV, head-initial/head-final, one creates problems in describing languages where either the object does not speak for all of the complements of a verb, or where the logical object can appear on either side of the verb. Both of these problems occur in Chinese. As the examples in (1) show, even though the object appears to the right of the verb; other members of the VP appear to the left of the verb:

(1) a. *cóng yōu gú   chūlai*   (Li & Thompson 1973:200)
       from dark valley emerge
       'emerge from dark valley'
    b. *ta pian-le   Lisi*. (Huang 1982:27)
       he cheat-ASP Lisi
       'He cheated Lisi.'

Further, the object may also appear to the left of the verb in a *ba-* construction, as we can see in example (2):

(2) *ta ba Lisi pian-le*. (Huang 1982:27)
    he *ba* Lisi cheat-ASP
    'He cheated Lisi.'

This array of facts presents problems for typologists who see languages in the black and white categories of OV/VO and has created a discussion in the literature as to whether Chinese is OV or VO. As a first pass, we might say that bare NPs follow the verb while PPs precede the verb, presuming that *ba* is a preposition. However, as reported by Li and Thompson (1975), some PPs follow the verb while others precede. In fact, two prepositions, *zai* and *gei*, can both precede and follow but there are meaning differences depending on the position. The relevant distinctions are given in examples (3) and (4):

(3) a. *tā gěi wǒ mài le   chēzi le*. (Li & Thompson 1975:180)
       he for me sell ASP car   ASP
       'He sold a car for me.'
    b. *tā mài gěi wǒ chēzi le*.
       he sell to me car   ASP
       'He sold a car to me.'
(4) a. *Zhāng-sān tiào   zài zhuōzi-shang*. (Li & Thompson 1975:182)
       Zhang-san jump at   table-on
       'Zhang-san jumped onto the table.'
    b. *Zhāng-sān zài zhuōzi-shang tiào*.
       Zhang-san at   table-on   jump
       'Zhang-san is jumping (up and down) on the table.'

In example (3) we see that *gei* is benefactive when preverbal and dative when postverbal. In example (4) we see that *zai* has a locational reading when preverbal and a directional reading when postverbal. I suggest the correct generalization is that post-verbal constituents are θ-marked by the verb, while preverbal constituents are VP-internal adjuncts. While benefactives and locational PPs may appear with any verb, datives and directionals depend directly on the lexical selection of a verb.

[. . .]

The word order of Chinese is presented schematically in (11), where PP_1 is assigned its θ-role directly from the V and PP_2 is an adjunct PP or identifies an argument position through linking:

(11) PP_2 V NP PP_1

Assuming that all the elements to the right of the verb are directly θ-marked by the verb, one could describe this word order by having a directionality parameter for θ-role assignment. Chinese, then, would assign θ-roles to the right. To account for the placement of adjuncts to the left of the verb, there would be a default specification of head-final. The head-final parameter would have an effect only on those elements not already covered by a previously mentioned parameter.

1.2 Kpelle

We have seen that by supposing that θ-role assignment direction may be parametrically specified, overriding a parameter of headedness, one type of head-internal word order may be accounted for. However, there are head-internal word orders that cannot be described through the direction of θ-role assignment. In Kpelle, within the VP, only objects precede the verb while all PPs follow the verb. Examples are given in (12) and (13).

(12)   a.   *galoŋ a   pérɛ   tɔɔi.*   (Gay and Welmers 1971:5)
              chief AGR house build
              'The chief is building a house.'
(13)   a.   *e   pa   dipɔ.*   (Gay and Welmers 1971:31)
              he come them-to
              'He came to them.'
        b.   *e   sɛŋ-kâu tèe kâloŋ-pɔ́.*   (Givon 1975:50)
              AGR money   sent chief-to
              'He sent the money to the chief.'

This means that, unlike Chinese, argument and non-argument PPs fall on the same side of the head. Schematically, this word order is as shown in example (14).

(14)   NP V PP_1 PP_2

The direction of θ-role assignment will not solve this problem, since θ-assigned elements appear both the left and to the right of the head. However, by assuming

that the direction of Case assignment, like the direction of θ-role assignment, may also be set as a word order parameter, we will be able to capture these word order facts (see Koopman 1984 for a similar account for Mahou). What we would say about Kpelle, then, is that Case is assigned to the left, explaining the position of the object NP to the left of the V. Again, a default specification of headedness is needed. If we say that heads are initial, all the PPs will be placed to the right of the verb. The Case direction will have an effect only on those elements dependent on the verb for Case (i.e. the object), and the head-initial parameter will have an effect on all other elements with the VP (i.e. the PPs).
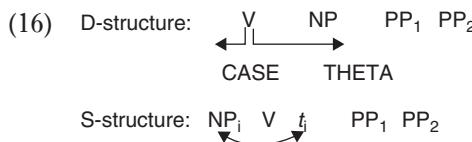
1.3 The realization of the parameters

The above view of word order parameters basically provides tools by which maximal projections may be divided into smaller domains. For instance, a VP can be seen as containing a Case domain and a θ-domain.

(15) $[_{VP} \underbrace{\underline{V\ NP}}_{\text{CASE}}\ PP_1\ PP_2]$
$\phantom{(15) [_{VP} \underbrace{V\ NP}}\underbrace{}_{\text{THETA}}$

No new primitives have been introduced to the theory, since Case relations and θ-relations are needed independently in the grammar. By allowing parameters to be sensitive to these relations, however, the system has become more powerful. Where before only non-heads could be positioned with relation to the head, now the precedence relation between a Case-assigner and a Case assignee, or a θ-role assigner and a θ-role assignee may also be given a parametric value.

Before investigating the problems that this new power brings, I would like to address the question of where the effect of any of these parameters takes place. Following GB theory strictly, the parameter of direction of θ-role assignment and the parameter of the direction of Case assignment will affect different levels of syntax. Since D-structure is seen as the pure representation of GF-θ, the direction of θ-role assignment will affect the way in which a D-structure is constructed. In other words, the D-structure of Chinese will have the same order as the S-structure. Elements θ-marked by the verb will be base-generated to the right of the verb, while all other elements will appear on the left. In Kpelle, however, the surface order will not be directly base-generated. Objects will be generated to the right of the verb in order to be θ-marked. However, since Case is assigned to the left, they will not receive Case in their base-generated position and they will be forced to move to a Case-marked position to the left of the verb. This is shown in example (16):

(16)   D-structure: V   NP   PP$_1$ PP$_2$

                        CASE   THETA

       S-structure: NP$_i$ V t_i PP$_1$ PP$_2$

Seen this way, the direction of θ-assignment affects D-structure, while the direction of Case assignment affects S-structure.

2 Restricting the system

By introducing these two new parameters, we have increased the power of our system and are running the risk of raising the same problems as those raised by the system of Phrase Structure Rules. In other words, while accounting for a larger number of word orders than allowed by the headedness parameter, we may also be predicting the possibility of word orders that are; in fact, nonexistent. With three parameters having two values each, we predict eight possible word orders within the VP. These are given below in terms of the parameters and then in terms of word order.

(17) Parameters:

	HEADEDNESS	THETA–DIRECTION	CASE–DIRECTION
a.	Final	Left	Left
b.	Final	Left	Right
c.	Final	Right	Left
d.	Final	Right	Right
e.	Initial	Left	Left
f.	Initial	Left	Right
g.	Initial	Right	Left
h.	Initial	Right	Right

(18) Word Orders:

a. $PP_2\ PP_1\ NP\ V$
b. $PP_2\ PP_1\ V\ NP$
c. $PP_2\ NP\ V\ PP_1$
d. $PP_2\ V\ NP\ PP_1$
e. $PP_1\ NP\ V\ PP_2$
f. $PP_1\ V\ NP\ PP_2$
g. $NP\ V\ PP_1\ PP_2$
h. $V\ NP\ PP_1\ PP_2$

Cases (17a) and (17h) are those that could be covered by the headedness parameter, since they are head-final and head-initial respectively. These are clearly evidenced in natural language by, for instance, Japanese (head-final) and English (head-initial). Cases (17d) and (17g) represent the languages already discussed in this paper, where (17d) is Chinese and (17g) is Kpelle. The question now is raised: are the other four predicted word orders exemplified in natural language? Cases (17b) and (17e) are the mirror images of Kpelle and Chinese respectively, suggesting that perhaps these orders should exist since they simply represent the opposite values for each parameter.

The two remaining orders, (17c) and (17f), however, are a type very different from the other six. In all the other six word orders, the same results may be achieved by stating the value of one or two parameters, (i) either Case or θ-direction, and/or

(ii) the headedness parameter. For example, Chinese (17[d]) assigns θ-roles to the right but is otherwise head-final, and Japanese (17a) need specify only head-final. All three parameters must be specified to describe the word orders in (17c) and (17f), however. Because of this difference, and the lack of evidence for languages of the type of (17c) and (17f), I propose that all six other orders are possible, but that (17c) and (17f) must be ruled out in principle.

    [. . .]

As noted above, word orders of types (17c) and (17f) are very different in kind, since *both* subdomain parameters must be specified as well as the headedness parameter. This raises one question: how many of these parameters may be set for any language? Another question that is raised, within this system as well as with the headedness parameter, is: Can parameters be set for every different category? What I claim is that only one subdomain parameter (i.e. Case or θ-direction) may be set per language, and that only the default parameter of headedness may vary from category to category. We will see below in more detail how this works. What is important is that the power of the parameters and the range of possible language types will be restricted.

    [. . .]

2.2 The restrictions

In talking of Chinese and Kpelle, we have been discussing two parameters for each language: Chinese assigns θ-roles to the right but is otherwise head-final. Kpelle assigns Case to the left but is otherwise head-initial. It was suggested above that perhaps only two parameters at most might be specified, a subdomain direction and then a default parameter of headedness. While this works and would rule out (17c) and (17f), it adds a complication to the framework. It is generally assumed that grammatical rules cannot "count." In other words, a grammatical rule could never refer to "three bounding nodes" or "the fifth Comp." A restriction could only refer to "one" or "many." For example, subjacency restricts movement to one bounding node. If we were allowed to specify only one parameter, however, it would seem that we could not distinguish Chinese from English, or Kpelle from Japanese. The relevant word orders are given in (24):

(24)

		CASE	THETA	HEADEDNESS
a.	PP$_2$ V NP PP$_1$ Chinese	_____	right	final
b.	V NP PP$_1$ PP$_2$ English	_____	right	initial
c.	NP V PP$_1$ PP$_2$ Kpelle	left	_____	initial
d.	PP$_1$ PP$_2$ NP V Japanese	left	_____	final

Without specifying the headedness parameter as well as the direction of θ-role assignment for Chinese and Case assignment for Kpelle, there would be no way to distinguish (24a) from (24b) or (24c) from (24d). However, one should also note that neither Japanese nor English needs to specify a subdomain, since their VPs are purely head-final and head-initial respectively. Let us assume, then, that subdomains

need to be stipulated only when they are separated from other elements of the maximal projection. This means that a specification of Case assignment to the right carries with it the information that the language is otherwise head-final. If the language were head-initial, the Case domain would not need to be specified in a directional parameter.

We can now postulate the following rest[r]ictions on word order parameters:

(25) a. If a subdomain direction is specified, this is all that may be specified.
          (This would be the marked case.)
      b. If no subdomain direction is specified, then a value must be given for
          headedness. (This is the unmarked case.) This value may vary across
          categories.

Below, the chart given in (17) is reinterpreted in the context of our restrictions. Possible word orders may all be distinguished by specifying only one parameter (capitalized in (26)). Further, word orders that are assumed to be impossible cannot be characterized using the restrictions above.

(26) Parameters:

	HEADEDNESS	THETA	CASE	LANGUAGE
a.	FINAL	Left	Left	Japanese
b.	Final	Left	RIGHT	Chinese (future)
c.	Final	Right	Left	*
d.	Final	RIGHT	Right	Chinese (present)
e.	Initial	LEFT	Left	Kpelle (past)
f.	Initial	Left	Right	*
g.	Initial	Right	LEFT	Kpelle (present)
h.	INITIAL	Right	Right	English

] . . .]
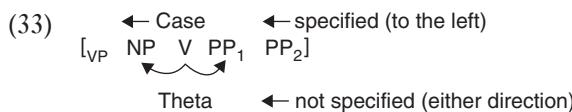
2.4 Unspecified parametric values

One final note should be made concerning parameters, in particular parameters which have a possible setting but for which no value has been given. For example, in Kpelle Case direction has been set, and headedness direction follows; however, nothing has been said about the direction of θ-role assignment. My assumption is that unspecified parameters have the option of both possible values. This is important for languages like German, where θ-marked complements and Case-marked complements may appear either to the left of their heads (as in VPs) or to the right (as in PPs). Since German specifies neither subdomain, nothing restricts the direction in which Case-marking or θ-marking apply.

This also allows for an alternative analysis concerning which level is affected by the setting of word order parameters (see §1.3). Davis (1987) suggests that Case assignment is just as important as θ-role assignment in creating initial phrase structures. In her view, the direction of Case assignment would be operative

immediately in determining the placement of complements of the verb. Direct objects in Kpelle, then, would be base-generated in a position to the left of the verb. Since θ-roles can be assigned in either direction, no problem is created. Theta-assignment would take whichever direction was necessary in order to "find" the appropriate complement. This is shown in (33) (cf. (16)):

(33)   ← Case   ← specified (to the left)
        [$_{VP}$ NP V PP$_1$ PP$_2$]

              Theta   ← not specified (either direction)

Conclusion

In this paper I have investigated the headedness parameter with the intention of solving the problem of head-initial languages like Chinese and Kpelle. I began by proposing two new parameters. These broke maximal projections into smaller domains, created by relations already incorporated into the theory such as Case assignment and θ-role assignment. While adding enough flexibility to the system to account for head-internal orders, these new parameters created a problem of overgeneration. Not only could Chinese and Kpelle be described, but nonexistent word orders would be predicted as well. In order to avoid this problem, restrictions on word order parameters were proposed to constrain their generative power. By restricting parameters in this way, I hope to have not only solved the empirical problem of word order in Chinese and Kpelle, but also met my own requirements for a parametric system outlined at the beginning of this paper.

16.3   Questions pertaining to Travis (1989)

1   The typological tradition that originates with Greenberg (1966) makes use of terms like "VO language" and "OV language." Travis shows that some languages cannot be classified in this way, and that recourse to finer-grained properties is necessary. To what extent might Travis find Diesing's (1992, 1997) later work on definite and indefinite objects congenial, and why?

2   Travis is very conscious of the danger of overgeneration when it comes to crosslinguistic word order patterns, and suggests ruling out in principle "PP$_2$ NP V PP$_1$," where PP$_2$ is an adjunct PP, and PP$_1$ is a PP of the sort that can be thought to bear a theta-role (e.g. a directional PP). Choose two Indo-Iranian languages and discuss the extent to which they do or do not display "PP NP V PP" order.

3   Travis proposes parameters specifying the direction of Case assignment and the direction of theta-assignment, and takes those parameters to be set in a consistent direction (either left or right) for any given language. Discuss the relation between this proposal of Travis's and the idea of Borer (1984) and Chomsky (1995, Ch. 3) that parameters are limited to features of functional heads. (Extra credit: Discuss the question of how we might try to restrict that set of features, bringing in Rizzi 2011.)

4 Baker's (1988) UTAH principle and much of Cinque (1999) amount to saying that external merge order is uniform across languages, in the sense that (within a simple sentence) if A is merged prior to B in one language, then A must (whenever A and B co-occur) be merged prior to B in all languages. Against this background, discuss the question whether Kayne's (1994, 2011) antisymmetry proposal necessarily leads to the conclusion that all crosslinguistic word order differences are reducible to crosslinguistic differences concerning internal merge (movement).

5 With the previous question in mind, how should we evaluate the following claim?: All crosslinguistic morpheme order differences are reducible to crosslinguistic differences concerning internal merge (movement).

6 Crosslinguistic comparisons, no matter what the theoretical framework, require knowing what morpheme in one language corresponds to/is the counterpart of what morpheme in the next language. Pick two not very closely related language families and then pick one language from each of them. With those two languages in hand, take five functional (as opposed to lexical) morphemes from one of them and find their counterparts in the other. Discuss the importance of whatever difficulties you may have had.

7 As a variant of the preceding question, pick two very closely related languages and proceed in the same way. To what extent was the task easier with more closely related, as opposed to less closely related, languages? Discuss the significance of whatever difference in degree of difficulty you might have discovered.

8 Latvian has its genitives preceding the associated noun, as opposed to Russian, whose genitives follow the noun. To what extent should this Latvian–Russian difference be assimilated to the English-internal difference between *my best friend's sister* and *the sister of my best friend*? To what extent does this English-internal pair make the same point as Travis's discussion of Chinese and Kpelle VP-internal word order? Discuss, in addition, the relevance to Greenbergian word order generalizations of the English-internal contrast between *The assassination of John Smith took place yesterday* and ??*The car of John Smith was involved in an accident yesterday*.

9 Word order statements concerning PPs typically rest on the assumption that PP is a constituent formed by externally merging P with what we call its object. Kayne (1999, 2002, 2004) argues, in contrast, that (certain) prepositions are merged outside VP and are brought together with what we call their object by internal merge (movement). How might Kayne then approach the question of the difference between postverbal PPs and preverbal PPs (the latter being much more common in, for example, German and Dutch than in English)? (Partial hint: Look at Kayne 1998.)

10 As Travis emphasizes, the position of argument-like PPs (such as directionals) is often different from the position of adjunct PPs. Why exactly is that, do you think?

11 Evaluate the argument given in Kayne (1994, preface) to the effect that (even) Japanese cannot be uniformly head-final, in any sense of the term. In addition, discuss the possibility of reinterpreting Ivana and Sakai (2007) on Japanese

honorific *o* in such a way that *o* would be a head that stays initial relative to the associated (nominalized) verb, while having all the arguments of the verb move past it (i.e., move past honorific *o*).

12 Whether or not Diesing (1992, 1997) is correct in taking indefinite objects to remain in situ, her argument that definite objects must end up higher than indefinite ones leads one to wonder why there are few, if any, languages that rigidly have OV order with definites and VO order with indefinites. Taking Diesing to be right in her claim about the relative height of definite and indefinite objects, propose an account of why the verb does not rigidly separate them. Bring into the discussion Jayaseelan (2001, 2010).

13 Although English is often called a VO language, it readily allows OV order in what are called compounds, for example, in *John is an avid magazine-reader, That magazine-reading linguist over there is John*. How might one account for **an avid read-magazine-er*, **that read-magazine-ing linguist*? (Hint: Bring in Biberauer et al. to appear.) Given that English allows OV in such compounds, why exactly does it not allow OV as a matter of course?

14 Although questions concerning the order of elements within a word are often considered distinct from word order questions per se, Greenberg (1966, Universal 27) had already noted a correlation between affix order (prefix vs. suffix) and adposition order (preposition vs. postposition). In that spirit, and whether or not you think of compounds as being words, investigate the following conjecture: OV order within compounds (as in English) is never found in V-initial languages. (Extra credit: Why might this correlation hold?)

15 Williamson and Blench (2000, 31ff.) give Bantoid, Cross River, Central Nigerian, and West Benue-Congo as containing (some) SVMO languages (where M = modal). How might one integrate such languages into Simpson's (2001) account of South East Asian languages having VP-movement past (neg+)modal?

References

Baker, M. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Biberauer, T., A. Holmberg, and I. Roberts. to appear. A syntactic universal and its consequences. *Linguistic Inquiry*.

Borer, H. 1984. *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Dordrecht: Foris.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. New York and Oxford: Oxford University Press.

Davis, M. 1987. Base-generating NP-structure. *McGill Working Papers in Linguistics 1*, Montreal: Dept. of Linguistics, McGill University, pp. 85–120.

Diesing, M. 1992. *Indefinites*. Cambridge, MA: MIT Press.

Diesing, M. 1997. Light verbs and the syntax of aspect in Yiddish. *The Journal of Comparative Germanic Linguistics* 1: 119–156.

Gay, J., & W. Welmers. 1971. *Mathematics and Logic in the Kpelle Language; and a First Course in Kpelle*. Ibadan, Ivory Coast: Institute of African Studies, University of Ibadan. Occasional Publication 21.

Givón, T. 1975. Serial verbs and syntactic change: Niger-Congo. In C. Li (ed.), *Word Order and Word Order Change*, Austin: University of Texas Press, pp. 47–112.

Greenberg, J. H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (ed.), *Universals of Language*, Cambridge, MA: MIT Press, pp. 73–113.

Huang, C.-T. J. 1982. *Logical relations in Chinese and the theory of grammar*. Doctoral dissertation, MIT.

Ivana, A., and H. Sakai. 2007. Honorification and light verbs in Japanese. *Journal of East Asian Linguistics* 16: 171–191.

Jayaseelan, K. A. 2001. IP-internal topic and focus phrases. *Studia Linguistica* 55: 39–75.

Jayaseelan, K. A. 2010. Stacking, stranding, and pied-piping: A proposal about word order. *Syntax* 13: 298–330.

Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. S. 1998. A note on prepositions and complementizers. Paper contributed to the Chomsky Virtual Celebration. [Also in revised form in Kayne (2000, Ch. 15).]

Kayne, R. S. 1999. Prepositional complementizers as attractors. *Probus* 11: 39–73. [Reprinted in Kayne (2000, Ch. 14).]

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2002. On some prepositions that look DP-internal: English *of* and French *de*. *Catalan Journal of Linguistics* 1: 71–115. [Reprinted in Kayne (2005, Ch. 7).]

Kayne, R. S. 2003. Antisymmetry and Japanese. *English Linguistics* 20: 1–40. [Reprinted in Kayne (2005, Ch. 9).]

Kayne, R. S. 2004. Prepositions as probes. In A. Belletti, *Structures and Beyond: The Cartography of Syntactic Structures* 3, New York and Oxford: Oxford University Press, pp. 192–212.

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2011. Why are there no directionality parameters? In M. Byram Washburn, K. McKinney-Bock, E. Varis, A. Sawyer, and B. Tomaszewicz (eds), *Proceedings of the 28th West Coast Conference on Formal Linguistics*, Somerville, MA: Cascadilla Proceedings Project, pp. 1–23.

Koopman, H. 1984. *The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar*. Dordrecht: Foris.

Lasnik, H., and M. Saito. 1992. *Move-alpha: Conditions on Its Application and Output*. MIT Press, Cambridge.

Li, C., and S. Thompson. 1975. The semantic function of word order: a case study in Mandarin. In C. Li (ed.), *Word Order and Word Order Change*, Austin: University of Texas Press, pp. 163–95.

Li, C., and S. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley & Los Angeles: University of California Press.

Rizzi, L. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Rizzi, L. 2011. On the elements of syntactic variation. In *Working papers of Centro Interdipartimentale di Studi Cognitivi sul Linguaggio (CISCL)*. Università di Siena. Distributed by MIT Working Papers in Linguistics.

Simpson, A. 2001. Focus, presupposition and light predicate raising in East and Southeast Asia. *Journal of East Asian Linguistics* 10: 89–128.

Sportiche, D. 1981. Bounding nodes in French. *The Linguistic Review* 1: 219–246.

Stowell, T. 1981. *Origins of phrase structure*. Doctoral dissertation, MIT.

Williamson, K., and R. Blench. 2000. Niger-Congo. In B. Heine and D. Nurse (eds), *African Languages: An Introduction*, Cambridge: Cambridge University Press, pp. 11–42.

17

Negative Heads and the Neg Criterion

Liliane Haegeman and Raffaella Zanuttini

1991

17.1   Introduction

As mentioned in the introduction to Kayne (1989, Ch. 14 of this volume), the notion of agreement refers to the presence of a morpheme expressing the same value of the same feature on multiple elements in a given local domain. In cases of subject–verb agreement and past participle agreement, an argument and a verbal form exhibit the same value for certain ϕ-features. Agreement in number and gender is also found in the nominal domain, on nouns, determiners, demonstratives, and adjectives. Is agreement in natural language limited to ϕ-features?

Many languages exhibit the presence of a morpheme expressing a negative feature on multiple elements, in a given local domain. Such cases, often referred to as NEGATIVE CONCORD, are exemplified in (1):

(1) a. . . .da Valère *niemand nie* kent. (West Flemish)
         . . .that Valère nobody not knows
         '. . .that Valère does not know anybody.'
     b. . . .da Valère *niemand nie en*-kent.
         . . .that Valère nobody not neg knows
         '. . .that Valère does not know anybody.'

In these examples, multiple negative constituents co-occur in a clause yet do not contribute multiple instances of negation to the interpretation of the sentence.

In *Negative Heads and the Neg Criterion*, Haegeman and Zanuttini pursue a connection between agreement and negation, and ultimately negative concord. They notice that, in West Flemish, the negative scope marker *en* (cf. (1b)) must

co-occur with another negative element that has sentential scope (or else it leads to ungrammaticality). They express this observation in terms of a licensing requirement: *en*, a functional head that bears a negative feature, must be licensed by the presence of another element with a negative feature in its specifier. This Spec-head configuration can be instantiated either by overt or covert movement of a negative operator. Moreover, they observe that a negative constituent can take sentential scope in West Flemish only if it raises out of the VP. They express this observation by suggesting that such a constituent needs to raise to the specifier of a functional head bearing a negative feature. These two requirements are combined in the Neg Criterion:

(2)   The Neg Criterion
       a.   Each Neg X^o must be in a Spec-Head relation with a Negative operator.
       b.   Each Negative operator must be in a Spec-Head relation with a Neg X^o.

In this view, scope bearing elements are not simply required to c-command the portion of the clause over which they have scope (as in May 1985: 17); they must also instantiate a Spec-head configuration with a particular functional head. Negative concord, this paper assumes, has as a precondition that the negative constituents have the same scope. In West Flemish, all the negative constituents that have sentential scope move (or adjoin) to the specifier of the negative head, where they undergo a process of factorization that results in a single instance of negation (see Zanuttini 1991; Haegeman 1995; Haegeman and Zanuttini 1996). The relation between negative concord and agreement, in particular as instantiated in the formal notion of AGREE, has later been explored in detail in Zeijlstra (2004) (see also Haegeman and Lohndal 2010).

    The proposal that an operator is required to move to the specifier of a head bearing the same kind of feature has also been made for other operators. Building on May's (1985: 17) "*wh*-criterion," Rizzi (1991, 1996) proposes the following:

(3)   The wh Criterion (Rizzi 1996: 64)
       a.   A *wh*-operator must be in a Spec-Head configuration with $X^o[+wh]$.
       b.   An $X^o[+wh]$ must be in a Spec-Head configuration with a *wh*-operator.

(In later work, Rizzi suggested that a Spec-head relation is required for other elements, as well, such as focalized and topicalized constituents, and subjects; see Rizzi 1997, 2006b; Rizzi and Shlonsky 2007). This kind of Spec-head requirement has been applied in Beghelli and Stowell (1997, Ch. 23 of this volume) to produce a precise account of the scope-taking possibilities of different classes of quantifiers. Focusing on the syntactic distribution of negative operators, Kayne (1998) has taken it one step further, making the radical claim that, in English and across languages, they always raise to the specifier of a negative head *overtly*; when this movement is not easily detectable from the linear order, it is because other movements have taken place, making it hard to see.

17.2 From "NEGATIVE HEADS AND THE NEG CRITERION"

1 Introduction

Languages differ in at least two ways with respect to the expression of sentential negation (cf. Ouhalla 1990; Zanuttini 1991). Some languages, like French, have a bipartite negation consisting of a pre-verbal negative clitic and another negative marker (1a,b); in other languages, like standard Dutch (1c,d) sentential negation is expressed by a single negative marker:

(1) a. Elle n'a   pas   vu   son père.
         she NEG-has NEG seen her father
         'She did not see her father.'
     b. Elle n'a   vu   personne.
         she NEG-has seen no one
         'She did not see anyone.'
     c. . . . dat zij haar vader niet gezien heeft.
              that she her   father not   seen   has
         '. . . that she did not see her father.'
     d. . . . dat zij niemand gezien heeft.
              that she no one   seen   has
         '. . . that she did not see anyone.'

A second difference concerns the way in which multiple occurrences of negative constituents are interpreted. In the Romance type languages two or more negative constituents in a clause express one single instance of sentential negation. This phenomenon is referred to as Negative Concord (NC) and illustrated in (2a). In Germanic type languages multiple negative constituents often each carry their own negative force, which means that one negative constituent will cancel the next one, resulting in Double Negation (DN) readings as in standard Dutch (2b).

(2) a. Je n'ai   jamais rien   dit   à   personne.
         I NEG have never   nothing said to nobody
         'I never told anyone anything.'
     b. Ik heb niemand niet uitgenodigd.
         I have no one   not   invited
         'I did not invite no one.'

We draw a distinction between pre-verbal negative elements such as French *ne* and the other negative constituents such as *jamais*, *rien* and *personne*. French *ne* is a clitic, i.e. a head; negative constituents such as *personne* and *jamais* are phrasal constituents. The relation between *ne* and a negative constituent is one between a head and a phrase; that between the negative constituents which enter into a NC reading is one between phrases. We reserve the term Negative Concord for the relation between phrasal negative constituents.

On the basis of an analysis of the distribution of the negative head and the negative constituents in West Flemish (WF) we propose an account for sentence negation as part of a theory of scope. The licensing of negative elements is interpreted in terms of the Neg Criterion, a well-formedness condition on LF. The Neg Criterion is an instantiation of a general condition on the distribution of scope-bearing elements, the Affect Criterion, proposed in Rizzi (1991).

Section 2 offers a description of sentential negation in WF, section 3 introduces the Neg Criterion and section 4 summarizes the paper.

2 Negation in West Flemish

2.1 Sentential negation

Sentential negation can be expressed by the negative marker *nie*:

(3) a. . . . da Valère gisteren nie tegen zen voader geklaapt eet.
            that Valère yesterday not against his father talked   has
         '. . . that Valère did not talk to his father yesterday.'

Nie precedes VP-complements such as the PP *tegen zen voader* in (2). This suggests that *nie* is in a position left-peripheral to VP (cf. sect. 2.2). Sentential negation can also be expressed by negative constituents:

(4) a. . . . da Valère niemand/geen mens kent.
            that Valère nobody/no   person knows
         '. . . that Valère does not know anybody.'
     b. . . . da   Valère over niemand klaapt.
            that Valère about no one   talks
         '. . . that Valère does not talk about anyone.'

One might be tempted to assume that the negative constituents in (4) occupy their base position. However, we will argue that the negative constituents in (4) are subject to the generalized scrambling which affects negative constituents with sentential scope.

The first argument for the scrambling of negative elements with sentential scope is based on NC. In (5a) *nie* enters in a NC relation with *niemand*:

(5) a. . . . da   Valère niemand nie kent.
            that Valère nobody   not knows
         '. . . that Valère does not know anybody.'
     b. . . . da Valère nie niemand kent.
         '. . . that Valère doesn't know nobody.'   (DN)

If *nie* is left-peripheral to VP, then *niemand* in (5a) must have been scrambled out of its base position and adjoined to a dominating maximal projection. When *niemand* occurs in a position to the right of *nie* it does not enter into a NC relation with *nie*; rather, as (5b) illustrates, this order gives rise to a DN interpretation.

Haegeman & Zanuttini [1996] propose that one condition for negative constitu-
ents to enter into a NC relation is that their scope domains be identical. In (5b) NC
is blocked because the scope of *niemand* is confined to the containing XP, whereas
nie takes sentential scope. In its scrambled position, the negative constituent takes
the same scope as *nie* and the two negative constituents enter into NC. We propose
that when negative constituents like *niemand* take sentential scope they have been
scrambled out of the VP.

[. . .]

2.2   The functional projection NegP

In addition to the negative elements discussed above, WF allows for a preverbal
negative clitic *en*:

(8)   a.   . . .   da   Valère die   boeken nie an zijn voader (en-)toogt.
                   that Valère those books   not to his father   en- shows
             '. . . that Valère does not show his father those books.'
      b.   . . .   da   Valère niemand (en-)kent.
                   that Valère nobody   en- knows
             '. . . that Valère does not know anyone.'
      c.   . . .   da Valère niemand nie (en-)kent.
      d.   . . .   da   Valère zou   willen nie (*en)-werken.
                   that Valère would want   not en work
             '. . . that Valère would like not to work.'
      e.   *. . . da Valère die boeken an zen voader en-toogt.   (cf. (8a))
      f.   *. . . da   Valère en-wist da   zen voader geen geld   oat.
                   that Valère en knew that his father no   money had
      g.   *. . . da   Valère an niemand zei da   Marie ziek en-was.
                   that Valère to nobody   said that Marie ill   en was

En co-occurs with *nie* (8a), with a single negative constituent (8b), or with multiple
negative constituents in a NC relation (8c). In all three examples the negative constitu-
ents have sentential scope. *En* cannot co-occur with an infinitive (8d) (see Haegeman
1991b). (8e) shows that *en* cannot express sentential negation on its own: in the absence
of another negative constituent the use of *en* is ungrammatical (cf. Haegeman 1991b).

There are locality constraints on the relation between *en* and the negative con-
stituent. In (8f) the intervening clause boundary apparently blocks the relationship
between *en* and the negative constituent. Similarly, in (8g) the negative constituent
in a higher clause fails to relate with *en* in the lower clause. Let us say as a first
approximation that the negative clitic *en* is licensed by a clausemate negative con-
stituent with sentential scope. The WF negative pattern is like the French pattern
(cf. Ashby 1981; Moritz 1989):

(9)   a.   Jean (n') a *(pas) téléphoné.
           Jean ne has pas telephoned
           'Jean did not telephone.'

For WF we assume the clause structure in (10) proposed in Haegeman (1991a,1991b). The analysis adopts the split Infl framework of Pollock (1989) in the adaptation of Belletti (1990). WF *en*, like French *ne*, heads NegP, which expresses sentential negation. NegP dominates TP. The relation between T and Neg is independently argued for in Zanuttini (1990). We also propose, following Haegeman & Zanuttini [1996] that in the absence of overt *en/ne*, Neg^0 is non-overt. In the remainder of this paper we only discuss the overt head *en*. Our tacit assumption throughout is that the discussion applies equally to the non-overt head Neg^0. For our purposes it is immaterial whether NegP is head-final, as in (10), like TP and AgrP, or head-initial, like CP.

(10) $[_{CP} [_{AgrP} [_{NegP} [_{TP} [_{VP} V] T] Neg^0] Agr]]$

2.3 The relation between the negative head and the negative constituents

2.3.1 Clausemate negative constituents

WF *en* is licensed by a clausemate negative constituent:

(11)   . . . da Valère die boeken *(nie) an zen voader en-toogt. (cf. (8a))

Only markers of sentential negation license *en*. In (12) constituent negation *nie* cannot license *en*:

(12)   . . . dan 'k $[_{PP}$ tegen $[_{NP}$ nie al d' joengers]] (*en)-klapen.
              that-I   against   not all the children   en   talk
       '. . . that I don't talk to all the children.'

   [. . .]

We have argued that all negative constituents with sentential scope scramble out of VP. If *en* is licensed by a negative constituent with sentential scope, then it can only be licensed by a scrambled negative constituent:

(15)   a.   . . . da Valère me niets   ketent (en)- is.
                   that Valère with nothing satisfied en   is
             '. . . that Valère is not satisfied with anything.'
       b.   . . . da Valère ketent me niets (*en)-is.

For two or more negative constituents to enter into NC, they must all scramble:

(15)   c.   . . . da Valère me   niets   nie ketent   en-is.
                   that Valère with nothing not satisfied en is
             '. . . that Valère is not satisfied with anything.'   (NC)
       d.   . . . da Valère nie ketent me niets en-is.
             '. . . that Valère is dissatisfied with nothing'   (DN)

The S-structure requirement on the position of negative constituents must be explained, though. Note that (15e) shows that the VP-internal S-structure position allows for LF-raising in the case of *wh*-constituents.

(15) e.   k weten nie wien dat er   nie ketent   me   wien   is.
           I know not who that there not satisfied with whom is
           'I don't know who is dissatisfied with whom.'

                                                   (cf. discussion of (37a))

   [. . .]

The starting point for our analysis is the contrast in (20). *En* can be licensed by a negative constituent in SpecCP; it cannot be licensed by an extraposed negative constituent:

(20) a.   Over niemand en-klaapt Valère tegen myn.
           about no on   en talks Valère against me
           'Valère does not talk about anyone to me.'
      b.   . . . da Valère (*en)-klaapt over niemand.
               that Valère en talks about no one
           '. . . that Valère talks about NO ONE.'

Our hypothesis is that the distinctive property of (20a) is that the negative constituent *over niemand* and the negative head *en* are in a Spec-Head relation.

3.2 The Neg criterion

In section 3.1. [in the full article] we identify the Spec-Head relation as crucial for *en*-licensing. The link between *en* and the negative constituent in (20a) is like that between a *wh*-complementizer and a *wh*-phrase. It is a general property of the grammar of English that *wh*-constituents give interrogative force to a sentence as a result of S-structure movement to SpecCP. Consider (21a). *Wonder* selects an interrogative CP, whose head C^0 carries the feature [+*wh*] (cf. Rizzi 1991). The *wh*-operator *what* moves to SpecCP where it ends up in a Spec-Head relation with C^0. In French (21b) *quoi* is not in SpecCP at S-structure: we assume that the *wh*-constituent moves to SpecCP at LF.

(21) a.   I wonder $[_{CP}$ what$_i$ $[_C$ [+*wh*] $[_{IP}$ you did t]]]
      b. $[_{IP}$ Tu as fait quoi]?

To account for the distribution of *wh*-phrases, Rizzi (1991:2) proposes the *wh Criterion* (cf. May 1985):

(22) *The wh Criterion*
      a.   Each *wh*-X^0 must be in a Spec-Head relation with a *wh*-operator;
      b.   Each *wh*-operator must be in a Spec-Head relation with a *wh*-X^0.

Rizzi says "[22] can be looked upon as a criterion of wellformedness on LF express-ing the way in which *wh*-expressions are assigned scope" ([1990]:378). (22) accounts for the root I-to-C movement in (21c):

(21) c. Who did you see?

The assumption is roughly (cf. Rizzi 1991 for details) that the feature [+*wh*] is instantiated on I; as a result of the movement of I to C, *who* will be in the correct Spec–Head configuration with the [+*wh*] head *did*. Rizzi proposes that in English the *wh Criterion* applies as early as S-structure (1991:3). The question arises how to account for *wh*-in situ in English multiple *wh*-questions.

(21) d. Who did you tell what?

In (21d) *what* is not in a Spec–Head relation with the *wh*-feature on *did*. Before discussing the example we need to add the following definitions:

(23) *Wh operator*:
      A *wh*-phrase in a scope position (Rizzi 1991:10)

(24) *Scope position*:
      a left–peripheral A′ position (either a Spec or an adjoined position). This excludes right–peripheral positions and the base-generated position of VP adverbials. (1991:10)

[. . .] Rizzi says:

> It is also necessary to assume that the functional definition . . . holds at DS and SS, whereas at LF it is superseded by a stronger principle according to which all elements endowed with intrinsic quantificational force are operators at this level, and must be moved to an appropriate scope position. (1991:22)

By (23) – (24) positions created by Heavy NP Shift are not scope positions (cf. Rizzi 1991:10):

(25) Which of the students borrowed *t* from you which of these theses?

Rizzi's (33)

Consider (26):

(26) a. I would do that in no case.
      b. *In no case I would do that.
      c. In no case would I do that.

Rizzi (1991:11) proposes that the I to C movement in (26c) can be interpreted as the result of a Spec–Head requirement on the relation between a negative head and a negative operator, which we formulate in (27):

(27)   *The Neg Criterion*
        a.   Each Neg X^0 must be in a Spec-Head relation with a Negative operator;
        b.   Each Negative operator must be in a Spec-Head relation with a NegX^0.

Like the *wh*-Criterion, the Neg Criterion (27) is a condition on LF. The Neg Criterion may apply as early as S-structure in some languages, while for others LF is the relevant level. WF illustrates the former, French or Italian illustrate the latter type.

3.3   The Spec-Head configuration at S-structure

In WF (28) the Neg Criterion is satisfied at S-structure. The Spec-Head relation obtains between *over niemand* in SpecCP and *en*, which is part of the complex C head containing the V, its Agr and T-inflection, and the negative head.

(28)   Over   niemand   en-klaapt   Marie   tegen   myn.
        about   nobody   en talks   Marie   against   me
        'Marie doesn't talk about anyone to me.'

[. . .]

For negative constituents which are not in SpecCP the Spec-Head configuration is between the base position of *en*, Neg^0, and SpecNegP. The obligatory scrambling of the negative constituents suggests that in WF the Neg Criterion applies as early as S-structure. If the Neg Criterion could be met at LF it would be surprising that a negative constituent in its base position, or in the extraposed position, two sites which allow *wh*-raising at LF, cannot license *en*.

[. . .]

In WF, the Neg Criterion has to be satisfied in full at S-structure. Either the negative constituent occupies SpecCP and is in a Spec-Head configuration with the S-structure position of negative head. Alternatively, the negative constituent is in SpecNegP or adjoined to NegP and is in a Spec-Head configuration with the trace of the negative head.

Looking at these two possibilities more schematically we note the following. The negative head, *en* in its overt form, is cliticized to the finite verb. In root clauses it is dominated by C; in embedded clauses it is dominated by the highest functional head, say Agr. Let us say that the movement of Neg^0 creates a Neg–chain.

[. . .]

Summary and problems for future research

On the basis of the distribution and interpretation of negative constituents in WF we develop a theory of negation as part of a more general theory of scope. The Neg Criterion, requires that negative operators and negative heads achieve a specifier-head configuration. In scrambling languages this requirement has to be met at S-structure; in others the relevant level is LF. In WF scrambled negative operators either move to SpecNegP or adjoin to NegP. When several negative operators cluster

around NegP they will be subject to Neg-factorization, similar to *wh*-absorption in multiple *wh*-questions, which results in a NC reading.

17.3   Questions pertaining to Haegeman and Zanuttini (1991)

1   Haegeman and Zanuttini, in discussing West Flemish, note the existence of (embedded) negative sentences like . . .*da Valère niemand nie kent* ('. . . that V nobody not knows'). As in many languages, the presence of negative *nie* ('not') here does not lead to a double negation reading; the interpretation is that of . . .*that Valère knows nobody*. Haegeman and Zanuttini take the fact that *nie* intervenes between object *niemand* ('nobody') and the verb to indicate that *niemand* has, in such sentences, scrambled leftward out of its original position. Discuss the relation between this proposal of Haegeman and Zanuttini's and Whitman's (2005) analysis of 'SONegV' languages, as well as the relation between Whitman's use of remnant movement and Kayne's (1998) analysis of English sentences containing *nobody*.

2   Whitman (2005) says that preverbal negative morphemes in OV languages are always uninflected for agreement (much as *nie* is uninflected in West Flemish). This would contrast with Finnish, which does have negation inflected for agreement, as discussed by Mitchell (1991). On the other hand, Vilkuna (1998: 216) has Finnish having OV order in nonfinite adjuncts. Suggest one or more finer-grained reformulations of Whitman's proposal that would make it compatible with Finnish; bring in Brattico and Huhmarniemi (2006).

3   Discuss the possible relation between Finnish agreeing negation and Roberts's (2000) proposal that English *n't* is restricted to finite contexts. To what extent does Roberts's proposal cover the other West Flemish negation morpheme *en*, as discussed by Haegeman and Zanuttini? Find as many other languages as possible in which certain negative morphemes are restricted to finite contexts and discuss the learnability challenge associated with such a restriction.

4   Haegeman and Zanuttini argue that "all negative constituents with sentential scope scramble out of VP" in order to meet the Neg Criterion, which requires that "each Negative operator must be in a Spec-Head relation with a NegX⁰." Haegeman and Zanuttini take this obligatory scrambling to be, in West Flemish, an instance of overt movement. For languages like English, in which an object negative phrase seems to remain within VP, as in *Today, John has seen nobody*, they take the required scrambling to take place covertly, at the level of LF (Logical Form). Evaluate Kayne's (1998) subsequent counterproposal, which takes object *nobody* in such sentences to have scrambled overtly, despite initial appearances.

5   Haegeman and Zanuttini's Neg Criterion is inspired by Rizzi's (1996) *wh*-Criterion (despite the later official publication date of the latter). Haegeman and Zanuttini discuss certain discrepancies between their Neg Criterion and Rizzi's *wh*-Criterion; find as many others as you can and discuss possible reasons for these discrepancies, including the ones mentioned by Haegeman and Zanuttini.

6   To what extent does Rizzi's ([2006]) subsequent proposal for a Subject
    Criterion affect our understanding of the Neg Criterion?

7   French is a VO language in the usual informal sense of the term. Yet in certain
    cases, its closest counterpart of English 'nothing' must be preverbal, as
    in *Jean n'a rien lu* ('J neg has nothing read'). Contrary to West Flemish, in
    French this happens only when the negative phrase contains no lexical noun,
    as seen in *Jean n'a lu aucun livre* ('J neg has read no book') vs. **Jean n'a aucun
    livre lu*. How might this French distinction be expressed? Bring in Kayne's
    (1975, Ch. 1) discussion of French *tout* ('everything'), as well as Poletto and
    Pollock (2005), and the syntax of English *who else*.

8   What are the key differences, in the area of negation, between West Flemish,
    as discussed by Haegeman and Zanuttini, and the fairly closely related
    Afrikaans, as discussed by Biberauer (2008)? (Extra credit: Bring in Nkemnji
    1995.)

9   In simple negative sentences French has two negative morphemes, *ne* and *pas*,
    e.g., *Jean n'a pas lu ce livre* ('J neg has neg read this book'), one of which (*ne*)
    precedes the finite verb (and can sometimes fail to appear at all), with the
    other (*pas*) following that verb. In one very particular case of an adjunct,
    French allows both negative morphemes to precede the complementizer *que*,
    e.g., *pour ne pas qu'il parte* ('for neg neg that he leave' = '(in order) for him not
    to leave'/'in order that he not leave'), in a way that would not be possible with
    English *not* – **in order not for him to leave*, **in order not that he leave*.
    Piedmontese *nen*, a close counterpart of French *pas*, can mimic the French
    behavior, as in *per nen ch'a parta* ('for neg that. . .'), while Italian *non* cannot –
    *per non che (lui) parta.

    The cross-Romance generalization seems to be that for a single negative
    morpheme to be able to precede the complementizer, as above, that negative
    morpheme must be of the sort that follows the finite verb in simple sentences,
    as does *pas* in the first example given, and as does *nen* in Piedmontese. Italian
    non cannot precede the complementizer because Italian *non* always precedes
    the finite verb and never follows it. In addition French *ne* can precede the
    complementizer *que* only if *pas* does – **pour ne qu'il parte pas*.

    What might be the reason for this last property of French *ne*? And what
    might be the reason that prevents negative morphemes that are consistently
    preverbal (like Italian *non*) from preceding the complementizer (bring in
    Pollock 1989 and Zanuttini 1997)?

10 Payne (1985: 207) says that negative verbs are not found in verb-medial lan-
    guages. After explicating the notion "negative verb," discuss whether or not
    Finnish (as mentioned above in question 2) is a true counterexample to
    Payne's generalization. To what extent could English *ain't* be taken to be a
    counterexample? To what extent should negative verbs be related to sentences
    like *It's not the case that Mary is in Paris*?

11 Payne (1985: 234) says that Persian sentences of the form . . .many. . . neg +V'
    allow neg to scope over 'many' much more readily when 'many' is part of the
    object than when 'many' is part of the subject, despite the fact that neg
    follows both the subject and the object. If you thought that Kayne (2011)

was not right, how would you approach this problem from a directionality parameter perspective? How would you approach it from an antisymmetry perspective? In addition, discuss whether this asymmetry in Persian is or is not relatable to the English contrast (with stressed negation) between *We did not see nobody* (possible in contradiction to a preceding *You saw nobody*) vs. **Nobody did not see us* (hardly possible at all in contradiction to a preceding *Nobody saw you*).

12 The proposal by Roberts (2000) mentioned in question 3, to the effect that English *n't* is limited to finite contexts, does not by itself account for the incompatibility of *n't* with subjunctives, as seen in *It's essential that they not leave before noon* vs. **It's essential that theyn't leave before noon*. How might one try to find a common account for this restriction on *n't* and the following one: *They're not here* vs. **They'ren't here*?

13 The overt movement of negative phrases discussed by Haegeman and Zanuttini for West Flemish has a clear counterpart in Scandinavian languages – for recent discussion, see Engels (2012). Kayne (1998) argues that English *John has seen nobody* vs. **John has nobody seen* (the reverse of the judgments in Scandinavian) is compatible with such overt movement (and that there are advantages to taking English to have it), as long as English follows that overt negative phrase movement with an instance of VP-preposing past the landing site of the negative phrase. Why, though, does Scandinavian not mimic English in this respect? Why does West Flemish not mimic English in this respect (and the same for Dutch and German)?

14 Alongside English *too many books, too few books*, French has *trop de livres, trop peu de livres*. In the French pair, there is a *peu* that matches English *few*, whereas the *many* of *too many books* has no overt counterpart in French *trop de livres* (the *de* in French corresponds to English *of*, as a first approximation). A natural proposal, modeled on Jackendoff (1977: 152), would take French *trop de livres* to contain a silent *MANY*. Why does no variety of French have a silent *FEW* instead of a silent *MANY*? To what extent is this question relatable to the following fact?: English has *Mary took the book to/from John*; yet *Mary took John the book* can correspond only to the sentence with *to*, and not at all to the sentence with *from*. (Extra credit: To what extent is there a link to *They jumped off/on the table* vs. *They jumped off of/*on of the table*?)

15 In English, negative polarity items can generally not precede the associated negation, e.g., **Anybody didn't give us flowers yesterday*. On the other hand, a fairly similar example is given by Holmberg (2005: 547) as being acceptable in Finnish. How might one distinguish a language that in this respect is like English from one that in this respect is like Finnish? To what extent might there be a link to English sentences like *Tabs are being kept on us*, with an idiom chunk subject (bring in Chomsky 1995, Ch. 3)? (Extra credit: How might it be important that in idiom chunk sentences of this sort verb agreement is in number, but not in person?)

16 Discuss the way in which Leu's (2012) proposed decomposition of English *not* into (at least) three morphemes bears on the limitation of *n't* to finite contexts mentioned in questions 3 and 12. In addition, discuss the potential

relevance of Leu's proposal to Beghelli and Stowell's (1997) idea that *not* originates in the Event argument position.

17   Referring to work that goes back to Zwicky and Pullum (1983), Pullum (1997) states that "Even in the slowest and most careful American English speech, the negative auxiliaries are not just acceptable but effectively obligatory; the periphrastic substitutes *do not*, etc., sound forced and unnatural in speech." Whether or not this is true for all American English, it is certainly true for some, with the interesting twist that, even for those who reject (except with stress on the negation) **They do not like linguistics*, **He will not be there*, etc., it is possible to have *They're not there, He's not home*, etc., with reduced forms of *be*.

Discuss the possible relevance of the proposal of Leu's mentioned in the previous question to this set of facts, bringing in Bernstein and Tortora (2005) on reduced forms of *be*.

References

Ashby, W. 1981. The loss of the negative particle *ne* in French. *Language* 57: 674–687.

Beghelli, F., and T. Stowell. 1997. Distributivity and negation: The syntax of *each* and *every*. In A. Szabolcsi (ed.), *Ways of Scope Taking*, Dordrecht: Kluwer, pp. 71–109.

Belletti, A. 1990. *Generalized Verb Movement*. Turin: Rosenberg & Sellier.

Bernstein, J., and C. Tortora. 2005. Two types of possessive forms in English. *Lingua* 115: 1221–1242.

Biberauer, T. 2008. Doubling and omission: Insights from Afrikaans negation. In S. Barbiers, O. Koeneman, M. Lekakou, and M. van der Ham (eds), *Microvariations in Syntactic Doubling. Syntax and Semantics* 36, London: Emerald, pp. 103–140.

Brattico, P., and S. Huhmarniemi. 2006. Finnish negation, the EPP feature and the valuation theory of morphosyntax. *Nordic Journal of Linguistics* 29: 5–44.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Engels, E. 2012. Scandinavian negative indefinites and cyclic linearization. *Syntax* 15: 109–141.

Haegeman, L. 1991a. *Negative concord, negative heads*. Paper delivered at the Going Romance Conference, Utrecht.

Haegeman, L. 1991b. *Negation in West Flemish*. Ms., University of Geneva.

Haegeman, L. 1995. *The Syntax of Negation. Cambridge Studies in Linguistics* 75. Cambridge: Cambridge University Press.

Haegeman, L., and T. Lohndal. 2010. Negative concord and (multiple) Agree: A case study of West Flemish. *Linguistic Inquiry* 41: 181–211.

Haegeman, L., and R. Zanuttini. 1996. Negative concord in West Flemish. In A. Belletti and L. Rizzi (eds), *Parameters and Functional Heads: Essays in Comparative Syntax*. Oxford: Oxford University Press, pp. 117–179.

Holmberg, A. 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36: 533–564.

Jackendoff, R. 1977. *X-bar Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Kayne, R. S. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

Kayne, R. S. 1989. Facets of Romance past participle agreement. In P. Benincà (ed.), *Dialect Variation and the Theory of Grammar*, Dordrecht: Foris, pp. 85–103. [Reprinted in Kayne (2000, Ch. 3).]

Kayne, R. S. 1998. Overt vs. covert movement. *Syntax* 1: 128–191.

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2011. Why are there no directionality parameters? In M. Byram Washburn, K. McKinney-Bock, E. Varis, A. Sawyer, and B. Tomaszewicz (eds), *Proceedings of the 28th West Coast Conference on Formal Linguistics*, Somerville, MA: Cascadilla Proceedings Project, pp. 1–23.

Leu, T. 2012. The indefinite article – Indefinite? Article? In J. Fruehwald (ed.), *Penn Working Papers in Linguistics: Proceedings of PLC 35*, pp. 161–168. [Available at http://repository.upenn.edu/pwpl/vol18/iss1/19; accessed January 2013.]

May, R. 1985. *Logical Form*. Cambridge, MA: MIT Press.

Mitchell, E. 1991. Evidence from Finnish for Pollock's theory of IP. *Linguistic Inquiry* 22: 373–379.

Moritz, L. 1989. *Aperçu de la syntaxe de la negation en français et en anglais*. Thèse de licence, University of Geneva.

Nkemnji, M. 1995. *Heavy pied-piping in Nweh*. Doctoral dissertation, UCLA.

Ouhalla, J. 1990. Sentential negation, relativized minimality and the aspectual status of auxiliaries. *The Linguistic Review* 7: 183–231.

Payne, J. R. 1985. Negation. In T. Shopen (ed.), *Language Typology and Syntactic Description 1: Clause Structure*, Cambridge: Cambridge University Press, pp. 197–242.

Poletto, C., and J.-Y. Pollock. 2005. On Wh-clitics and Wh-doubling in French and some North Eastern Italian Dialects. *Probus* 16: 241–272.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Pullum, G. 1997. The morpholexical nature of English *to*-contraction. *Language* 73: 79–102.

Rizzi, L. 1990. Speculations on verb second. In J. Mascaro and M. Nespor (eds), *Grammar in Progress*, Dordrecht: Foris, pp. 375–386.

Rizzi, L. 1991. Residual verb second and the Wh-criterion. *Technical Reports in Formal and Computational Linguistics* 2, University of Geneva.

Rizzi, L. 1996. Residual verb second and the Wh-criterion. In A. Belletti and L. Rizzi (eds), *Parameters and Functional Heads*, Oxford: Oxford University Press, pp. 63–90. [Reprinted in Rizzi (2000, Ch. 9).]

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 281–337.

Rizzi, L. 2000. Comparative Syntax and Language Acquisition. London: Routledge.

Rizzi, L. 2006. On the form of chains: Criterial freezing and ECP effects. In L. Cheng and N. Corver (eds), *Wh-Movement: Moving On*, Cambridge, MA: MIT Press, pp. 97–134.

Rizzi, L., and U. Shlonsky. 2007. Strategies of subject extraction. In U. Sauerland and H. M. Gärtner (eds), *Interfaces + Recursion = Language?: Chomsky's Minimalism and the View from Syntax-Semantics*, New York: Mouton de Gruyter, pp. 115–160.

Roberts, I. 2000. Auxiliary reduction and negation reduction: A rough sketch. *Snippets* 1: 15–16.

Vilkuna, M. 1998. Word order in European Uralic. In A. Siewierska (ed.), *Constituent Order in the Languages of Europe*. Berlin: Mouton de Gruyter, pp. 173–233.

Whitman, J. 2005. Preverbal elements in Korean and Japanese. In G. Cinque and R. S. Kayne (eds), *The Oxford Handbook of Comparative Syntax*, New York and Oxford: Oxford University Press, pp. 880–902.

Zanuttini, R. 1990. Two types of negative markers. In Carter, J., R.-M. Déchaine, B. Philip, and T. Sherer (eds), *Proceedings of NELS 20*, Amherst: GLSA, University of Massachusetts, vol. 2, pp. 517–530.

Zanuttini, R. 1991. *Syntactic properties of sentential negation. A comparative study of Romance languages*. Doctoral dissertation, University of Pennsylvania.

Zanuttini, R. 1997. *Negation and Clausal Structure: A Comparative Study of Romance Languages*. New York and Oxford: Oxford University Press.

Zeijlstra, H. 2004. *Sentential Negation and Negative Concord*. Utrecht: LOT Publications.

Zwicky, A., and G. Pullum. 1983. Cliticization vs. inflection: English *n't. Language* 59: 502–513.

18

Romance Clitics, Verb Movement, and PRO

Richard S. Kayne

1991

18.1 Introduction

The discovery of correlations is particularly important for the work of a linguist: if two phenomena go hand in hand, they might stem from the same grammatical property. If we can prove that they do, we make one step forward in understanding the nature of the grammatical system. In *Romance Clitics, Verb Movement, and PRO*, Kayne discusses a well-known difference observed within Romance languages and a little-studied contrast that we see in Romance (as well as within English), both concerning embedded infinitival clauses. In so doing, he discovers the existence of a previously unknown correlation, and points to the property of the grammar responsible for it.

The well-known difference that this article tackles anew is that in some Romance languages pronominal clitics precede infinitival verbs, whereas in others they follow them, as we see in (1) and (2):

(1) *Lui parler*   serait   une erreur.   (French)
     him to-speak would.be an   error
     'To speak to him would be a mistake.'
(2) *Parlargli*   sarebbe   un errore.   (Italian)
     to-speak-him would.be an error
     'To speak to him would be a mistake.'

The little-studied difference is the contrast between infinitival clauses that allow the co-occurrence of a complementizer-like element and a null subject, and those

that do not. This contrast can be seen in English, where *whether* co-occurs with a null subject in the infinitival clause but *if* does not:

(3)   a.   He doesn't know *whether* to go to the movies.
      b. *He doesn't know *if* to go to the movies.

Finally, the previously unknown correlation is the following: the Romance languages that exhibit clitic-infinitive order (e.g., French, Occitan, Sardinian) have a complementizer that patterns like English *if* in that it cannot occur in an infinitival clause with a null subject (cf. (4)). In contrast, the ones that exhibit infinitive-clitic word order (e.g., Italian, Catalan, Spanish) have a complementizer that allows such co-occurrence, as in (5):

(4) *Marie ne sait pas *si* aller au cinéma (ou non). (French)
     Marie neg knows neg if to-go to-the cinema (or not)
(5)   Maria non sa *se* andare al cinema. (Italian)
     Maria neg knows if to-go to-the cinema

Having uncovered this correlation, the challenge is that of determining whether there is a single grammatical property that gives rise to both the order of the clitic pronoun with respect to the infinitival verb and the possibility of having the counterpart of *if* in an infinitival clause with a null subject.

Whereas previous studies treated the contrast in word order as resulting from the clitic adjoining to the right of the verb in certain languages and to its left in others, Kayne argues that clitics always adjoin to the left of a functional head. He offers a new analysis for the well-known contrast in word order, arguing that it arises from the infinitival verb moving to a different extent in the two groups of languages. More precisely, Kayne assumes (following Raposo 1987) that the infinitival suffix of French and Italian corresponds to a functional head (INFN) with nominal properties. In French, the verb raises and adjoins to Infn (but does not raise to T), and the clitic adjoins to the complex head V + Infn:

(6)   ...T ...Cl + [$_{INFN}$ V + Infn] ...[$_{VP}$ [$_V$ e]]... (French)

In Italian, the verb adjoins to Infn and then the complex head V + Infn further raises to a position higher than T (it adjoins to T''); the clitic adjoins to T:

(7)   ...V + Infn...Cl + T...[$_{INFN}$ e][$_{VP}$ [$_V$ e]]...(Italian)

   If possible, replace '…' by '. . .'.>
This difference in verb movement yields the different relative order of the infinitival verb and the clitic pronouns: clitic-infinitive in French, infinitive-clitic in Italian.

How does this relate to the possibility of having the counterpart of *if* in an embedded infinitival clause? For English, Kayne takes *whether* to be a *wh*-phrase (following Katz and Postal 1964 and Larson 1985) and *if* a lexical head (cf. Emonds 1985). He then argues that this difference in their status underlies their different behavior in infinitival clauses. Since *if* is a lexical head, it governs the subject; assuming Chomsky's (1981) proposal that the null subject is PRO, an element that cannot be governed, it follows that *if* cannot co-occur with PRO. In contrast, *whether* is not a lexical head, but a phrase (in the specifier of CP); hence it does not govern the subject, and can co-occur with PRO. For Romance, Kayne shows that French *si* and Italian *se* share a number of properties with *if*, concluding that they are also lexical heads. It is then straightforward to assume that French *si* (like its counterparts in Occitan and Sardinian) is not compatible with the null subject of an infinitival clause, PRO, for the same reason given for English *if*, namely because it governs PRO. But it is surprising that Italian *se* (like its counterpart in Catalan and Spanish) is compatible with PRO. Kayne relates this difference to the different position of the infinitival verb in Italian as opposed to French. He argues that in languages with infinitive-clitic order (like Italian, Catalan, and Spanish), where the infinitival verb moves to a position higher than T, it blocks government of PRO by the complementizer. This is why in these languages, and only in these languages, the lexical head that is the counterpart of *if* can co-occur with a null subject.

Given the framework adopted in Kayne's article, the question arises of whether the infinitival verb itself governs PRO, when it adjoins to T'. Kayne provides an answer to this question that is positive and yet compatible with the presence of PRO, by appealing to the version of the Binding Theory developed in Chomsky (1986). In a theoretical framework that dispenses with the notion of government, like that of the Minimalist program, the question becomes how to express the idea that the higher structural position of the infinitival verb in some languages (e.g., Italian, Catalan, and Spanish) makes possible the presence of a null subject, while the lower position of the infinitival verb in others (e.g., French, Occitan, and Sardinian) does not. Regardless of framework, once a correlation has been discovered, our model of grammar needs to capture it, as elegantly as possible (cf. Kayne to appear).

18.2 From "ROMANCE CLITICS, VERB MOVEMENT, AND *PRO*"

2 PRO

2.1 English

There is in English a contrast between *whether* and *if* with respect to control:

(51)   He doesn't know whether to go to the movies.
(52)   *He doesn't know if to go to the movies.

Both *whether* and *if* are of course possible in the finite counterparts to these:

(53)   He doesn't know whether he should go to the movies.
(54)   He doesn't know if he should go to the movies.

The grammaticality of (51) can be straightforwardly assimilated to that of other *wh*-infinitive constructions such as (55) if, following Katz and Postal (1964, 96) and Larson (1985, 238), we take *whether* to be a *wh*-phrase:

    [. . .]

Conversely, the *if* of (54) almost certainly bears some relation to that of conditionals:

(61)   If you had not left, he would have been a lot happier.

Since this *if* does not alternate with *wh*-phrases, it is not surprising that *whether*, a *wh*-phrase, is not found:

(62)   *Whether you had not left, he would have been happier.

The conclusion I would like to draw from all this is that the primary difference between *whether* and *if* is that the former is a *wh*-phrase and the latter is not, and furthermore, that this difference in syntactic status is responsible for the contrast in behavior with respect to control seen in (51) versus (52).

As for the exact status of *if*, I will, in agreement with Emonds (1985, 287), take it to be a complementizer, and more precisely, to be an X^0 element. Emonds takes *if*, like other complementizers, to be of category P^0, as opposed to Chomsky's (1986a) C^0. I will call it C^0, while keeping in mind that P^0 might perhaps be compatible with what follows, in particular a non-Case-assigning P^0.

The basic proposal will be that control is incompatible with the presence of a lexical complementizer, and hence incompatible with *if*. Control is, on the other hand, compatible with *whether* since *whether* is not a lexical complementizer, but a *wh*-phrase (that is, it is not a C^0, but a phrase in the Specifier position of CP); nor is there any element in (51) that is a C^0. As for the exact reason why a lexical complementizer inhibits control in (52), let us adopt as a first approximation the theory of control developed in Chomsky ([1981]) (*LGB*), which takes the controlled subject NP to be the element PRO, having the features [+anaphoric] and [+pronominal]. Principles A and B of the binding theory combine to yield the so-called PRO theorem, which states that PRO must be ungoverned. Assume now that a lexically filled C^0 counts as a governor for the PRO in subject position, but that a nonlexical C^0 position does not. [. . .] Then the contrast between (51) and (52) follows from the *LGB* theory of control, via the PRO theorem.

2.2 French

French is substantially like English with respect to the phenomena of the previous section, once we abstract away from a major difference, namely, that

French lacks any counterpart to English *whether*. Corresponding to (53) and (54) French has only (63):

(63)   Marie ne sait pas si elle devrait aller au cinéma.

This alone is not sufficient to tell us whether French *si* corresponds more to English *if* or to English *whether*. However, if we run through the various distinguishing properties noted above, we see that *si* corresponds strongly to *if* and not at all to *whether*. First, the control counterpart of (63) is ungrammatical, like *if* in (52):

(64)   *Marie ne sait pas si aller au cinéma (ou non).

    [. . .]
Finally, conditionals in French do use *si* as English uses *if* (see (61)):

(67)   Si vous n'étiez pas parti, il aurait été plus heureux.

    [. . .]
    That *si* is a complementizer (see Huot (1974, 47)) and more specifically a C^0 makes it possible to account for (64) in exactly the same way as proposed earlier for English (52), that is, in terms of the PRO theorem and government of PRO by *si*.
    [. . .]

2.3   Italian

There is no single word for *either* (or *neither*) in Italian, and, as we would then expect, no word corresponding to *whether*. There is, on the other hand, a word *se*, which resembles French *si*, and which, like French *si*, has much in common with English *if*. Like *si* and *if*, Italian *se* occurs both in embedded interrogative contexts and in conditionals:

(75)   Gianni non sa   se dovrebbe andare al   cinema.
        Gianni NEG knows if he-should to-go   to-the movies
(76)   Se Gianni avesse fatto questo, Paolo . . .
        if Gianni had   done this   Paolo

[. . .] There thus appears to be every reason to take Italian *se* to be an instance of C^0.
    Support for this position comes from dialects like those described by Ganzoni (1983, 160) and Poletto (1990) in which subordinating conjunctions, as well as embedded *wh*-phrases, are invariably followed by the complementizer *chalche* 'that', with one exception: *schalse* 'if'. I interpret this to reflect the C^0 status of *schalse* versus the non-C^0 status of subordinating conjunctions and *wh*-phrases.
    Additional support for this hypothesis comes from clitic-climbing considerations. As noted in Kayne ([1989], 245), *se* blocks clitic climbing into a matrix sentence more strongly than *wh*-phrases do in general.
    [. . .]

Despite these many ways in which Italian *se* seems definitely to be a C^0 like French *si* and English *if*, there is one major unexpected disparity in behavior. Unlike *si* and *if*, Italian *se* is compatible with control:

(80)   Gianni non sa   se andare al   cinema.
       Gianni NEG knows if to-go to-the movies

In light of the first three paragraphs of this section, it would be totally implausible to try to interpret *se* as an Italian equivalent of *whether*. But if so, the contrast between (80) and its French counterpart (64), repeated here as (81), seems mysterious:

(81)   *Marie ne sait pas si aller au cinéma.

The analysis developed here so far would lead us to expect (80) to be ungrammatical, too –*se*, being a C^0, should govern PRO across IP and thereby induce a PRO theorem violation.

2.4 Romance

In the spirit of the comparative syntax work of the past ten or more years, we must ask whether this Italian-French difference is related to any other, in the hope that if a correlation is discovered, it will point the way toward a solution to the problem. In Kayne ([1989], 252) I suggested a correlation with the null subject parameter, but consideration of additional Romance languages seems to indicate that that was incorrect.

Though it is true that the null subject languages Catalan and Spanish appear to pattern with Italian as far as (80) versus (81) is concerned, the null subject languages Occitan and Sardinian pattern instead with French; that is, they do not allow control with their counterpart to *if* (*se* in Occitan (82), *si* in Sardinian (83)):

(82) *Sabi   pas se anar al   cinema.
       I-know not if to-go to-the movies
(83) *No'isco   si andare.
       NEG I-know if to-go

I conclude that being a null subject language is not a sufficient condition for permitting control with *if* and therefore that there must be some other factor at issue in the Italian/French contrast between (80) and (81).

The question, then, is to figure out what Italian, Catalan, and Spanish have in common that sets them off from French, Occitan, and Sardinian. I propose that the key property is that of infinitive-clitic order, which holds for the first three, but not for the last three, which show clitic-infinitive order.

   [. . .]

2.5   Infinitive Adjunction interferes with C^0-Government

In section 1 [in the full article] I took infinitive–clitic languages to differ from clitic–infinitive languages in having their infinitive left-adjoin to the I′ just below the C projection, the clitic itself being left-adjoined to the corresponding I (which I took to be T):

(84)   . . . V_{inf} + [$_{I'}$ . . . Cl +] . . .

The order clitic–infinitive in the other class of languages involved no such adjunction to I′, but rather movement of the infinitive into some I position and adjunction of the clitic either to that I position or to some higher one.

Recall now that I have suggested interpreting the ungrammaticality of control with *if/si/se* in French, Sardinian, Occitan, Gardenese (and English) as due to the government of PRO by the lexical C^0 and to the consequent violation of the PRO theorem:

(85)   . . . *if* . . . [$_{IP}$ PRO . . .

In clitic–infinitive languages, the infinitive ends up in an I position below PRO. In the absence of *if*, control is perfectly possible and the standard conclusion is that the infinitive there does not govern PRO. In the presence of *if*, the infinitive moves to the same I and the same conclusion holds. In other words, in (86) PRO is governed by *si/se* and is not governed by the infinitive (independently of whether any clitic is present):

(86)   . . . *si* . . . [$_{IP}$ PRO . . . V_{inf} + I . . .

By virtue of being governed by C^0, PRO in (86) violates the PRO theorem, that is, the conjunction of Principles A and B of the *LGB* binding theory.

Fleshing out (84) to show PRO and to show where the lexical C^0 is (when it is present), we have (87):

(87)   . . . *se* . . . [$_{IP}$ PRO . . . [$_{I'}$ V_{inf} + [$_{I'}$ . . . (Cl +)I . . .

I have taken the infinitive to left-adjoin to I′ in these languages, whether or not a clitic is present. Put another way, in the infinitive–clitic languages like Italian, the infinitive will in the general case move into a position that is hierarchically closer to PRO than the position it moves into in the clitic–infinitive languages. I would like to propose, now, that in so doing the infinitive in (87) blocks off government of PRO by C^0 and thereby eliminates the potential PRO theorem violation induced by that C^0.

[. . .]

Summing up, the idea that I am pursuing is that a lexical C^0 will be expected to induce a PRO theorem violation when PRO is the subject of the IP sister of that C^0. However, the government relation between C^0 and PRO that would be the cause of such a violation can be blocked by the presence of a closer governor.

In languages that have the order infinitive-clitic, and only in those, the infinitive itself can be the required closer governor, having moved into an appropriate position by adjoining to I′.

It should be noted that this account of the correlation between control with a lexical C^0 and infinitive-clitic order, insofar as it depends crucially on the sensitivity of PRO to government by that C^0, supports the very postulation of a category PRO, that is, of a type of empty NP with a particular position in the syntactic structure and with the features [+ anaphoric] and [+ pronominal] given it by the *LGB* binding theory.

In effect, we can think of the process of looking at a set of Romance languages, moving from one with clitic-infinitive order to the opposite type and back, as a kind of experiment in which we hold the basic structure of a language – Romance – (relatively) constant, while varying the position of the infinitive. What we learn is that as we so vary its position, the grammaticality of control sentences with *si/se* varies in step. If my theoretical proposal is correct, then we can interpret this covariance as reflecting the sensitivity of PRO to the position of the infinitive, that is, to the presence versus absence of a government relation with *si/se*.

2.6   Binding Theory and PRO

The question arises why the infinitive adjoined to I′ in the Italian-type languages does not itself induce a PRO theorem violation. [. . .]

Let us adopt the paradoxical position that infinitive adjunction in Italian does create a configuration in which the infinitive comes to govern PRO, that the PRO theorem continues to play an important role in UG, and yet that there is no PRO theorem violation here.

Consider the revision of binding theory suggested by Chomsky in *Knowledge of Language* (*KL*) [Chomsky 1986b] (pp. 170 ff.) in which a slight discrepancy is introduced (in terms of BT-compatibility) between the governing category for an anaphor and the governing category for a pronoun. This discrepancy concerns in particular anaphors and pronouns in subject position. It is relevant when the subject position in question is governed by a lexical category that is found inside (rather than outside, as is more usual) the X^{max} of which the anaphor or pronoun is the subject. In that case the governing category of the pronoun would be X^{max}, the smallest category containing both the governor and a subject position.

However, in the case of an anaphor in such an internally governed subject position, the governing category is not X^{max}, but rather the next category up containing a subject position, the reason being that although X^{max} contains the governor of the anaphor, its subject position is not a potential binder for the anaphor (informally put, it would be unreasonable to require an anaphor to be bound within a category containing no position that could contain a potential binder – comparable unreasonableness is not an issue in the case of pronouns).

It follows from the simplest interpretation of this revision that the PRO theorem should no longer hold in full generality, although it will continue to hold over a restricted (but still wide) range. This is so since the PRO theorem follows from the strict parallelism between Principles A and B of the binding theory. To the extent that strict parallelism fails to hold over some range of environments, the PRO

theorem will fail to hold for that range. More specifically, it will fail to hold for any subject PRO governed by a lexical category found within the category of which PRO is the subject, since in such a case the governing category for PRO qua anaphor will not be identical to the governing category for PRO qua pronoun.

On the other hand, the PRO theorem will continue to hold, as in *LGB*, for all object PROs as well as for all subject PROs governed by an element outside the category of which PRO is the subject.

In particular, when a lexical complementizer governs PRO, a PRO theorem violation continues to hold, since the complementizer is outside the IP of which PRO is in subject/Spec position. This is what excludes . . *if PRO to go to the movies* and the comparable examples discussed above for French, Occitan, Sardinian, and Gardenese (see (81)–(83)).

The difference between the *KL* binding theory and the *LGB* binding theory becomes important when we turn to the languages like Italian in which the infinitive left-adjoins to I′:

(88)   . . . *se* . . . $[_{IP}$ PRO . . . $[_{I'} V_{inf} + [_{I'} . . .$

By hypothesis, *se* no longer governs PRO in this configuration, but V_{inf} does. In the *LGB* theory, this would have led to a PRO theorem violation. In the *KL* theory, on the other hand, that is not the case, as follows: The governing category for PRO qua pronoun is IP, since that is the smallest category that contains a subject position and contains the governor of the pronoun.

This is not yet different from the *LGB* state of affairs. The crucial difference lies in how the two theories determine the governing category of PRO qua anaphor in (88). For the *LGB* theory, it is again IP, the same as for PRO qua pronoun, leading to a typical PRO theorem violation. For the *KL* theory, that is not the case. IP in (88) does contain the governor, but it does not contain a suitably accessible potential binder and so does not qualify as governing category for PRO qua anaphor. Rather, the governing category for PRO qua anaphor will be the next category up containing a subject position, in effect, the next IP up (not shown in (88)). Since this governing category is distinct from that assigned to PRO qua pronoun, there is no violation of the PRO theorem sort, as desired.

Thus, the *KL* binding theory is capable of distinguishing the Italian construction represented by (88) from the corresponding French and English one.

In assigning to PRO qua anaphor the next IP up as governing category, the binding theory adopted here excludes the possibility that the antecedent of PRO in (88) could be taken to be a subject NP two IPs up. This accounts correctly for the fact that in (89) the antecedent of PRO must be *Gianni* and cannot be *Maria*:

(89)   Maria pensa che Gianni non sappia se andare al   cinema.
        Maria thinks that Gianni NEG knows if to-go   to-the movies

This pattern is of course widespread for control infinitivals that are verb complements, as, for example, in (90), in which again the antecedent of PRO must be the subject of 'decide' and cannot be that of 'thinks'.

(90)   Maria pensa che Gianni abbia deciso   di andare.
       Maria thinks that Gianni has   decided DI to–go

This resolves a paradox noted by Lasnik (1989), namely, that the *LGB* binding theory accounts for the distribution of PRO (by excluding it from governed positions), but at the same time fails to assign it a governing category and so makes no claim at all about the location of its antecedent. My extension of the *KL* binding theory to PRO retains the distributional account (by excluding PRO from all governed positions except those subject positions governed by an element inside the XP of which PRO is the subject) and at the same time does assign PRO a governing category and so does make some claim about the location of the antecedent.

This approach to PRO, in having binding theory determine a governing category for PRO and hence delimit the possible positions for the antecedent of PRO, is significantly similar to that of Manzini (1983), but has the advantage that there is no need to add to binding theory any notion of domain-governing category. From our perspective, the same effect is achieved in the Italian infinitive cases by the basic characterization of Principle A as picking out as governing category the smallest category containing a governor and an accessible subject. Since where PRO is the subject of an infinitive, that subject position does not count as accessible, Principle A will look for the next largest category containing one, which, in the case of the infinitive as complement of V, will straightforwardly be the next IP up (and there will be no PRO theorem violation, as discussed).

The approach developed here has the further advantage of allowing an account of the Italian-French contrast with respect to control in the presence of *se/si*, which depends on the *KL* binding theory and in particular on the analysis of PRO as simultaneously anaphoric and pronominal, whereas Manzini took PRO here to be a pure anaphor.

2.7   Levels

My account of the Italian-French contrast with respect to control in the presence of *se/si* 'if' depended in part on postulating a rule of leftward infinitive adjunction to I′ that applies in Italian, but not in French. The left-adjoined infinitive governs PRO in Italian, with the consequences noted in the previous two sections. The absence of comparable infinitive movement in French means that in French the infinitive does not govern PRO – this is precisely what allows a lexical C^0 in French to induce a PRO theorem violation. In the absence of a lexical C^0, as in (91), French PRO is therefore ungoverned:

(91)   Jean veut aller au   cinéma.
       Jean wants to–go to–the movies

This is of course expected within the *LGB* perspective and is perfectly compatible with what I have said so far. This is so, in the sense that I have argued that PRO can be governed under certain very specific conditions, but have in effect left open the possibility that it can also be ungoverned.

A problem arises, however, with respect to the paradox adduced by Lasnik that was mentioned earlier. I argued that his paradox is resolved for Italian by the fact that PRO there is governed by the preposed infinitive, hence gets a governing category, so that binding theory actually does provide an indication of where the antecedent of PRO must be. But if PRO remains ungoverned in French, Lasnik's paradox reappears there. I would like to propose, then, that French is to Italian with respect to leftward infinitive adjunction to I′ as Chinese is to Italian with respect to *wh*-movement, in other words, that French actually does have such infinitive movement, but only at the level of LF.

This leads to the following proposal:

(92)　All controlled PROs are governed at some level of representation.

(92) holds even though the PRO theorem is largely true. This is so in the sense that the PRO theorem continues to hold for all PROs other than those that are in subject position and governed by an internal governor. On the other hand, if I am correct in putting forth (92), then any controlled PRO that is ungoverned at all levels of representation is equally excluded.

I take the reason for the existence of (92) to be that it is via government that PRO qua anaphor receives a governing category. Assuming further that an antecedent for PRO must be within PRO's governing category (that is, that an ungoverned PRO would not be able to be associated with any antecedent at all), (92) follows. In effect, I have reached the conclusion that PRO is less exotic than it was in the *LGB* framework, since PRO is now like other empty categories in being licensed in part via government; at the same time, the present theory maintains the specificity of PRO, and in particular its exclusion from most governed positions (see footnote 77 [in the full article]).

(92) is stated in such a way as to allow for the possibility that there exist instances of ungoverned noncontrolled PRO, that is, instances of ungoverned PRO_{arb}. However, PRO_{arb} seems to exist in Italian with infinitives, as, for example, in (93) (also see Manzini (1979)):

(93)　Tu　conosci il　modo migliore per comportarsi　　a tavola.
　　　you know　the way　best　　for to-behave-self$_{arb}$ at table

But by my analysis, the infinitive in (93) has moved into a position from which it governs PRO (notice the clitic in (93) following the infinitive and serving as a visible indication of that general movement). Therefore, the PRO_{arb} of (93) cannot be ungoverned, which suggests in turn that (92) should be taken to extend to all instances of PRO – in other words, that PRO_{arb} is really a subcase of controlled PRO, as proposed by Epstein (1984), who argues that many instances of PRO_{arb} should be taken to be controlled by a hidden dative (also see Higginbotham (1989, 324)). The most recalcitrant cases are those of (94) and (95):

(94)　?John knows how to get oneself elected.
(95)　a.　John knows the best way to get oneself elected.
　　　　b.　John knows the best way of getting oneself elected.

The fact that these seem best when embedded within a larger NP (as suggested by Petrovitz (1990)) might indicate that these instances of PRO_{arb} must, in the spirit of Lebeaux (1984) and Authier (1989), be bound by some null operator sitting in a position provided by the NP.

Returning to the idea that controlled PRO is governed even in French (at LF), let us reconsider two kinds of examples:

(96) *Jean ne sait   pas si aller au   cinéma.
     Jean NEG knows not if to-go to-the movies
(97) Jean veut aller au   cinéma.
     Jean wants to-go to-the movies

My idea has been that (96) is excluded because the lexical C^0 *si* governs PRO and induces a PRO theorem violation. Yet I am now proposing that in (97) PRO is governed by the infinitive at LF. There is no contradiction, since in (97) government will be of the internal type (that is, the governor is internal to the IP of which PRO is the subject), whereas in (96) it is of the external type (*si* is external to that IP), and in my analysis the (revised) PRO theorem holds for subject PRO only over the domain of external government configurations.

It is important, however, to ensure that LF movement of the infinitive does not have the undesirable consequence of making (96) legitimate, the point being that subsequent to such LF movement PRO in (96) will be governed by the infinitive and will no longer be governed by *si*. I conclude that a PRO theorem–type violation at S-Structure, as in (96), cannot be neutralized at LF. Considering more closely the exact nature of the violation in (96), note that by virtue of being governed by *si*, PRO qua anaphor receives as governing category the matrix IP, which is perfectly reasonable – if (96) were grammatical, that is where we would expect the antecedent to be. The problem with (96) is really that PRO qua pronoun also receives the matrix IP as governing category, yielding the familiar contradiction. If LF movement of the infinitive were able to neutralize such a violation, it would have to be by virtue of changing what counts as the governing category of PRO qua pronoun. Since the violation remains, I conclude that a governing category assigned by Principle B to a given pronominal element must be taken to stick to it.

Put more perspicuously, a given indexing must respect Principle B at all levels. Thus, if PRO in (96) is coindexed with *Jean*, a violation will ensue since Principle B will not have been respected at S-Structure. On the other hand, if I am correct in thinking that PRO cannot be assigned an antecedent without having a governing category, then in (97) PRO has an antecedent only at LF. In other words, Principle A must be met at some level of representation, but does not need to be met at all levels. This asymmetry between Principle A and Principle B recalls the conclusion reached in Belletti and Rizzi (1988, 318).

If we now ask why there should exist such an asymmetry, the following answer suggests itself: Binding principles are properly thought of as applying to a set of levels of representation associated with a given sentence. Principle A has intrinsically existential character (for a given anaphor, there must exist an antecedent

within the appropriate syntactic domain). Interpreting this existential character consistently yields: For a given anaphor, there must exist some antecedent at some level (that is, somewhere in the set) within the appropriate syntactic domain. Principle B, on the other hand, has intrinsically universal character (a given pronoun must be free from all antecedents within the appropriate syntactic domain). Interpreting this consistently yields: A given pronoun must be free from all antecedents at all levels (that is, everywhere in the set) within the appropriate syntactic domain.

18.3 Questions pertaining to Kayne (1991)

1   In establishing his generalization that control with *if* is possible in Romance only in infinitive-clitic languages, Kayne takes into account a dozen or so Romance languages. Is that a satisfactory number? (Discuss the reasons for your answer.) What is the significance of the fact that many Romance languages spoken in the past are now extinct?

2   To what extent is it essential that Kayne's proposals be tested by examining language families other than Romance? Discuss some of the difficulties that might arise in attempting to do so.

3   Some of the languages taken into account in this paper, such as English, French, and Italian, have large numbers of speakers and have quite a number of syntacticians working on them who are native speakers, respectively, of English, French, or Italian. Other languages taken into account, such as Piedmontese, Milanese, and Paduan have many fewer native speakers and few syntacticians are native speakers of them. Discuss the significance of this difference between the two groups of languages, bringing in the notion of replicability. In addition, discuss the limiting case of languages for which no native-speaking syntactician is available.

4   In discussing the contrast between *We don't know whether to leave now or not* and **We don't know if to leave now or not*, Kayne does not attempt to take into account facts such as *We don't know in what sense we should be satisfied with his answer vs. *We don't know in what sense to be satisfied with his answer*. How likely is it that this second contrast is closely related to the first? Give your reasons and discuss the possible implications.

5   In Romance languages, object clitics typically precede, rather than follow, finite verbs, and this is typically true even when the subject is the impersonal *si/se* studied in detail in Cinque (1988). For example, Italian has, with object clitic *lo*, *Lo si vede* ('one sees it/him') vs. **Si vedelo*. Yet Friulian allows the counterpart of the latter example; more specifically it allows *lo* and other object clitics to follow the finite verb only when the subject is impersonal *si*.

In allowing object clitics an atypical position in the presence of a specific other element, here *si*, Friulian bears a certain resemblance to Italian itself, which, atypically for it, allows object clitics to precede infinitives only in negative imperatives. To what extent, and in what way, can this Italian phenomenon, discussed in Kayne (1992), be linked to the Friulian one?

6 In taking Romance object clitics to invariably left-adjoin, rather than right-adjoin, to a head, this paper is a partial precursor to *The Antisymmetry of Syntax*. One type of sentence mentioned in this paper as supporting the availability of right-adjunction in other cases involves subject postposition in French, as had been discussed in Kayne and Pollock (1978). Pinpoint the key ingredients of Kayne and Pollock (2001) that now make those French postverbal subjects compatible with antisymmetry.

7 In taking object clitics to adjoin to heads, Kayne in this paper was assuming that object clitics are themselves heads. In subsequent work, though, Kayne (2008, note 11) suggests, by extension from English expletive *there*, that at least some Romance object clitics might be remnants, i.e., phrasal constituents only one piece of which is visible. Discuss some of the implications of the possibility that all object clitics are remnants. (Extra credit: Extend the discussion to what are called agreement morphemes, that is, to the implications of the possibility that all agreement morphemes are remnants.)

8 The most robust exception to the generalization that Romance object clitics precede finite verbs is found in Borgomanerese, as studied recently by Tortora (2002). Discuss how exactly Cinque (1999) makes it easier to fit Borgomanerese into the general picture of Romance object clitics.

9 Object clitics raise out of past participial phrases in Romance more consistently than out of infinitival phrases. Put another way, to take French, *Jean les a vus* ('Jean them has seen') contrasts with **Jean les veut voir* ('Jean them wants to-see'). Why might this contrast between past participles and infinitives hold?

10 This paper had V-adjunction to T-bar play a crucial role in control sentences with *if*. To what extent could such V-adjunction be mimicked within bare phrase structure? To what extent is the paper's analysis of control compatible with Chomsky and Lasnik (1993)?

11 The approach to control that Kayne adopted in this 1991 paper primarily followed Chomsky (1981, 1986). Subsequently, O'Neil (1995, 1997), Hornstein (1999, 2001), and Kayne (2002) suggested approaches to control based on movement. The O'Neil/Hornstein approach involved, in familiar sentences like *John tried to solve the problem*, movement of *John* from one theta-position to another, higher one. (Kayne's did not require allowing one DP to bear two theta-roles; rather *John*, in such sentences, originates as a double of PRO.) To what extent do the problems pointed out by Modesto (2010) for the O'Neil/Hornstein approach carry over to Kayne's approach? (See also the articles in *Syntax 9.2*, 2006, a special issue of the journal on Raising and Control.)

12 Either type of movement approach mentioned in the previous question will in principle lead to new perspectives on the control-with-*if* question illustrated in *We don't know whether/*if to leave now or not*. In particular, the restriction with *if* could now be thought of in terms of a restriction against extraction from preverbal subject position. State at least one major challenge for that thought.

13 (Extra credit) As (the answer to) the preceding question shows, English by itself does not provide much support here for the movement approach to

control. Yet Romance might, insofar as the French vs. Italian contrast (with French acting like English, vs. Italian allowing control with *if*) recalls one discussed in the context of pro-drop, namely that Italian "freely" allows postverbal subjects in a way that French does not.

More specifically, thinking of Rizzi (1982, Ch. 4), one might take Italian control with *if*, as in *Gianni non sa se partire* ('Gianni neg knows if to-leave') to involve movement from a postverbal, rather than preverbal, subject position. That option would not be available to the French counterpart of such a sentence. A problem that immediately arises is that French does allow control with other embedded interrogatives, e.g., *Jean ne sais pas quoi faire* ('Jean neg knows not what to-do').

Discuss a possible solution to this problem that would invoke the limited way in which French does admit some postverbal subjects, as discussed in Kayne and Pollock (1978, 2001).

14 Toward the end of section 2.5, Kayne says ". . . we can think of the process of looking at a set of Romance languages, moving from one with clitic-infinitive order to the opposite type and back, as a kind of experiment in which we hold the basic structure of a language – Romance – (relatively) constant, while varying the position of the infinitive. What we learn is that as we so vary its position, the grammaticality of control sentences with *si/se* varies in step. . . . we can interpret this covariance as reflecting the sensitivity of PRO to the position of the infinitive. . . ." How might this have to be restated if a movement approach to control were to turn out to be correct?

15 In a footnote associated with the text of the preceding question, Kayne says "it is advantageous to work with a set of closely related languages, much as in any experiment one tries to keep the number of variables as low as possible." How might this idea transpose to English?

References

Authier, J.-M. P. 1989. Arbitrary null objects and unselective binding. In O. Jaeggli and K. J. Safir (eds), *The Null Subject Parameter*. Dordrecht: Kluwer.

Belletti, A., and L. Rizzi 1988. Psych-Verbs and 0-Theory. *Natural Language and Linguistic Theory* 6: 291–352.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1986a. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 1986b. *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N., and H. Lasnik. 1993. The theory of principles and parameters. In J. Jacobs, A. von Stechow, W. Sternefeld, and T. Vennemann (eds), *Syntax: An International Handbook of Contemporary Research*, Berlin: de Gruyter, pp. 506–569. [Reprinted in Chomsky (1995, Ch. 1).]

Cinque, G. 1988. On *si* constructions and the theory of *arb*. *Linguistic Inquiry* 19: 521–581.

Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York and Oxford: Oxford University Press.

Emonds, J. E. 1985. *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.

Epstein, S. D. 1984. Quantifier-pro and the LF Representation of PRO_{arb}. *Linguistic Inquiry* 15: 499–505.

Ganzoni, G. P. 1983. *Grammatica ladina: Grammatica sistematica dal rumantsch d'Engiadina Bassa*. Samedan: Lia Rumantscha.

Higginbotham, J. 1989. Reference and control. *Rivista de Linguistica* 1: 301–326.

Hornstein, N. 1999. Movement and control. *Linguistic Inquiry* 30: 69–96.

Hornstein, N. 2001. *Move! A Minimalist Theory of Construal*. Oxford: Blackwell.

Huot, H. 1974. Les relatives parenthétiques. In C. Rohrer and N. Ruwet (eds), *Actes du Colloque Franco-allemand de Grammaire Transformationnelle* I, Tübingen: Niemeyer., pp. 31–62.

Katz, J. J., and P. M. Postal. 1964. *An Integrated Theory of Linguistic Descriptions*. Cambridge, MA: MIT Press.

Kayne, R. S. 1989. Null subjects and clitic climbing. In O. Jaeggli and K. J. Safir (eds), *The Null Subject Parameter*, Dordrecht: Kluwer, pp. 239–261.

Kayne, R. S. 1992. Italian negative infinitival imperatives and clitic climbing. In L. Tasmowski and A. Zribi-Hertz (eds), *De la musique à la linguistique. Hommages à Nicolas Ruwet*, Ghent: Communication Cognition, pp. 300–312. [Reprinted in Kayne (2000, Ch. 6).]

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2002. Pronouns and their antecedents. In S. D. Epstein and T. D. Seely (eds), *Derivation and Explanation in the Minimalist Program*, Oxford: Blackwell, pp. 133–166. [Reprinted in Kayne (2005, Ch. 6).]

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2008. Expletives, datives, and the tension between morphology and syntax. In T. Biberauer (ed.), *The Limits of Syntactic Variation*, Amsterdam and Philadelphia: John Benjamins, pp. 175–217. [Reprinted in Kayne (2010, Ch. 6).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Kayne, R. S. to appear. Comparative syntax. *Lingua*.

Kayne, R. S., and J.-Y. Pollock. 1978. Stylistic inversion, successive cyclicity and move NP in French. *Linguistic Inquiry* 9: 595–621.

Kayne, R. S., and J.-Y. Pollock. 2001. New thoughts on stylistic inversion. In A. C. Hulk and J.-Y. Pollock, *Inversion in Romance*, New York and Oxford: Oxford University Press, pp. 107–162.

Larson, R. K. 1985. On the syntax of disjunction scope. *Natural Language and Linguistic Theory* 3: 217–264.

Lasnik, H. 1989. Two notes on control and binding. Paper presented at the MIT Workshop on control. [See also: Lasnik, H., 1992, Two notes on control and binding. In R. Larson, S. Iatridou, U. Lahiri, and J. Higginbotham (eds), *Control and Grammar*, Dordrecht: Kluwer, pp. 235–251.]

Lebeaux, D. 1984. Anaphoric binding and the definition of PRO. In C. Jones and P. Sells (eds), *Proceedings of NELS 14*, Amherst: GLSA, University of Massachusetts.

Manzini, M. R. 1979. Una teoria del controllo. *Rivista di Grammatica Generativa* 4: 139–163.

Manzini, M. R. 1983 On control and control theory. *Linguistic Inquiry* 14: 421–446.

Modesto, M. 2010. What Brazilian Portuguese says about control: Remarks on Boeckx & Hornstein. *Syntax* 13: 78–96.

O'Neil, J. 1995. Out of control. In J. N. Beckman (ed.), *Proceedings of NELS 25*, Amherst: GLSA, University of Massachusetts, vol. 1, pp. 361–371.

O'Neil, J. 1997. *Means of control: Deriving the properties of PRO in the Minimalist Program*. Doctoral dissertation, Harvard University.

Petrovitz, W. 1990. *Argument opacity*. Doctoral dissertation, CUNY.

Poletto, C. 1990. L'inversione soggetto clitico/verbo nei dialetti veneti. Ms., Università degli Studi di Padova e Venezia/Université de Genève.

Raposo, E. 1987. Case theory and Infl-to-Comp: The inflected infinitive in European Portuguese. *Linguistic Inquiry* 18: 85–109.

Rizzi, L. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Tortora, C. 2002. Romance enclisis, prepositions, and aspect. *Natural Language and Linguistic Theory* 20: 725–757.

19

The Position of Subjects

Hilda Koopman and Dominique Sportiche

1991

19.1 Introduction

Subjects in SVO languages are generally thought of as being in the specifier of the head of the clause, identified first as INFL and, later, as T. One question that arises is whether subjects originate in this position or move there. For passives, and later for unaccusatives (Perlmutter 1978; Burzio 1986), it was argued that the subject originates as an object of the verb and raises to a preverbal subject position. For subjects that are not underlying objects, the most common answer used to be that they are base-generated in the specifier of INFL. But in the 1980s, researchers began to question this assumption and provide evidence that subjects actually originate in a lower structural position and raise to the specifier of INFL, [Spec,IP] (see Contreras 1987; Kitagawa 1986; Koopman and Sportiche 1991; Kuroda 1988; Rosen 1990; Speas 1986; Wible 1990; Woolford 1991; Zagona 1982). In the paper included in this volume (originally presented at GLOW 1985), Koopman and Sportiche argue the selection relation between a predicate and its arguments is always local, and therefore subjects always start out within the maximal projection of the V. Underlying subjects, they argue, originate in the specifier of a projection they name V^{max}, which immediately dominates the VP, across languages.

Koopman and Sportiche present a great deal of evidence for their conclusion; here we will mention only one piece, which has to do with scope ambiguities. The sentence in (1) has two possible readings, given in (2):

(1) A griffin seems to be lurking on the 25th level.
(2) a. It seems that there is a griffin lurking on the 25th level.
     b. There is a griffin such that it seems to be lurking on the 25th level.
(3) seems [a griffin to be lurking on the 25th level]

May (1977, 1985) argued that these two readings are possible because *seem* is a raising verb; that is, the underlying structure of the sentence is the one illustrated in (3), where the subject is in the embedded clause. The ambiguity arises because the subject can be interpreted either in its base position (reading 2a) or in its surface position (reading 2b). Koopman and Sportiche point out that, if we replace *seem* with a modal element like *might*, as in (4), we get a similar ambiguity:

(4)   A griffin might be lurking on the 25th level.
(5)   a.   It might be that there is a griffin lurking on the 25th level.
      b.   There is a griffin and it might be lurking on the 25th level.

This ambiguity suggests that in sentence (4), as well as in the case of (1), the subject can be interpreted in one of two positions: one structurally lower than *might*, and one structurally higher. For Koopman and Sportiche, the former is the base position of the subject, the specifier of V*max*, while the latter is the position to which the subject raises, the specifier of INFL.

   In the second part of the paper (not included in this volume), Koopman and Sportiche discuss the implications of their proposal that external subjects receive their theta-role in V*max*. In view of the fact that [Spec,IP] is the position in which a deep object surfaces in a passive structure, that is, when it lacks objective case, Koopman and Sportiche analyze it as the position in which nominative Case is assigned. Hence, they argue that the reason why the thematic subject obligatorily raises to [Spec,IP] in certain cases is because of Case. This, however, raises questions with regard to (active) examples in which the subject does not raise to [Spec,IP], as for instance in VSO languages like Irish and Welsh. Koopman and Sportiche propose that the way INFL assigns Case is subject to parametric variation (cf. also Travis 1989, Ch. 16 of this volume). INFL can assign Case in one (or both) of two ways: by agreement, in a specifier-head configuration (see also Kayne 1989, Ch. 14 of this volume); or via government. In languages like English and French, nominative Case is assigned in a specifier-head agreement configuration; this is why we find SVO order, even when the verb raises to INFL, as it does in French (Pollock 1989, Ch. 15 of this volume). In languages like Irish and Welsh, on the other hand, INFL assigns Case only under government, to an intermediate subject position, and this is why we find VSO order and absence of agreement with (nonpronominal) subjects. (Note that these two strategies are similar to the options of Agree with and without an EPP feature on T in the later developments of the theory discussed in Chomsky 2000.) Finally, languages like Arabic are particularly interesting, in that they allow both options, VSO and SVO, with full subject agreement in the latter but not in the former word order.

   The general proposal put forth in this paper and in other work from that time, namely that subjects start out in a projection that immediately dominates VP (e.g., the specifier of a second VP shell, as proposed in Sportiche 1990), is often referred to as the "VP-Internal Subject Hypothesis." It has been widely accepted in the field, and further elaborated in very influential papers like Kratzer (1996). The particular

implementation of the idea has since been recast in different terms, invoking the notion of a νP, so-called "little VP" (Chomsky 1995: 315 and much other work since).

19.2   From "THE POSITION OF SUBJECTS"

Grammatical theories all use in one form or another the concept of canonical position of a phrase. If this notion is used in the syntax, when comparing the two sentences:
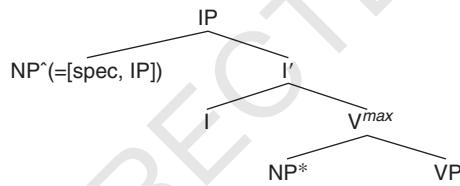
(1)   (a)   John will see Bill.
       (b)   Bill John will see.

we say that *Bill* occupies its canonical position in (1a) but not in (1b). Adopting the terminology of the Extended Standard Theory, we can think of the canonical position of a phrase as its D-structure position.

Since the concept of canonical position is available, it becomes legitimate to ask of each syntactic unit in a given sentence what its canonical position is, relative to the other units of the sentence.

The central question we address in this article is: what is the canonical position of subjects. Starting with English, we propose that the structure of an English clause is as in (2):

(2)

where NP* is the canonical or D-structure position of the subject, NP^ = (spec, I) is its S-structure position in simple declarative clauses, and V^{max} is a small clause whose predicate is VP.[2]

More generally, we propose that the constituent structure in (2) (linear order aside) can be generalized to hold of many (and perhaps all) languages, and that these languages fall in two classes. In Class 1 languages, such as English, French, Vata . . ., a subject generated in position NP* must move to position NP^. In a Class 2 language, such as possibly Italian, Welsh, Japanese, . . ., a subject generated in the NP* position does not have to raise to the position NP^. We will suggest that raising may also be obligatory in such a language, although not necessarily to NP^.

1   INFL as a raising category

Consider our proposal for English. It is equivalent to saying that INFL is a raising category. In this section, we argue that departure from this analysis of INFL needs arguments, arguments that are presently lacking.

1.1 Raising verbs

Start with a prototypical raising verb like *seem*. Why is it taken to be a raising verb?

1.1.1 Syntactic arguments

Here is a list of the classical arguments for analyzing the verb *seem* as a raising verb:

(1) *seem* imposes no selectional restrictions on its subject;
(2) *seem* can take expletive *it* as subject (*it seems that John sleeps all the time*) or non-expletive subjects (*John seems to sleep all the time*);
(3) *seem* allows as subject an NP licensed by the predicate of the clause embedded under it:
     weather *it* (*it seems to rain*)
     idiom chunks (*the cat seems to be out of the bag*)
     existential *there* (*there seems to be a griffin on the 22nd level*)

In contemporary terms, we see that all these properties have as necessary (but not sufficient) precondition the fact that *seem* does not assign an external theta role.

Why are all these observations considered arguments for treating *seem* as a raising verb? The answer is that the expression of certain lexical or grammatical relations requires a configurational or syntactic closeness: lexical relations such as selection, subcategorization or theta assignment can hold only of items that are structurally close. To put it differently, lexical relations are projected from the lexicon into the syntax subject to strict locality requirements.

Let us look at each case in turn:

In *John seems to sleep all the time, John* is licensed by *sleep*, by virtue of the theta relation between the two. Theta relations are considered to be local in such a way as to exclude skipping over the verb *seem*.

In the weather verb case, the idiom chunk case or the existential sentence case, the reasoning is the same, even though it does not have to do with theta relations. In each case, the subject of the main clause is licensed by the bottom predicate by syntactically projecting a property of the bottom verb according to rules that disallow skipping over the main verb.

1.1.2 Semantic representation of raising verbs

There is another, less straightforward reason why *seem* is treated like a raising verb. Consider a pair of sentences as:

(3) (a) It seems that John sleeps all day.
     (b) John seems to sleep all day.

Informally speaking, we want to say that their common semantic representation is SEEM(SLEEP(John)). It is quite straightforward to get this representation from the first sentence: *John* appears as a syntactic argument of the verb *sleep* as well as

an argument of the corresponding semantic predicate. This parallelism breaks down in the second case. How then is the structure of its semantic representation computed?

If *seem* is a raising verb, the relevant syntactic representation of (3), when looked at appropriately, is in fact *seems John to sleep all day* (due to the presence of the trace of *John*), i.e. essentially identical to the first sentence. The identity of semantic representations follows.

If *seem* is not a raising verb, two options arise for treating the second sentence. The first option consists in lowering the subject when computing the semantic representation of the sentence. This option is not really an alternative to the raising analysis. It is its mirror image. It is a notational variant that can be implemented consistently (e.g. locality conditions on projection of lexical properties will apply after lowering rather than before raising, etc.). The second option consists in denying identity of semantic representation for the two sentences and in arguing that the representation for the second sentence is (SEEM(SLEEP))(John). This comes down to treating *seem* as a function mapping a predicate P into a predicate P′ such that P′ inherits all the properties of P relating to its external argument if any. This option, although not impossible, requires strong empirical motivation, for it claims that there are two verbs *seem*. Ignoring the experiencer, the verb *seem* appearing in the second sentence would be a function mapping a predicate onto a predicate and taking one or no individual as argument, depending on whether or not the embedded verb has an external argument. The verb *seem* that appears in the first sentence would be a one-place predicate taking a proposition as argument. This empirical motivation is, to our knowledge, lacking.

[. . .]

1.2   Raising INFL

1.2.1   The raising properties of INFL

We now turn to the case of INFL. We want to show that the minimal analysis of tensed INFL is that it is a raising category. In order to make the point, let us consider some prototypical material in INFL, the modal *will*. Everything that follows applies equally well to anything else that might appear in tensed INFL. We have the following observations:

(i)   *will* does not assign an external theta role;
(ii)   *will* allows as subject an NP licensed by the predicate embedded under it:
          external argument of a predicate (*John will sleep*)
          weather *it* (*it will rain*)
          idiom chunks (*the cat will be out of the bag*)
          existential *there* (*there will be a griffin on the 22nd level*)

These are diagnostic properties of raising items. We therefore conclude that *will* is a raising verb. By the same arguments all the modals are raising categories, *do* is a raising verb, and more generally tensed INFL is a raising category. By the same argument, if negation heads a NEGP complement of INFL as recently suggested,

negation is a raising category. By the same argument, if INFL is split between Tense heading a TP and AGR heading an AGRP (Pollock 1989), both T and AGR are raising categories. By the same argument, aspectual verbs (perfective *have* and *avoir*, passive *be* and *être*, progressive *be*), which are analyzed as heading their own VP and taking VP complements are raising verbs.

This is a strong consequence. Suppose, as we claim, that aspect is determined outside the maximal projection of the verb, and that the maximal projection of the verb contains all its arguments. This would suggest that the thematic properties of a predicate are independent of the aspectual properties of the clause it appears in, although aspect could be dependent on some thematic properties. There are superficial examples of such dependencies: if a verb takes an agent, it can be put in the progressive. We expect no dependencies going the other way (although Campbell 1989 argues otherwise).

[. . .]

1.2.3 The existence argument

So far, we have shown that there is some motivation for taking tensed INFL (and aspectual verbs . . .) to be a raising verb, and furthermore, that not assuming this leads to some complications. We now show that the grammar of English (or French) as it stands generates raising structures with INFL already.

This argument is an elaboration of remarks found in Stowell (1983). Stowell shows that small clauses of the type $Y = [NP \ X^n]$ exist for X = A, P, N and V. The interesting point is that he establishes directly (i.e. by exhibiting) and indirectly (i.e. by generalizing the structures of the type Y with all the lexical X's) the existence of small clauses with verbal heads, In other words, the kind of constituent we postulate as complement of INFL has been argued to exist independently of our proposal.

In such small clause structures, the external argument of the predicate of the small clause *is* syntactically projected as sister of a projection of this predicate. This establishes that a subject (or, more precisely an external argument) can be projected as sister to its predicate.

Let us apply this to VP's. If a VP with an external argument can realize this argument as its sister in small clauses, nothing prevents this projection rule from applying in clauses as well. How can we avoid projecting the subject of a VP as sister to this VP in a clausal structure? In other words, the very existence of small clauses triggers the generation of structures like (2). The real question becomes whether the standard clause structure in (5) [see below, p. 00] is ever justified. In the absence of arguments in favour of it, it has become superfluous. Note finally that the discussions in the previous subsections all point toward the same conclusion: the hypothesis that theta roles are assigned under sisterhood indicate that NP* and VP in (2) are indeed sister nodes and that I and NP* are not; the fact that small clauses exist independently indicates that the sequence NP VP in (2) forms a constituent excluding I, assuming the standard approach to small clauses according to which the predicate and its subject form a constituent excluding the rest (an assumption presumably having to do with the general relationship between semantic types and syntactic constituents). This is what we will continue to assume.

1.3   Additional arguments

1.3.1   The X-bar theoretic argument: VSO languages

First, we suppose that the null assumption concerning language variation is that it does not exist. In the absence of (learnable) evidence to the contrary, language structure does not vary. This puts an upper bound on how complex or remote from the primary linguistic data parameters can be. To be sure, detailed claims about the nature of this upper bound are intimately tied to the theory of how language acquisition actually proceeds. However, it appears plausible to suppose that covert structural differences such as that between the standard clause structure and the one we propose (or more generally differences of hierarchical organization of constituents) will be beyond this upper bound. Consequently, if some language can be shown to have a clausal structure of the type illustrated in (2), English will be assumed to have it too and reciprocally.

Consider a VSO language like Welsh or Irish. Assume the correctness of the standard clausal structure repeated below:

(5) $[_{IP}$ [spec, IP] $[_{I'}$ I $[_{VP}$ $[_{V'}$ V NP]]]]

Then, S(ubject) is generated as the specifier of I; V and O are generated as part of a constituent in the VP excluding S, and the VP is the complement of I. This means that a VSO language cannot be base generated as VSO, since V and O must form a constituent excluding S. In other words, a VSO structure involves movement. This conclusion, which extrapolates constituent organization from the way it is in English or French to languages that look superficially different is supported empirically in various ways. As noted by many (Emonds [1980], Jones and Thomas 1977, Harlow 1981), one supporting fact is that Welsh and Irish are also AuxSVO languages if there is an auxiliary. This suggests that the VSO structure involves movement of the V to the position that an overt Aux otherwise occupies. What is this position? Verb movement is by now fairly well documented. This documentation suggests that two landing positions for this verb movement are a priori plausible: it could be INFL, or it could be COMP. In the well-documented cases (the Germanic languages, Den Besten, 1983 and others), movement of V to COMP takes place only in the absence of an overt complementizer. If a C-position is filled, it cannot be a landing site for V movement. If the C-position is not filled, it is available as landing site: this is why, in general, V to COMP (i.e. V-second) is only observed in root contexts. In Irish or Welsh, the VSO order is observed both in root clauses, and in non root clauses. This in itself is not significant. However, it occurs in clauses containing overt complementizers, e.g. embedded clauses. This indicates that the simplest assumption is that the VSO order is derived by V-movement to INFL. This kind of V-movement, clearly found in French (cf. Emonds 1978) or in Vata (cf. Koopman 1984) is unaffected by the presence of complementizers. From this, we may conclude that the AuxSVO order is simply the base generated order, with V and O being part of a VP excluding S. So the structure of Irish/Welsh is: INFL SVO. The problem with this is apparent. If structural variation is preferably

avoided, S should be specifier of INFL as in (5) above, and the VP containing VO the complement of I. If the AUX SVO is base generated as such, S, the specifier of I, intervenes between INFL and its complement VP. But this contradicts X-bar theory which claims that a head and its complement form a constituent excluding the specifier. It also contradicts the facts of Irish/Welsh, which do not allow a specifier to so intervene. Keeping to minimal assumptions, there are several possibilities: (i) S is specifier of I, but the INFL S VO structure itself involves movement; or (ii) VP is not a complement of INFL; or (iii) S is not specifier of I. The problem with (i) is that it is unclear what kind of movement it could be. For example, there is no plausible landing site between INFL and VP that would c-command the launching site given the right branching character of the language. (ii) also raises questions. INFL is not a lexical category. A priori, then, we do not expect radical language variation in its complement structure. If INFL takes a VP complement in English or French, it should do so here too. Koopman (1987) provides independent evidence based on Bambara that INFL subcategorizes for an XP complement, and VP in particular.

Suppose then that S is not specifier of I. We are led to a structure of the sort:

(6)   INFL NP VP

It turns out that there is substantial evidence internal to Irish and Welsh supporting this conclusion as Chung and McCloskey (1987) have shown for Irish, and Harlow (1981) for Welsh (cf. also Sproat 1985, Koopman 1984). Assume that, more precisely, the structure is: INFL [NP VP], with NP and VP forming a constituent, as Chung and McCloskey propose. Given this conclusion, we now face the reverse problem: if S is not specifier of INFL in these languages, it should not be so in English either. If we cannot make Irish like English, we should, at the appropriate level of linguistic representation, try to make English like Irish: the simplest assumption seems to be that S is not generated as specifier of INFL in English either. Rather, it is generated as a sister to VP. By the same reasoning as above, if we observe the surface word order S INFL VO, it must be because S moved from its base-generated position to pre-INFL position. Basically this leads to the adoption of the structure given in (2). S is base generated in the position NP*, and is moved to specifier of INFL position NP^.

1.3.2   Agreement in Arabic

Standard Arabic is a language in which the orders VSO and SVO are both observed in simple clauses. The interesting feature here is the agreement pattern. In the order VSO, the verb only exhibits a default number agreement (3rd person singular). With the SVO order, the verb fully agrees with S. How does agreement in clauses function? Let us assume the following standard version. Agreement is the morphological reflex of a relation between INFL and its specifier, or more generally, between a head and its specifier. This property of INFL is realized on the verb because the verb moves into INFL. As Mohammad (1989) observes, the Standard Arabic agreement pattern can be analyzed quite

simply under the above treatment of VSO order coupled with assumption (2). VSO in Arabic, Mohammad argues, results from the obligatory movement of the verb to INFL, i.e. from an underlying INFL SVO. If nothing further takes place, INFL has a silent expletive specifier as Mohammad argues (or no specifier position at all) and agreement gets the default value, namely 3rd person singular.

In the case of the SVO order, however, the derivation from an underlying INFL S VO comprises one more step. First, the V obligatorily moves to INFL position. In order to reestablish the SVO surface order, it must be assumed, just like in English, that S moves to specifier of I. Consequently, V in INFL has a specifier and fully agrees with it: we get subject/verb agreement (see Mohammad, 1989, for a detailed analysis).

1.3.3 Q-float

The argument for structure (2) based on the distribution of Q's rightward 'floated' from their NP is extensively developed in Sportiche ([1988]). We limit ourselves here to a short outline of the logic of the argument.

In the first instance, such Q's appear between INFL and VP. It behaves with respect to the NP it is related to just like an anaphor does with respect to its antecedent. Yet, it can be shown that it is not an anaphor. These central observations can be explained by the following scenario.

(i)   The clausal structure (2) is correct.
(ii)   Movement takes place from NP* to NP^, leaving a trace.
(iii) This movement optionally pied pipes Q. If not, we have a floated Q next to the trace NP*.

This explains why Q appears between INFL and VP: simply because NP* does (subjects of small clauses may precede their predicate). This also explains the anaphoric relation. This relation really holds between NP^ and its trace NP*. Q, being adjacent to this trace, gives the illusion of being anaphorically dependent on NP^. More generally, any time a structure contains an empty category *e* dependent upon another NP**, if *e* is modified by Q, the illusion that the Q has floated off NP** will be created. Reciprocally, if a Q is stranded, we may suspect the existence of a covert NP adjacent to it.

1.3.4 Q-lowering

Consider the following sentence:

(7)   A griffin seems to be lurking on the 25th level.

In such an example, the indefinite subject can be understood either outside the scope of the verb *seem* (i.e. *There is a griffin such that it seems to be lurking on the 25th level*) or inside it (i.e. *It seems that a griffin is lurking on the 25th level*). As May (1977, 1985) has discussed, this is due to the fact that the main verb is a raising

verb, so that the second interpretation can be reconstructed by 'lowering' the main subject into its base position. Aoun and Li [1989] propose a different treatment of this ambiguity. However, their approach shares with May's approach the idea that the crucial factor is the presence of a trace of the main subject in the embedded clause.

The same 'lowering' effect is observed with the following structures:

(8)   A griffin might be lurking on the 25th level.

The two interpretations here are one with main subject outside the scope of the modal: *there is a griffin and it might be lurking on the 25th level*, and one with the main subject inside the scope of the modal: *It might be that a griffin is lurking on the 25th level*.

According to (2), *might* in INFL is a raising category, i.e. followed by a trace of the subject: the facts follow.

1.3.5   Possible idioms

We can elaborate on the argument for raising based on the distribution of idiom chunks. Idiomatic expressions are extremely common in English. What possible shape can they take? The semantics of idioms must be stated in some component of the grammar. The natural place is the lexicon, the repositiory [sic] of idiosyncratic information. It would appear rather natural at first that idioms are simply fully specified constituents with an idiomatic reading. It is quite easy to exhibit counterexamples to this generalization. Many idioms contain open positions: e.g. *lose one's cool*.

Surely however, there are constraints as to the open or variable positions that an idiom can contain. Otherwise, nothing would prevent an idiom only containing the italicized elements as part of it:

(9)   A *pale* man *slowly* put flowers next to John.

Since such idioms do not exist, it is not obvious what this one could mean. Assume the following: only if *pale* modifies a subject and *slowly* co-occurs in the same proposition does the following idiomatic interpretation arise: pale means unknown to the speaker and the action was done in a roundabout way. In other words, when uttering (9), the speaker means that the man unknown to me put flowers next to John in a roundabout way. More generally, *pale X slowly verbed* . . ., stands for *X unknown to speaker verbed in a roundabout way*.

In order to exclude impossible idioms, we may propose that idioms must at least meet a condition stronger than (10) below:

(10)   If X is the minimal constituent containing all the idiomatic material, the head of X is part of the idiom.

Applied to (9), this would mean that the head of S should be part of the idiom, which it is not. So this idiom is excluded. Similarly, the specifier of a noun and the complement of a noun cannot form an idiom together, with the head noun an open position. A verb taking an NP and a PP as complements cannot be an open position of the idiom made up of the NP and the PP.

Consider such idioms as *the shit will hit the fan*. For such idioms, (10) predicts that the minimal constituent containing the idiomatic material S = IP should have its head fixed. But there is no constraint whatever on the content of INFL in such sentences. If the structure of S is SUBJECT [$_I$, I VP], these idioms constitute a systematic class of counterexamples to (10). Note that it is not the case that non-lexical categories cannot be part of idioms. For many French speakers examples such as *les carottes sont cuites*/(the carrots are cooked) 'all is lost' in which INFL is fixed, or *Que le diable l'emporte*/(let the devil take him away) 'let him be damned', in which the complementizer is obligatory, are idioms. Of course, because of movement processes, a condition like (10) must be understood to hold at D-structure. Now, it is clear that adopting our proposal on clause structure in (2) removes this class of idioms as exceptions. At D-Structure, the minimal constituent containing all the idiomatic material is V^{max}, excluding I.

1.4   Conclusion: What is a subject?

Consider the superficial properties of the subject of a clause in English (or French). A (non-derived) subject:

(i)   occupies the position specifier of I, i.e. [NP, S] in usual terms;
(ii)   is the external argument;
(iii) triggers agreement with the verb.

These three properties are usually all analysed as a property of the specifier of INFL. Adopting (2), we see that the three properties do not correlate. Property (ii) is a property of [NP, V^{max}]. Properties (i) and (iii) do correlate due to the fact that we get specifier–head agreement, and we get a merger of V and I (by V to I or by Affix-hopping).

The distinction between (i) and (ii) sheds a different light on the question of what the head of S = IP is. Should the subject be contained in the maximal projection of V, or is the standard clause structure (as in (4)) correct? (See Marantz 1979, who suggests that V is the head of S, and Hornstein, 1977, who suggests that Aux is the head of S.) The answer according to (2) is positive, if we take subject as meaning external argument, negative, if we take subject as meaning the NP triggering agreement. In a sense, then, both positions are correct: the maximal projection of V contains the subject understood as external argument but does not contain the subject understood as the NP triggering agreement.

[. . .]

Note

2 Throughout the paper, we will use NP^ and NP* to refer either to these particular positions or to the content of these positions. The context will make it clear what is meant in each particular case. We also need to distinguish between the various projections of V. We call VP the phrasal projection of V, while V^max is its maximal projection.

19.3   Questions pertaining to Koopman and Sportiche (1991)

1 Surface subject position had long been considered to be a derived position in the case of passives such as *John was arrested by the police yesterday*, in which *John* raises from object position. Prior to Koopman and Sportiche's paper, the same kind of idea had come to be widely accepted for unaccusative sentences like *The letter arrived yesterday*, in which *the letter* raises from object position (cf. Perlmutter 1978; Burzio 1986). Koopman and Sportiche's paper further generalizes the idea that surface subject position is moved into. If they are right, then surface subject position is invariably filled by internal merge, rather than by external merge. To what extent can this view of the sentential surface subject position be generalized to derived nominals? Give your reasons, bringing in Chomsky (1970), Kayne (2008), Taraldsen (2012), and the contrast between *There has arrived a letter for you* and **there's arrival of a letter for you*.

2 Thinking of the very last part of the previous question, we can note that English gerunds containing a lexical subject not followed by *'s* act like finite sentences, insofar as expletive *there* is possible, as in *We weren't aware of there having been a problem*. Yet gerunds with *'s* have a different status, with ?*We weren't aware of there's having been a problem* often rejected, but sometimes accepted. How might this difference between the two types of gerunds be expressed? Bring in the crosslinguistic question of the status of derived nominals vs. gerunds vs. infinitives.

3 In the case of transitives and unergatives, Koopman and Sportiche take the external argument that moves into surface subject position to originate in a specifier-like position sister to VP. More recently there has developed a widely held view to the effect that the external argument originates as the specifier of a functional element called little v. Discuss the pros and cons of taking this v, in the case of transitives and unergatives, to be a silent counterpart of the lexical verb *do*, or of the auxiliary verb *do*. (Bring in Haddican 2007.)

4 Collins (2005) takes the unifying position that the external argument should be merged in the same position in passives as in actives. If so, then the object in passives, in raising to surface subject position, must apparently cross the position of the external argument. Collins shows that this potential intervention problem is neutralized if the object in passives is "smuggled" past the external argument by movement into Spec,VoiceP of the past participle phrase that contains the object. Discuss the extent to which Collins's proposal can or should be generalized to passive-like derived nominals such as *its destruction*

by the soldiers. Bring in the fact that in some languages (e.g., Italian, German) the preposition that appears before the external argument/agent in sentential passives (*by* in English, *da* in Italian, *von* in German) cannot appear in derived nominals. (Extra credit: Discuss possible generalizations of Collins's proposal to French sentences like *Ils ont fait arrêter Jean par la police* ('they have made arrest John by the police' = 'they had John arrested by the police'), in which there is an infinitive ('arrêter') rather than a past participle, yet in which the embedded sentence is nonetheless passive-like in the ways discussed by Kayne 1975, sections 3.5, 3.6.) (Double extra credit: Relate this kind of French sentence to the German and Dutch IPP phenomenon.)

5 The little v in whose Spec the external argument is merged is higher than VP. But what is its hierarchical position relative to other projections below Tense? Bring in Travis (1991, 1992) and Cinque (1999).

6 If Kayne (1999, 2002a, 2004) is correct to take certain prepositions to be merged above VP (rather than with what we call their objects), we can ask, much as in the previous question, where those Ps are merged relative to little v. Discuss considerations bearing on the answer.

7 Sportiche (1988) provides evidence based on Q-floating in favor of the idea that surface subjects invariably originate in a lower position, insofar as he takes the position of a floated quantifier to reflect a position through which its antecedent has passed. Thus in *The students have all done their homework*, the phrase *the students* must have passed through a position adjacent to that of *all*. Given standard assumptions, there arises a problem with passives, namely that sentences like **The students haven't been arrested all* and (for some speakers) *(*)The students haven't been all arrested* are unexpectedly not possible. Discuss the possibility of solving this problem by bringing together the obligatoriness of negative phrase movement as discussed by Haegeman and Zanuttini (1991, Ch. 17 of this volume) with the proposals of Beghelli and Stowell (1997, Ch. 23 of this volume) on quantifiers.

8 Some varieties of English allow floated quantifier phrases to contain an overt pronoun, as in *They've all of them decided to join us*. To what extent are these a problem for Sportiche (1988) (bring in Kayne 2002b)? (Extra credit: Make a proposal as to why the presence of this pronoun widens the class of possible floated/stranded quantifiers, as illustrated by (in at least some varieties of English) *They haven't any *(of them) decided to join us, They've none *(of them) read your paper*? Bring in Kobuchi-Philip 2007.)

9 Taking external arguments to originate in Spec,vP, would one expect those external arguments to be able to bear oblique Case, or not? Give your reasons.

10 Some speakers of English (cf. Kimball and Aissen 1971; Kayne 2003, (28)) accept to some extent (with *think* rather than *thinks*, and especially with stress on *all*) sentences like *Those children, who John all think should be invited, are much too young*. Focusing on *all*, discuss the ramifications of such sentences for how successive cyclic movement works and for questions concerning what phrases count as phases.

11 VSO languages give the initial impression that, from Koopman and Sportiche's perspective, the subject might have remained in situ, that is, in

its external merge position. On the other hand, Kayne (2011, (6)) has suggested that "All arguments must move at least once." Discuss the pros and cons of Kayne's proposal, including in your discussion the question of the syntax of compounds and incorporation. (Extra credit: What might the requirement that all arguments move at least once follow from?)

12 Discuss the interaction between Sportiche's (1988) analysis of floating quantifiers as stranded quantifiers and the fact that in derived nominals the derived subject (followed by *'s*, if lexical) cannot be directly associated with a floated quantifier, as illustrated by examples like *the children's all refusal to leave*, *the bridges' all destruction by the enemy*. Bring in the acceptability of *the children's refusal to all leave at the same time* (cf. Baltin 1985), as well as the unacceptability, against the background of Ross (1969), of *the all realization that they were going to lose*.

13 Despite the similarity between *all of the cake* and *the whole cake*, there is a sharp contrast between *The cake/it has all been ruined* and *The cake/it has whole been ruined*. How might this be made to follow from Sportiche (1988), or, alternatively in part, from Shlonsky (1991)? Somewhat similarly, we have both *All linguists are intelligent* and *Three linguists are intelligent*, but a contrast between *Linguists are all intelligent* and *Linguists are three intelligent*. Why exactly is that? (Extra credit: To what extent should Sportiche 1988 or Shlonsky 1991 be extended to cover "numeral+classifier" floating of the sort found in various East Asian languages? Bring in Simpson 2005 and Watanabe 2012.)

14 How exactly does taking the external argument of transitives and unergatives to be merged in Spec, v bear on the question of ergative Case? Bring in Koopman (2012).

15 Koopman and Sportiche emphasize that having the surface subject position systematically filled by movement makes all sentences instances of raising in a way akin to what had long been the standard view on sentences with verbs like *seem*. In English, predicate nominals are readily available, as in *Mary is/became a famous chemist*. Yet in at least some American English it is not possible to have *Mary seems/appears a famous chemist*. Why might that be, and why might it be that such sentences are acceptable in at least some British English?

16 External arguments (and perhaps arguments more generally) cannot be APs. How might this fact be related to the contrast in English between *They ended up happy* and *They ended happy up* (cf. Kayne 1985; den Dikken 1995)? (Extra credit: In contrast to English, Norwegian allows *Per ser nervøs ut* ('P sees nervous out' = 'P looks nervous'). How might this be related to the inverse contrast that has English (cf. Klima 1964; Kayne 1998), but no Scandinavian (cf. Engels 2012), allowing *John has seen nobody today*? Bring in the rest of Germanic.)

References

Aoun, J., and A. Li. 1989. Scope and constituency. *Linguistic Inquiry* 20: 141–172.

Baltin, M. 1985. Floating quantifiers, PRO, and predication. *Linguistic Inquiry* 26: 199–248.

Beghelli, F., and T. Stowell. 1997. Distributivity and negation: The syntax of *each* and *every*. In A. Szabolcsi (ed.), *Ways of Scope Taking*, Dordrecht: Kluwer, pp. 71–109.

den Besten, H. 1983. On the interaction of root transformations and lexical deletive rules. In W. Abraham (ed.), *On the Formal Syntax of the Westgermania. Linguistik Aktuell/Linguistics Today 3*, Amsterdam and Philadelphia: John Benjamins, pp. 47–131.

Burzio, L. 1986. *Italian Syntax: A Government-Binding Approach*. Dordrecht: Reidel.

Campbell, R. 1989. *The grammatical structure of verbal predicates*. Doctoral dissertation, UCLA.

Chomsky, N. 1970. Remarks on nominalization. In R. Jacobs and P. Rosenbaum (eds), *Readings in English Transformational Grammar*, Waltham, MA: Ginn, pp. 184–221.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2000. Minimalist inquiries: The framework. In R. Martin, D. Michaels, and J. Uriagereka (eds), *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, Cambridge, MA: MIT Press, pp. 89–155.

Chung, S., and J. McCloskey. 1987. Government, barriers and small clauses in Modern Irish. *Linguistic Inquiry* 18: 173–237.

Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York and Oxford: Oxford University Press.

Collins, C. 2005. A smuggling approach to the passive in English. *Syntax* 8: 81–120.

Contreras, H. 1987. Small clauses in Spanish and English. *Natural Language and Linguistic Theory* 5: 225–243.

den Dikken, M. 1995. *Particles: On the Syntax of Verb-Particle, Triadic, and Causative Constructions*. New York: Oxford University Press.

Di Sciullo, A. M., and C. Boeckx (eds). 2011. *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford Studies in Biolinguistics, Oxford and New York: Oxford University Press.

Emonds, J. 1978. The verbal complex V′-V in French. *Linguistic Inquiry* 9: 151–175.

Emonds, J. 1980. Word order in generative grammar. *Journal of Linguistic Research* 1: 33–54.

Engels, E. 2012. Scandinavian negative indefinites and cyclic linearization. *Syntax* 15: 109–141.

Haddican, B. 2007. On *egin: do*-support and VP-focus in Central and Western Basque. *Natural Language and Linguistic Theory* 25: 735–764.

Haegeman, L., and R. Zanuttini. 1991. Negative heads and the NEG criterion. *The Linguistic Review* 8: 233–251.

Harlow, S. 1981. Government and relativization in Celtic. In F. Heny (ed.), *Binding and Filtering*, London: Croom Helm and Cambridge, MA: MIT Press, pp. 213–254.

Hornstein, N. 1977. S and the X-bar convention. *Linguistic Analysis* 3: 137–176.

Jones, M., and A. Thomas. 1977. *The Welsh Language: Studies in Syntax and Semantics*. Cardiff: University of Wales Press.

Kayne, R. S. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

Kayne, R. S. 1985. Principles of particle constructions. In J. Guéron, H.-G. Obenauer, and J.-Y. Pollock (eds), *Grammatical Representation*, Dordrecht: Foris, pp. 101–140.

Kayne, R. S. 1989. Facets of Romance past participle agreement. In P. Benincà (ed.), *Dialect Variation and the Theory of Grammar*, Dordrecht: Foris, pp. 85–103. [Reprinted in Kayne (2000, Ch. 3).]

Kayne, R. S. 1998. Overt vs. covert movement. *Syntax* 1: 128–191.

Kayne, R. S. 1999. Prepositional complementizers as attractors. *Probus* 11: 39–73. [Reprinted in Kayne (2000, Ch. 14).]

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2002a. On some prepositions that look DP-internal: English *of* and French *de*. *Catalan Journal of Linguistics* 1: 71–115. [Reprinted in Kayne (2005, Ch. 7).]

Kayne, R. S. 2002b. Pronouns and their antecedents. In S. D. Epstein and T. D. Seely (eds), *Derivation and Explanation in the Minimalist Program*, Oxford: Blackwell, pp. 133–166. [Reprinted in Kayne (2005, Ch. 6).]

Kayne, R. S. 2003. Some remarks on agreement and on Heavy-NP-Shift. In M. Ukaji, M. Ike-Uchi, Y. Nishimura, and H. Nakajima (eds), *Current Issues in English Linguistics*. Vol. 2 of *Special Publications of the English Linguistic Society of Japan*, Tokyo: Kaitakusha, pp. 67–86. [Reprinted in Kayne (2005, Ch. 11).]

Kayne, R. S. 2004. Prepositions as probes. In A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures 3*, New York and Oxford: Oxford University Press, pp. 192–212.

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2008. Antisymmetry and the lexicon. In J. Rooryck and J. van Craenenbroeck (eds), *Linguistic Variation Yearbook* 8, Amsterdam and Philadelphia: John Benjamins, pp. 1–31. [Reprinted in Kayne (2010, Ch. 9) and in Di Sciullo and Boeckx (2011, Ch. 14).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2011. Why are there no directionality parameters? In M. Byram Washburn, K. McKinney-Bock, E. Varis, A. Sawyer, and B. Tomaszewicz (eds), *Proceedings of the 28th West Coast Conference on Formal Linguistics*, Somerville, MA: Cascadilla Proceedings Project, pp. 1–23.

Kimball, J., and J. Aissen. 1971. I think, you think, he think. *Linguistic Inquiry* 2: 242–246.

Kitagawa, Y. 1986. *Subjects in Japanese and English*. Doctoral dissertation, University of Massachusetts, Amherst.

Klima, E. S. 1964. Negation in English. In J. A. Fodor and J. J. Katz (eds), *The Structure of Language*, Englewood Cliffs, NJ: Prentice-Hall, pp. 246–323.

Kobuchi-Philip, M. 2007. Floating numerals and floating quantifiers. *Lingua* 117: 814–831.

Koopman, H. 1984. *The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar*. Dordrecht: Foris.

Koopman, H. 1987. On the absence of case chains in Bambara. Ms., UCLA.

Koopman, H. 2012. Samoan ergativity as double passivization. In L. Brugè, A. Cardinaletti, G. Giusti, N. Munaro, and C. Poletto (eds), *Functional Heads: The Cartography of Syntactic Structures* 7, New York and Oxford: Oxford University Press, pp.168–180.

Koopman, H., and D. Sportiche. 1991. On the position of subjects. *Lingua* 85: 211–258.

Kratzer, A. 1996. Severing the external argument from the verb. In J. Rooryck and L. Zaring (eds), *Phrase Structure and the Lexicon*, Dordrecht: Kluwer, pp. 109–137.

Kuroda, S.-Y. 1988. Whether we agree or not: A comparative syntax of English and Japanese. In W. J. Poser (ed.), *Papers from the Second International Workshop on Japanese Syntax*, Stanford, CA: CSLI, Stanford University, pp. 103–143. [Also in S.-Y. Kuroda, 1992, *Japanese Syntax and Semantics: Collected Papers*, Dordrecht:Kluwer, pp. 315–357.]

Marantz, A. 1979. Assessing the X-bar convention. Ms., MIT.

May, R. 1977. *The grammar of quantification*. Doctoral dissertation, MIT.

May, R. 1985. *Logical Form*. Cambridge, MA: MIT Press.

Mohammad, M. 1989. The problem of subject verb in Arabic: Towards a solution. In M. Eid (ed.), *Perspectives in Arabic Linguistics*, Amsterdam: John Benjamins, pp. 95–125.

Perlmutter, D. 1978. Impersonal passives and the unaccusativity hypothesis. In J. Jaeger (ed.), *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society, pp. 157–189.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Rosen, S. 1990. *Argument structure and complex predicates*. Outstanding Dissertations in Linguistics 4, New York: Garland Publishing, Inc.

Ross, J. R. 1969. On the cyclic nature of English pronominalization. In D. A. Reibel and S. A. Schane (eds), *Modern Studies in English: Readings in Transformational Grammar*, Englewood Cliffs, NJ: Prentice-Hall, pp. 187–200.

Shlonsky, U. 1991. Quantifiers as functional heads: A study of quantifier float in Hebrew. *Lingua* 84: 159–180.

Simpson, A. 2005. Classifiers and DP structure in Southeast Asia. In G. Cinque and R. S. Kayne (eds), *The Oxford Handbook of Comparative Syntax*, New York and Oxford: Oxford University Press, pp. 806–838.

Speas, M. 1986. *Adjunction and projection in syntax*. Doctoral dissertation, MIT.

Sportiche, D. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19: 425–449.

Sportiche, D. 1990. Movement, agreement, and case. Available at Lingbuzz/000020 [accessed January 2013]. [Also in D. Sportiche, 1998, *Partitions and Atoms of Clause Structure*, London: Routledge.]

Sproat, R. 1985. Welsh syntax and VSO structure. *Natural Language and Linguistic Theory* 3: 173–216.

Stowell, T. 1983. Subjects across categories. *The Linguistic Review* 2: 285–312.

Taraldsen, K. T. 2012. The structural object position of verbs and nouns. In L. Brugè, A. Cardinaletti, G. Giusti, N. Munaro, and C. Poletto (eds), *Functional Heads: The Cartography of Syntactic Structures 7*, New York and Oxford: Oxford University Press, pp. 263–274.

Travis, L. 1989. Parameters of phrase structure. In M. Baltin and A. Kroch (eds), *Alternative Conceptions of Phrase Structure*, Chicago and London: University of Chicago Press, pp. 263–279.

Travis, L. 1991. Inner aspect, derived objects, and the structure of VP. Paper presented at *NELS 22*, University of Delaware.

Travis, L. 1992. Inner aspect and the structure of VP. In M. Champagne and S. Ratté (eds), *Proceedings from the Second Annual Lexical-Syntactic Relations Workshop*, Montreal: Cahiers Linguistiques de l'Université du Québec à Montréal,1, pp. 130–146.

Watanabe, A. 2012. Measure words as nouns: A perspective from silent years. *Studia Linguistica* 66: 181–205.

Wible, D. 1990. *Subjects and the clause structure of Chinese and English*. Doctoral dissertation, University of Illinois, Urbana-Champaign.

Woolford, E. 1991. VP-internal subjects in VSO languages and nonconfigurational languages. *Linguistic Inquiry* 22: 503–540.

Zagona, K. 1982. *Government and proper government of verbal projections*. Doctoral dissertation, University of Washington, Seattle.

20

On Argument Structure and the Lexical Expression of Syntactic Relations

Kenneth Hale and Samuel Jay Keyser

1993

20.1   Introduction

Determining the exact scope of the syntactic component of the grammar is one of the leading research questions in generative linguistics. For example, a rich line of work has investigated whether or not syntactic principles govern the process of word formation. Pursuing this question, Baker (1988) proposed that, in polysynthetic languages, words with multiple roots are formed through a syntactic process labeled INCORPORATION (cf. also Baker and Hale 1990). This article by Hale and Keyser extends Baker's proposal to verbs that contain only one lexical (open class) element, arguing that they are also derived via a syntactic process. More generally, this work puts forth the view that the derivation of lexical items is constrained by syntactic principles, and therefore is a syntactic derivation.

Hale and Keyser's article argues that "unergative," or "true intransitive" verbs (Perlmutter 1978; Burzio 1981, 1986), like *laugh, sneeze, dance*, start out with a lexical projection that includes an abstract light verb and a nominal complement – the same kind of structure projected by the overt light verb *make* and its nominal complement in cases like *make trouble*. To form a denominal verb like *laugh*, the head of the nominal complement raises and incorporates into the abstract V head, obeying the syntactic principles that govern incorporation. This proposal is supported by the fact that, across languages, we find cases in which the denominal verb in one language is expressed by a combination of a light verb plus a nominal complement in another language (as in Basque, which exhibits the equivalent of *sleep do*), and by the existence of cases where the incorporation of the noun is visible.

In cases of denominal verbs that express location verbs, e.g., *shelve* and *box* (as in *shelve the books, box the apples*), Hale and Keyser argue that it is the complement

of the (abstract) preposition that incorporates, thus forming a transitive verb from an underlying ditransitive structure. For example, in the case of *shelve the books*, the denominal verb *shelve* is derived from the equivalent of $[_V$ *put* $[_{PP}$ *(on) the shelf]]* via incorporation of the noun *shelf* into the abstract light verb.

Evidence in support of the proposal that the derivation of denominal verbs is governed by syntactic principles comes from the fact that it is always the head of the structural complement of the light verb that can incorporate and gives rise to a denominal verb; the subject, in contrast, can never do so. For example, in English we see cases like (1a) that alternate with cases like (1b), which are argued to result from incorporation of the complement into the V. But there is no denominal verb like (1c), which would be the result of the incorporation of the subject into the abstract light verb:

(1) a.   A cow had a calf.
     b.   A cow calved.
     c.   *It cowed a calf.

On the syntactic view of word formation presented here, this gap results from the structural difference between the head–complement and the head–specifier relation. More precisely, incorporation of a complement into a c-commanding head is possible, but incorporation of a specifier is not.

The debate on the extent to which word formation is syntactic has a longstanding tradition in the field, and has been approached from various points of view. The influential *Distributed Morphology* framework of Halle and Marantz (1993) and Marantz (1997), as well as various other works (cf. Kayne 2000, 2005, 2010; Julien 2002; Brody 2003; among others) build on the idea that there is no specific morphological structure-building component separate from the syntax. The Distributed Morphology literature has also drawn attention to the difference between noun–based and root–based word formation (Marantz 1997; Arad 2003; Levinson 2007), arguing that only roots are present in the lexicon, and that categorial specifications are introduced as functional heads. Kayne (2008), in contrast, proposes that there is one and only one open class category in the lexicon, namely N, all verbs being light verbs, with the possibility of an incorporated noun – a "strong reading" of Hale and Keyser (1993).

20.2   From "ON ARGUMENT STRUCTURE AND THE LEXICAL EXPRESSION OF SYNTACTIC RELATIONS"

1   Introduction

For a number of years we have been investigating the relation between lexical items, particularly verbs, and the syntactic structures into which they enter. This is one part of a general program that seeks to explore and understand the implications of the thesis that syntax is projected from the lexicon (see, among other works, Chomsky 1981).
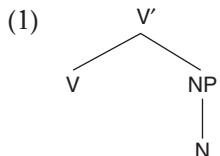
During the course of our investigations, we have become persuaded that the proper representation of predicate argument structure is itself a syntax. That is to say, as a matter of strictly lexical representation, each lexical head projects its category to a phrasal level and determines within that projection an unambiguous system of structural relations holding between the head, its categorial projections, and its arguments (specifier, if present, and complement). We will refer to these projections sometimes as *lexical argument structures* and sometimes as *lexical relational structures* (LRSs), and we will use the now conventional tree diagrams to represent them in our discussions here. The diagrams will also make use of the conventional labels for the lexical categories V, N, P, A, and their phrasal projections V′, VP, and so on, but these are to be understood in terms of a particular theory of lexical categories, to be introduced below.

We have been led to this syntactic view of lexical argument structure in large part through an investigation of denominal verbs of the type represented by *calve, lamb, shelve, bottle, saddle, hobble*, and the like. See Clark and Clark 1979 for an impressive array of denominal verb types, and see Talmy 1985 for a discussion of a wide range of lexicalization patterns, including so-called *conflation*, a term we sometimes use to refer to derivation of denominal verbs of the type under consideration here.
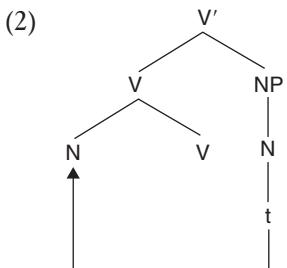
Assuming that these verbs are in fact derived from nouns, the process involved in their derivation is almost certainly *lexical*, in the widely accepted sense of that term (see Chomsky 1970). But, we argue, this is quite independent of the question of whether the process is *syntactic* in some equally accepted sense. Thus, for example, if established principles of syntax function to constrain denominal verb derivations, then the simplest assumption to make is that these derivations are in fact syntactic in nature.

The evidence we hint at in the foregoing paragraph only makes sense, of course, within a particular *theory* of denominal verb formation. We assume that verbs like *shelve* and *saddle* are formed by means of the "head movement" variant of Move α – more specifically, by means of the process known as *incorporation*, whose theoretical properties have been studied in detail by Baker (1988). If denominal verb formation takes place by means of incorporation, then it is to be expected that it would be subject to syntactic principles that govern the application of incorporation (e.g., those identified in Baker 1988 and in Baker and Hale 1990). This would be the "evidence" in favor of the syntactic theory (cf. Walinska de Hackbeil 1986, 1989, for a conception of denominal verb formation closely similar to ours).

The so-called unergative verbs (see Perlmutter 1978, Pullum 1988), all called simply *(true) intransitive* verbs (Burzio 1981), represent by far the simplest class of denominal verbs derived by incorporation. For English, these include, among many others, the verbs *laugh, sneeze, neigh, dance, calve*. As shown in (1), their initial lexical projection is simply that of a verb and a nominal complement.

(1)

This structure is the same as that projected by verbs such as *make* (as in *make trouble*, *have* (as in *have puppies*), and *do* (as in *do a jig*). The difference is that the lexical structure representation of an unergative verb, like *laugh*, involves incorporation, into an abstract V, of the nominal head N of its NP complement. We assume, with Baker (1988) and others, that this process is as depicted in (2); that is, the head N of the NP governed by the V is moved and adjoined to the latter. The resulting "compound," of which only the N component is phonologically realized, corresponds to the denominal verb.

(2)

V′
├── V
│   ├── N
│   └── V
└── NP
    └── N
        └── t

The derivation depicted in (2) conforms to the principles that constrain the syntactic process of incorporation. In particular, it conforms to the Head Movement Constraint in (3) (from Travis 1984; also see Baker 1988).

(3)   *The Head Movement Constraint*

An X^0 may only move into the Y^0 that properly governs it.

To this extent at least, we are justified in our assumption that unergative verbs have an initial lexical structure of the simple transitive type. This position is strengthened by the observation that the unergatives of one language are matched in other languages either (i) by the simple transitive VP structure without incorporation (e.g., Basque *lo egin* (sleep do) 'sleep') or (ii) by the transitive VP modified by "visible" incorporation (e.g., Jemez *-záae-'a* (song-do) 'sing').
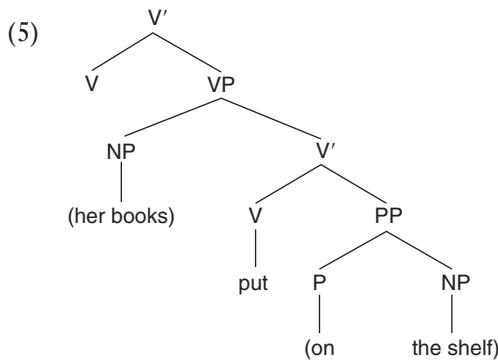
The relation between the simple transitive structure (1) and the incorporation structure (2) belongs to the class of phenomena sometimes known as *lexical alternations*, whose study has been so revealing in relation to the lexical representations of argument structure (see, for example, Levin 1991). If we are correct in our belief that derivations of the type represented by (2) involve a syntactic process, defined over syntactic objects, then this has clear implications in relation to the nature of argument structure. Argument structure is syntactic, necessarily, since it is to be identified with the syntactic structures projected by lexical heads.

A somewhat more complex class of denominal verbs is that represented by "location" verbs, like *shelve* (as in *shelve the books*), *corral* (as in *corral the horses*), *box* (as in *box the apples*), and "locatum" verbs, like *saddle* (as in *saddle the horse*), *hobble* (as in *hobble the mule*). We will assume that these also, like verbs of the simpler unergative type, are formed by incorporation.

We suppose that, abstractly speaking, the LRS representation of location verbs is identical to that of the English verb *put*, as used in such sentences as (4).
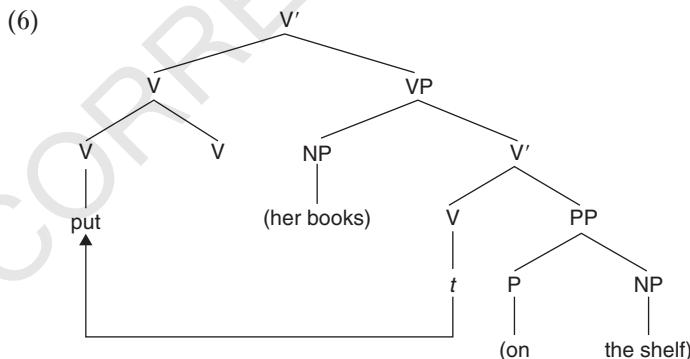
(4)   She put her books on the shelf.

And we assign to *put* the structure set out in (5).

(5)

This is the structure that Larson (1988) assigns to verbs of the type represented by *put*, as well as to "double object" or "dative" verbs like *give*. For Larson, this is the D-Structure representation of these verbs. For us, this is their Lexical Argument Structure representation (which, of course, determines the D-Structure representation). In both cases the representations are fundamentally syntactic in the sense that they are structures over which fundamental syntactic relations and principles are defined.
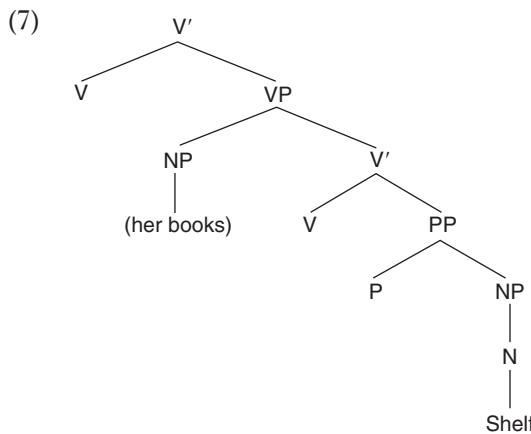
The complex structure (5) is the initial lexical representation of English *put*. The form that appears in the D-Structure representations of sentences containing this verb is derived by head movement, or incorporation, which, in this instance, moves the lower V up into the matrix "clause" and adjoins it to the matrix verb, as depicted in (6).

(6)
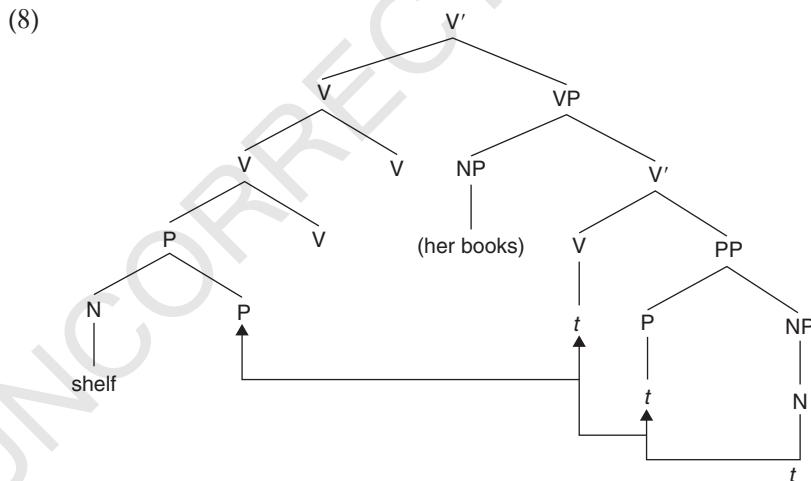
Like the noun incorporation process involved in (2), the verb incorporation of (6) is in conformity with the Head Movement Constraint, since the matrix verb properly governs the lower verb. Our reasons for assuming the structures (5) and (6) will be explicated in part later in this paper, but they are essentially the reasons found in Larson 1988, Hale 1989, and Hale and Keyser 1991.

It is a fundamental assumption of our account that English verbs like *shelve*, and other location verbs, are "denominal" precisely in the sense that they are derived

by head movement. Their initial LRS representations share their essential relational structure with verbs like *put*, with the exception that the morphological "constant" (i.e., the phonologically overt morpheme ultimately realized in the matrix verb position) is not a verb but a noun, heading the complement of the PP in the LRS representation. This is shown in (7).

(7)

As shown in (8), the surface form of the verb is derived by three applications of head movement, the first of which incorporates the lower N (*shelf*) into the P that governs it. The compound so formed is then moved into the verb that governs it, there forming a compound that makes the final move to incorporate into the matrix verb.

(8)

Each step in this derivation conforms to the Head Movement Constraint. At each point, incorporation involves movement into a head that properly governs the moving element.

With the background afforded by these two examples – denominal unergative and location verbs – we can illustrate the central point of this introduction. By hypothesis, these two verb types involve incorporation in their derivations, and

the process that effects the incorporation conforms to the Head Movement Constraint. It is appropriate to view this constraint as a special case of the Empty Category Principle (ECP) in (9) (see Chomsky 1981:273).
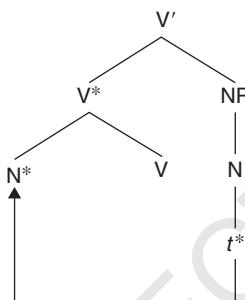
(9)   *Empty Category Principle*

[e] (an empty category) must be properly governed.

For present purposes, we will simply assume Baker's (1988: 51–68) argument that the Head Movement Constraint can be derived from the ECP, the trace of head movement being the relevant empty category [e] of (9). The ECP, then, is the effective principle constraining head movement. For reasons noted by Baker (1988: 54ff.), *antecedent government* is the relevant government relation for head movement in relation to the ECP.

In the incorporation structures of interest here, an empty category will be properly governed if, among other things, it is antecedent-governed by the relevant incorporated head, for example, by the incorporated N in (2), repeated here as (10).

(10)

We assume that this condition is met here and in other cases we have examined. The c-command requirement is met under the assumption that the adjunction node (e.g., the upper V-node in (10)) immediately dominating the incorporated element does not count as the first branching node relevant in defining the c-command relation (see Baker 1988: 54–55), a circumstance that will follow automatically if, as we will suppose, a zero-level adjunction node acquires the index (or indices, symbolized*) of the adjunct it dominates. In effect, the indexed adjunction node functions as the required antecedent, satisfying the government requirement of the ECP. And the locality requirement for government is met by virtue of the fact that no barrier intervenes between the antecedent and the trace.

Similarly, in more complex derivations (e.g., that of *shelve* depicted in (8), each instance of incorporation results in an antecedent-trace relation that satisfies the ECP. This follows, since each trace in a well-formed derivation is governed by a local c-commanding head that is coindexed with it, given our assumptions.

Let us imagine that the unergative and location verbs briefly examined here are in fact derived by incorporation, or head movement, in the manner indicated. Head movement is a process that is constrained by syntactic principles, and it is expected therefore to limit the range of theoretically possible incorporations. If this is empirically the case, then to that extent we are justified in our belief

that these lexical processes are syntactic in nature. And, most important, since the lexical processes we are examining affect the argument structures of lexical items, we are justified in our belief that argument structures are themselves syntactic objects.

To put this another way, if denominal verb formation were not constrained by syntactic principles—if it were simply a process of category change, say – then the range of possible denominal verb types would be expected to include verbs of the sort exemplified in (11). But English simply does not have verbs of this type – transitive verbs that take an expletive subject and have meanings corresponding more or less to the parenthetic paraphrases given here.

(11) a. *It cowed a calf.
          (cf. A cow had a calf. A cow calved.)
      b. *It mared a foal.
          (cf. A mare had a foal. A mare foaled.)
      c. *It dusted the horses blind.
          (cf. The dust made the horses blind. The dust blinded the horses.)
      d. *It machined the wine into bottles.
          (cf. A machine got the wine into bottles. A machine bottled the wine.)
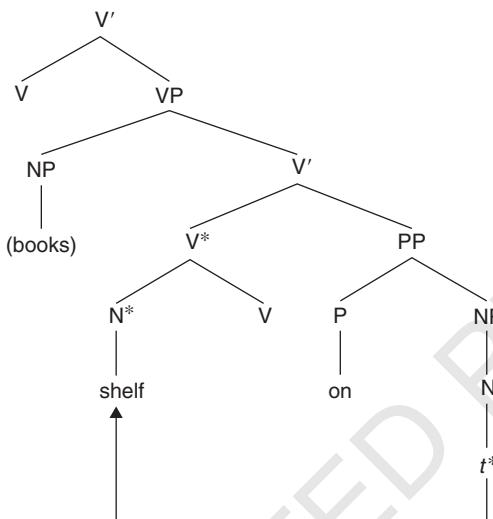
This gap in the English verbal lexicon can be explained within the incorporation theory of denominal verb formation under the natural assumption that the hypothetical verbs of (11) are formed by incorporation of a subject, rather than of a complement. It is well known that a subject (i.e., a subject that originates as an external argument) cannot incorporate into the verb that heads its predicate (see Baker 1988, Mithun 1984). Presumably, incorporation from the subject position, external to VP, would violate the ECP. The question may in fact be academic. We will argue later that the subject of verbs of the type represented in (11) is external in the sense that it is not present at all in Lexical Relational Structure. Lexical incorporation would therefore be impossible. In any event, the incorporation theory of denominal verb formation, a theory determined by syntactic principles, accounts for the nonexistence of the verbs of (11), and of their counterparts in other languages.

English has many lexical items of the form [V . . . P], where P is a prepositional particle, such as *take (the business) over, take (a stray cat) in, turn (the stove) on, plow (the corn) under*. Whatever the source and proper analysis of these items, there are no such [V . . . P] items that correspond to the hypothetical verbs of the sentences in (12).

(12) a. *He shelved the books on.
          (cf. He put the books on a shelf. He shelved the books.)
      b. *He corralled the horses in.
          (cf. He put the horses in a corral. He corralled the horses.)
      c. *He bottled the wine in.
          (cf. He put the wine in bottles. He bottled the wine.)

Each of these hypothetical items, *shelve (books) on*, *corral (horses) in*, *bottle (wine) in*, is derived by incorporation of the noun that heads the complement of the preposition, as shown in (13). The trace of incorporation is thus "too far" from its antecedent and is therefore not properly governed, violating the ECP.

(13)

```
                          V′
                        /    \
                      V        VP
                            /      \
                         NP          V′
                          |        /    \
                      (books)    V*        PP
                               /   \      /   \
                             N*     V    P      NP
                              |          |       |
                            shelf        on      N
                              ▲                   |
                              |                   t*
                              |_____|
```
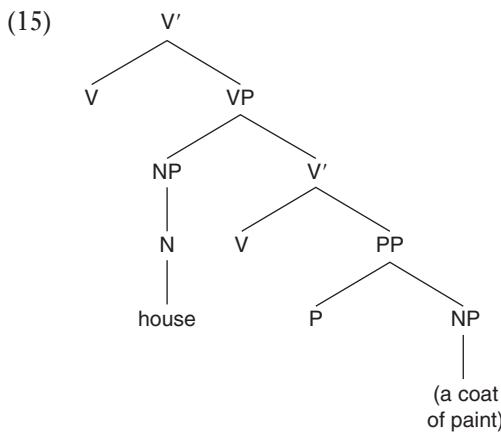
Although the trace is coindexed with the verb to which its antecedent is adjoined (as indicated by the asterisk notation), this verb does not govern the trace. The preposition is a "closer governor," defining PP as the minimal governing domain for the trace (see Chomsky [1986]). By Minimality, therefore, PP is a barrier to government from the more distant verb.
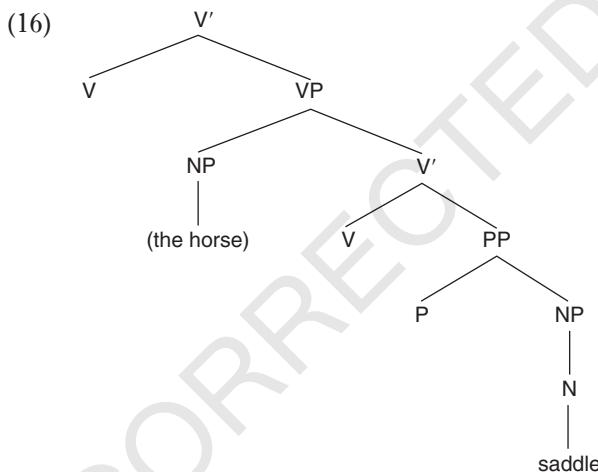
Minimality is also at issue in explaining why English lacks verbs of the hypothetical type in (14).

(14)   a.   *She churched her money.
              (cf. She gave a church her money.)
        b.   *He bushed a trim.
              (cf. He gave a bush a trim.)
        c.   *They housed a coat of paint.
              (cf. They gave a house a coat of paint.)

While *church*, *bush*, *house*, and many others exist as denominal verbs, they do not exist as denominal verbs having meanings comparable to those of the parenthetic sentences. While the verbs here may be impossible for a variety of reasons, there is a clear structural reason for their non-existence, on the assumption that their LRS representations would correspond to that depicted in (15) for the hypothetical *house* of (14c).

(15)

```
                        V′
                     ／     ＼
                    V        VP
                          ／      ＼
                        NP        V′
                        |       ／    ＼
                        N      V      PP
                        |            ／   ＼
                      house        P      NP
                                          |
                                      (a coat
                                      of paint)
```

The abstract P here would be a nonovert variant of the category appearing overtly in the expression *provide a house with a coat of paint*. This structure is in fact widely used in English – as (16) shows, it is the LRS representation for the large class of locatum verbs, like *saddle, blindfold, harness, shoe*.

(16)

```
                        V′
                     ／     ＼
                    V        VP
                          ／      ＼
                        NP        V′
                        |       ／    ＼
                   (the horse)  V      PP
                                     ／   ＼
                                    P      NP
                                           |
                                           N
                                           |
                                         saddle
```
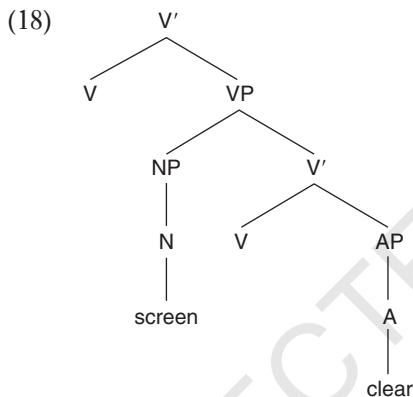
Thus, the verb *saddle* has a structure closely parallel to that of *provide* in *provide the horse with a saddle*. Here, the overt noun *saddle* may incorporate in cyclic fashion into each governing head in complete accordance with the principles of syntax.

But this structure cannot legitimately give rise to verbs of the type represented in (14). These would require incorporation from the "internal subject" position, that is, from Spec(ifier) of VP [(Spec, VP)]. Such incorporation would violate the ECP. Since the inner VP contains a governor (the V that heads it), that VP counts as the immediate governing domain in relation to the NP in its Spec. By Minimality, therefore, the inner VP is a barrier to government from the higher V. Movement of N from Spec position in the inner VP thus violates the ECP (see Baker and Hale 1990).

The same reasoning might explain why English also lacks verbs like those in (17).

(17)   a.   *She metaled flat.
              (cf. She flattened some metal.)
       b.   *He speared straight.
              (cf. He straightened a spear.)
       c.   *They screened clear.
              (cf. They cleared a screen.)

Again, these verbs exist, but not in the meanings indicated. Like the verbs of (14), those of (17) are ruled out by virtue of the ECP, on the view that their hypothetical lexical structures would correspond to that assigned to *screen* in (18).

(18)

This is essentially the structure of the analytic expressions *make a screen clear, get a screen clear*, and it is the source, by hypothesis, of the well-formed incorporation structure *clear a screen*, formed by successive-cyclic incorporation of the adjective *clear* in conformity with the principles of syntax. But this structure cannot be used to derive the verb of (17c), since that would require incorporation from the Spec position of the inner VP, violating the minimality requirement of the ECP.

In this discussion, we have maintained that certain verbal lexical items of English are derived through the operation of the head movement variant of Move α, that is, by incorporation. We have maintained that this is so because certain gaps in the lexicon can, we think, be explained on the assumption that the formation of the lexical items in question is subject to principles known to be operative in syntax. If this is true, it follows that the structures over which lexical derivations are defined are true syntactic structures, over which syntactic relations are defined. The final step in developing our position in regard to lexical representation is one we must simply assert. But we assert it with the belief that it is well supported by the kinds of linguistic material we have been considering. It is this: the notion "argument structure" is to be identified with the notion "lexical relational structure." Thus, the representation of the argument structure of a verb is a syntactic representation of the usual sort.

20.3   Questions pertaining to Hale and Keyser (1993)

1 What does Hale and Keyser's "syntax in the lexicon/l–syntax" have in common with LF (Logical Form)? To what extent are one or another or both redundant with respect to s–syntax/ordinary syntax?

2 How close is Hale and Keyser's lexical syntax to Chomsky's (1970) notion of the lexicon as a repository of idiosyncrasies?

3 How might one formulate the parameter distinguishing English (which allows verbs like *laugh*) from Basque (which requires nominal *laugh* plus an overt light verb)?

4 Contrary to the impression given by the preceding question, one can question whether English *laugh* can be a verb at all. If Hale and Keyser are right to take *They laughed* to involve incorporation of nominal *laugh* to a silent light verb, in what sense is or is not the resulting *laugh*+V constituent a V itself? (Hint: Bring in Chomsky's 1994 bare phrase structure.)

5 On page 96 of their paper, Hale and Keyser say that "most, probably all, superficially monomorphemic verbs are lexically phrasal." To what extent does Kayne's (2008: 13) statement that "all verbs are light verbs" differ from Hale and Keyser's formulation?

6 Hale and Keyser, like Baker (1988), take incorporation to be head movement. Now New York City has a borough called Brooklyn and another called the Bronx that differ in that English allows *several other Brooklyn-born linguists*, but not **several other the Bronx-born linguists*, even though the definite article is normally obligatory with *the Bronx*, as seen in *He lives in *(the) Bronx*. How might these facts follow from incorporation as head movement? (Extra credit: How close a link is there to the fact that English disallows plural *-s* in its OV compounds, e.g., *She's a real cat(*s) loving linguist*? To what extent is Massam's (2001) work on pseudo-incorporation relevant?)

7 The interpretation of *They shelved the book* is close to that of *They put the book on a shelf*, and that closeness is reflected in Hale and Keyser's analysis. There is almost certainly no variety of English in which the interpretation is, instead, close to that of *They took the book off a shelf*. Why might that be? How is *unshelve* relevant?

8 (Extra credit) How does the preceding question tie in with Kayne's (2006, section 7) brief discussion of the fact that *enough money* involves a silent MUCH (cf. Jackendoff 1977: 152) rather than a silent LITTLE? Is there any link to the fact that *We took them a book* corresponds to *We took a book to them* and not to *We took a book from them*?

9 (Lots of extra credit) How do (some) Scandinavian languages differ from English with respect to *take . . . from*, and what kind of parameter might be involved? What kind of parameter might be involved in American English *We prevented them from leaving* vs. British *We prevented them leaving* (cf. Landau 2002; Baltin 2009)? To what extent are the two parameters of this question similar to one another?

10   Hale and Keyser's analysis of *They shelved the book* involves a silent preposition plus incorporation of *shelf* (and the preposition) to the silent light verb. Having that silent preposition, but without any incorporation and with a pronounced verb would appear to lead to the impossible **They put the book a shelf*. How might one exclude such sentences?

11   If the preposition is overt, it cannot incorporate. Neither **They onput the book the shelf*, with a pronounced V, nor **They onshelved the book*, with a silent V, is possible. How might these be linked to the head movement vs. phrasal movement question?

12   There is a difference between the marginal*? He top-shelved the book* (cf. *He put the book on the top shelf*) vs. the sharply impossible **He sturdy-shelved the book* (cf. the well-formed *He put the book on a/the sturdy shelf*). What is the right descriptive generalization? Is there any connection to Cinque (2010) on adjectives? To *He had a good sleep* vs. *He good-slept?

13   Hale and Keyser show that incorporation in English (of the noun *bottle* to the silent light V) can't skip/strand the overt preposition in cases like *He bottled the wine (*in)*. Do you think that this restriction could be related to other restrictions on (English) preposition-stranding discussed in Kayne (1998, sect. 3.3)?

14   Hale and Keyser take sentences like *They gave John a horse* to contain a silent preposition akin to the pronounced preposition of *They provided J with a horse*. Is this idea compatible with Mulder's (1992) approach to the latter? How serious a problem is it for Hale and Keyser that **They gave John with a horse* is itself impossible?

15   Hale and Keyser note that Papago (Tohono O'odham) allows, with a pronounced causative verbal suffix *-cud*, a(n apparent) counterpart of **The grass sneezed the child*, whereas English allows only *The grass made the child sneeze*. Why might it be that English disallows this with a silent counterpart of *make*? How likely is it that there is a link to French *L'herbe a fait éternuer l'enfant* ('the grass has made sneeze the child') vs. **L'herbe a fait l'enfant éternuer* (cf. Kayne 1975; Rouveret and Vergnaud 1980; Burzio 1986)?

16   Hale and Keyser's example (14c) **They housed a coat of paint*, which is impossible in the sense of *They gave the/a house a coat of paint*, seems worse than*? Why don't you Red Cross the money you just inherited?* (in a sense close to that of *Why don't you give the Red Cross the money. . .?*). Is this example a counterexample to their proposal that incorporation of the subject of a double-object type small clause is impossible? To what extent might there be a link to sentences like *Why don't you Google it?*? Why is **Why don't you linguist the money you just inherited?* completely impossible?

17   Find three languages in which sentences like *Why don't you Google it?* are impossible, with *Google* looking like a verb. What parameter(s) might underlie the difference between those languages and English?

18   Pick two languages other than English that differ from English with respect to *They shelved the books* and discuss how one might try to account for those differences.

19 Discuss the degree of resemblance between Hale and Keyser's question "Why are there so few theta-roles?" and the question "Why do verbs have so few (nonadjunct) PP complements?"

20 Hale and Keyser take agents to be external arguments indirectly related by predication to their verb. How does this compare with the little *v* approach to agents (Kratzer 1996; Chomsky 1995)? How similar is the little *v* approach to Fillmore's (1968) deleted *by* and to Collins's (2005) use of *by*?

21 How might Hale and Keyser have attempted to integrate into their analysis nonagentive causers (cf. Pesetsky 1995)? Extra credit: What approach to nonagentive causers do you think could best account for contrasts such as *?the guy who they gave the sister of a copy of their paper* vs. **the guy who it* [the cold weather] *gave the sister of a bad case of the flu*? (Bring in Schäfer 2012.)

References

Arad, M. 2003. Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language and Linguistic Theory* 21: 737–778.

Baker, M. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Baker, M., and K. Hale. 1990. Relativized Minimality and pronoun incorporation. *Linguistic Inquiry* 21: 289–297.

Baltin, M. 2009. The properties of negative non-finite complements. Ms., New York University.

Brody, M. 2003. *Towards an Elegant Syntax*. London: Routledge.

Burzio, L. 1981. *Intransitive verbs and Italian auxiliaries*. Doctoral dissertation, MIT.

Burzio, L. 1986. *Italian Syntax: A Government–Binding Approach*. Dordrecht: Reidel.

Chomsky, N. 1970. Remarks on nominalization. In R. Jacobs and P. Rosenbaum (eds), *Readings in English Transformational Grammar*, Waltham, MA: Ginn, pp. 184–221.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 1994. *Bare Phrase Structure*. Cambridge, MA: MIT Occasional Papers in Linguistics 5. [Reprinted in Webelhuth (1995, Ch. 8).]

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Cinque, G. 2010. *The Syntax of Adjectives. A Comparative Study*. Cambridge, MA: MIT Press.

Clark, E., and H. Clark. 1979. When nouns surface as verbs. *Language* 55: 767–811.

Collins, C. 2005. A smuggling approach to the passive in English. *Syntax* 8: 81–120.

Di Sciullo, A. M., and C. Boeckx (eds). 2011. *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford Studies in Biolinguistics. Oxford and New York: Oxford University Press.

Fillmore, C. 1968. The case for case. In E. Bach and R. T. Harms (eds), *Universals in Linguistic Theory*, New York: Holt, Rinehart, and Winston, pp. 1–88.

Hale, K. 1989. The syntax of lexical word formation. In R. Carlson et al. (eds), *Proceedings of the Fourth Meeting of the Pacific Linguistics Conference*, Eugene: Department of Linguistics, University of Oregon, pp. 190–217.

Hale, K., and S. J. Keyser. 1991. *On the syntax of argument structure*. Lexicon Project Working Papers, Center for Cognitive Science, MIT.

Hale, K. L., and S. J. Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In K. L. Hale and S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press, pp. 53–109.

Halle, M., and A. Marantz. 1993. Distributed Morphology and the pieces of inflection. In K. L. Hale and S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press, pp. 111–176.

Jackendoff, R. 1977. *X-bar Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Julien, M. 2002. *Syntactic Heads and Word Formation*. New York and Oxford: Oxford University Press.

Kayne, R. S. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

Kayne, R. S. 1998. Overt vs. covert movement. *Syntax* 1: 128–191.

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2006. On parameters and on principles of pronunciation. In H. Broekhuis, N. Corver, R. Huibregts, U. Kleinhenz, and J. Koster (eds), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, Berlin and New York: Mouton de Gruyter, pp. 289–299. [Reprinted in Kayne (2010, Ch. 4).]

Kayne, R. S. 2008. Antisymmetry and the lexicon. In J. Rooryck and J. van Craenenbroeck (eds), *Linguistic Variation Yearbook* 8, Amsterdam and Philadelphia: John Benjamins, pp. 1–31. [Reprinted in Kayne (2010, Ch. 9) and in Di Sciullo and Boeckx (2011, Ch. 14).]

Kayne, R. S. 2010. *Comparisons and Ccontrasts*. New York and Oxford: Oxford University Press.

Kratzer, A. 1996. Severing the external argument from the verb. In J. Rooryck and L. Zaring (eds), *Phrase Structure and the Lexicon*, Dordrecht: Kluwer, pp. 109–137.

Landau, I. 2002. (Un)interpretable Neg in Comp. *Linguistic Inquiry* 33: 465–492.

Larson, R. 1988. On the double object construction. *Linguistic Inquiry* 19: 335–391.

Levin, B. 1991. English verb classes and alternations: A preliminary investigation. Ms., Northwestern University.

Levinson, L. 2007. *The roots of verbs*. Doctoral dissertation, New York University.

Marantz, A . 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In A. Dimitriadis, L. Siegel, C. Surek-Clark, and A. Williams (eds), *University of Pennsylvania Working Papers in Linguistics* 4, pp. 201–225.

Massam, D. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19: 153–197.

Mithun, M. 1984. The evolution of noun incorporation. *Language* 60: 845–895.

Mulder, R. 1992. *The aspectual nature of syntactic complementation*. Doctoral dissertation, Holland Institute of Linguistics.

Perlmutter, D. 1978. Impersonal passives and the unaccusative hypothesis. In J. Jaeger (ed.), *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society, pp. 157–189.

Pesetsky, D. 1995. *Zero Syntax: Experiencers and Cascades*. Cambridge, MA: MIT Press.

Pullum, G. K. 1988. Topic . . . comment: Citation etiquette beyond Thunderdome. *Natural Language and Linguistic Theory* 6: 579–588.

Rouveret, A., and J.-R. Vergnaud. 1980. Specifying reference to the subject: French causatives and conditions on representations. *Linguistic Inquiry* 11: 97–202.

Schäfer, F. 2012. Two types of external argument licensing – the case of causers. *Studia Linguistica* 66: 128–180.

Talmy, L. 1985. Lexicalization patterns. In T. Shopen (ed.), *Language Typology and Syntactic Description*, Cambridge: Cambridge University Press.

Travis, L. 1984. *Parameters and effects of word order variation*. Doctoral dissertation, MIT.

Walinska de Hackbeil, H. 1986. *The roots of phrase structure: The syntactic basis of English morphology*. Doctoral dissertation, University of Washington, Seattle.

Walinska de Hackbeil, H. 1989. *Theta-government, thematic government, and extraction asymmetries in zero derivation*. Report CS-R8918 April. Amsterdam: Center for Mathematics and Computer Science.

Webelhuth, G. 1995. *Government and Binding Theory and the Minimalist Program*. Oxford: Blackwell.

21

Reference and Proper Names: A Theory of N-movement in Syntax and Logical Form

Giuseppe Longobardi

1994

21.1 Introduction

Like all fields of scientific inquiry, generative syntax aims to find an explanation for certain empirical observations; it develops hypotheses to account for them, tests them against more data, evaluates them carefully, and then discards them or refines them. Since the mid 1950s, it has amassed a core of empirical observations (incorporating those made by previous traditions) and developed a set of theoretical proposals that, though still open to refinements in the details of their implementation, constitute a core around which the field has been growing. This can be seen very clearly in the literature on the syntax of the noun phrase, as in this article.

Based on data from Hungarian, Szabolcsi (1983/84, 1994) (the latter excerpted in Ch. 22 of this volume) expressed the intuition that noun phrases consist of an NP that is the complement of a functional head. This idea, also independently developed in Abney (1987), has come to be known as the "DP hypothesis." This initial body of work has stimulated a fruitful line of research in a variety of different languages, which has gathered relevant data and extrapolated evidence leading to a deeper understanding of the internal syntax of the noun phrase (cf. Delsing 1993; Bernstein 1991b, 1993; Taraldsen 1990; among others). In this paper, Longobardi further strengthens and refines the DP hypothesis based on evidence from Italian, comparing noun phrases that lack an overt determiner with those that have one.

Longobardi notes that, when a noun phrase that serves as an argument of a predicate is headed by a singular count noun, it requires an overt determiner; however, when it is not an argument (but a vocative, or a predicate), it does not:

(1) a. *(Un/Il) grande amico di Maria mi ha telefonato. (Italian)
         (a/the) great friend   of Maria me has called
         Intended meaning: 'A/the great friend of Mary's called me.'

     b. *Ho incontrato (un/il) grande amico di Maria ieri.
         have met   (a/the) great   friend of Maria yesterday
         Intended meaning: 'I met a/the great friend of Mary's yesterday.'

     c. Ti   credevo amico di Maria.
         you believed friend of Mary
         'I believed you a friend of Mary's.'

In contrast, when a noun phrase is headed by a mass noun or a plural count noun, it may lack an overt determiner, even if it serves as an argument, provided that it is the complement of a lexical head, as in (2). In such cases, it can only have an indefinite (often existential) interpretation:

(2) a. Bevo sempre vino.
         drink always   wine
         'I always drink wine.'

     b. Mangio patate.
         eat   potatoes
         'I eat/am eating potatoes.'

Since the distributional restriction to the complement of a lexical head is characteristic of a certain class of null elements, Longobardi takes this observation as evidence that the noun phrase does not lack a DP layer, but rather contains a D that is phonetically null (cf. Contreras 1986). He further argues that it is this null D that determines the interpretation of the noun phrase. Moreover, since this distributional restriction only applies when the noun phrase is an argument, Longobardi concludes that noun phrases that serve as arguments and those that serve as predicates have different requirements: argument noun phrases must be DPs, other noun phrases may be NPs or DPs (cf. Szabolcsi 1987; Stowell 1989).

The proposal that argument noun phrases must be DPs raises the issue of how to analyze proper names. In some languages, including Italian, they may be accompanied by a definite article (analyzed as a D), which has led to viewing them as the head of the NP (cf. Burge 1973). Yet, when they are not accompanied by an article, proper names do not exhibit the same properties as common nouns that lack an overt determiner: even when used as arguments, they are not restricted to the complement of a lexical head, and they do not have an indefinite existential interpretation. Longobardi reconciles these observations with the proposal that arguments must be DPs by arguing that noun phrases with proper names do not have a null D: they either have an overt D, or else the proper name raises from N to D. This proposal finds empirical support in Italian, where a proper name may co-occur with a determiner and follow an adjective, as in (3a), or else it may not have a determiner and precede the adjective, as in (3c); but it cannot at once fail to co-occur with a determiner and follow the adjective:

(3) a.   Il mio Gianni ha finalmente telefonato.
           the my Gianni has finally   called
     b.   *Mio Gianni ha finalmente telefonato.
     c.   Gianni mio ha finalmente telefonato.

This is because, where there is no determiner, the N raises to D.

Longobardi also compares Italian and English (and more generally Romance and Germanic languages) with regard to N–to–D raising, and its interaction with interpretive properties (referential, generic, existential) in different kinds of nominals (mass/count, singular/plural, common nouns/proper names). Concerning proper names, he notes that in English they do not raise to D, and yet do not exhibit the same distributional and interpretive restrictions that characterize nouns with a null determiner in Italian. This leads him to conclude that a null D in English is not necessarily associated with an indefinite interpretation, a property that can explain the existence of bare nouns denoting kinds (e.g., "Dinosaurs are extinct").

In sum, this article lays much of the groundwork for a parametric account of correlations between the internal syntax of argument noun phrases, their distribution, and the restrictions on their interpretation.

21.2   From "REFERENCE AND PROPER NAMES: A THEORY OF N-MOVEMENT IN SYNTAX AND LOGICAL FORM"

In recent years, formal syntactic theory has broadened its scope and has come to interact more and more closely with parallel domains of study, in particular with such well-established traditions of inquiry as comparative dialectology, language typology, and analytic philosophy, fruitfully exchanging insights and research techniques. As a result, it has become possible to raise and solve new, meaningful problems, which would have hardly been conceivable as recently as twenty years ago, and to sharpen the formulation of more traditional questions so as to provide them with adequate empirical answers.

Within such an enlarged framework of interests and methods, this article will consider evidence from Romance and Germanic suggesting the following theoretical conclusions:

- There exist instances of N-movement to D in the syntax of Western Romance, implying the correctness for such languages of the so-called DP analysis.
- The same type of movement is likely to take place only in LF in English and German.
- Head-to-head dependencies fall into essentially the same categories as those between maximal projections: they define chains or CHAINS (in Chomsky's ([1986]) terms), and chains are created either by substitution or by adjunction, with distinct properties.
- Various semantic types of articleless nominals (proper names, existentials, definite and indefinite generics, nonargument nominal phrases) are

distinguishable by their syntactic behavior at PF and LF, and a plausible theory of the semantic licensing of NPs and DPs can be envisaged: proper names and generics are thus also distinguished from definite descriptions, suggesting a possible syntactic answer (in the spirit of Kripke 1980 or Neale 1990) to long-standing philosophical questions.

• The definite article of many European languages can be shown to have two different functions, a substantive and an expletive one, a distinction morpho-logically manifested in some varieties.

• Finally, some empirical advantages of adopting certain basic tenets of Chomsky's (1993) minimalist program for explaining the generalizations reached in this article are suggested in the appendix.

1   DPs and NPs

To begin with, consider that in light of the generalization of X-bar theory to all lexical and nonlexical categories, two positions have recently emerged about the structure to be assigned to projections of determiners: one view locates Determiner Phrases inside Noun Phrases, in particular in their specifier position; the other, originally stemming from an intuition of Szabolcsi (1983/84 and subsequent work), conceives of the whole nominal construction as coinciding with DP and of NP as a complement of the head D (see in particular Abney 1986, 1987). Schematically, the two hypotheses can best be summarized as in (1) and (2), respectively.

(1) $[_{NP} DP [_{N'} N]]$
(2) $[_{DP}[_{D'} D NP]]$

Although the problem of choosing between the two views has proved not to be easy to solve on empirical grounds, one line of argument in favor of the structure advocated by Szabolcsi and Abney appears to be especially promising and has been explored in order to try to decide the issue conclusively in certain languages. Consider that, if movement can be argued to apply in some language from inside NP to a position inside DP (e.g., from specifier position to specifier position or from the position of N^0 to that of D^0), then the structure in (1) will immediately be discarded, under any current theoretical approach, by the ban against movement to a non-c-commanding position. Movement from Spec NP to Spec DP might be instantiated in English, if the pair in (3) is to be related transformationally.

(3)   a.   a very strange man
      b.   how strange a man

Head-to-head movement from N^0 to a higher functional head that may, at least in some cases, be identified with D^0 has been tentatively argued to apply in Semitic (see Ritter 1986, 1989, Ouhalla 1988, Fassi Fehri 1989, Siloni 1989, 1996, and references cited there) and Scandinavian (see Delsing 1988, Taraldsen 1990,

Holmberg 1992). Taraldsen (1990), for instance, analyzes the following Norwegian paradigm in terms of N-raising:

(4)   a.   hans bøker om   syntaks
           his   books about syntax
      b.   bøkene   hans om   syntaks
           book-s-the his   about syntax

(4a) shows the normal SNO structure of Germanic NPs (see Giorgi and Longobardi 1991), where the subject can be independently argued to asymmetrically c-command the object. (4b) instantiates an alternative N-initial order in which the subject can still be shown to asymmetrically c-command the object; thus, Taraldsen rejects the possibility of its being base-generated and proposes to derive it from (4a) by means of N-raising, thus supporting a DP analysis-for Norwegian nominals. In fact, the head N appears in (4b) to be morphologically adjoined to the article.

   It may also be possible to extend Taraldsen's analysis to all cases of so-called suffixed articles in Scandinavian, accounting for such common alternations as the following:

(5)   a.   en bok
           a   book
      b.   boken
           book-the

A similar approach was also successfully taken in the study of suffixed definite articles in Romanian (Grosu 1988, Dobrovie-Sorin 1987).

   However, evidence of this sort, as well as of other types discussed particularly by Abney (1987), can hardly be reproduced in the Western Romance languages (but see Bernstein 1991a), for which the choice between (1) and (2) has so far remained more undetermined (although the DP analysis has occasionally been employed to treat aspects of Romance nominal syntax; see, e.g., Torrego 1988, Battye 1989, Brito 1990). In what follows we will examine evidence of a completely different nature, suggesting even more directly that instances of N-to-D movement must be postulated in Western Romance as well and thus further supporting the structure in (2) and the theory of head movement. Before looking at this evidence, however, we must analyze certain semantic and distributional properties of Ns and Ds.

2   Bare nouns

Let us begin by noticing that a singular countable head noun may not occur in Italian in any of the major positions suitable for arguments (e.g., subject, direct object, prepositional object, inverted subject of either ergative or unergative

predicates) without being introduced by an overt determiner, most usually a definite or indefinite article, a quantifier, or a demonstrative.

(6) a.   *(Un/Il) grande amico di Maria mi ha telefonato.
          (a/the) great friend   of Maria called me up
     b.   Ho incontrato *(un/il) grande amico di Maria ieri.
          I met   (a/the) great friend   of Maria yesterday
     c.   Ho parlato con *(un/il) grande amico di Maria ieri.
          I spoke   with (a/the) great friend   of Maria yesterday
     d.   Ha telefonato/E' venuto *(un/il) grande amico di Maria.
          called up/came   (a/the) great friend   of Maria

   The constraint in question does not hold for nominals in typical nonargument function, as in vocative, predicative, or exclamatory contexts.

(7) a.   Caro amico, vieni a trovarmi.
          dear friend   come to visit me
     b.   Tenente, esegua   l'ordine!
          lieutenant perform the command

(8) a.   Gianni è tenente.
          Gianni is lieutenant
     b.   Gianni è amico di Maria.
          Gianni is friend of Maria
     c.   L'ho promosso tenente.
          I promoted him lieutenant
     d.   Ti credevo   amico di Maria.
          I believed you friend of Maria

(9) a.   Diavolo!
          devil
     b.   Maledetto tenente!
          damn'   lieutenant

   There are also some kinds of PPs that allow articleless singular nouns; but, pending further study, it is not implausible to assimilate them to predicative expressions on semantic grounds.

(10) a.   in abito lungo
           in long dress
      b.   di buona famiglia
           of good   family

   On the basis of these observations, we may tentatively propose the following principle of Italian grammar:

(11)   A "nominal expression" is an argument only if it is introduced by a
        lexically filled D position.

Although plausible and basically correct in spirit, (11) has two shortcomings,
one conceptual and one empirical. First, reference to the lexical (i.e., phonetic)
content of the category D seems to be inappropriate and unparalleled in an essen-
tially semantic licensing condition. Second, (11) appears to be simply too strong
under this formulation. In fact, as noted also by Benincà (1980), three types of
bare nouns occur in Italian in argument function: singular mass nouns, plural
count nouns (bare plurals), and even some rarer cases of singular count nouns in
the scope of a sentential negation, although it is not clear whether the latter are to
be considered real arguments or rather quasi-idiomatic expressions. Consider the
following examples:

(12)   a.   Bevo sempre   vino.
              I always drink wine
        b.   Mangio   patate.
              I eat/am eating potatoes
        c.   Non c'era   studente in giro. (from Benincà 1980)
              there wasn't student   around

In all these cases the interpretation of the nominal seems to be roughly similar to
that of an indefinite, existentially quantified NP. In this sense, bare nouns appear
to bear some semantic similarity to the so-called partitive article (formed by *di*
'of' + a definite determiner) of Italian (and French); also the latter in fact is
limited to mass head nouns and to plurals, for which it seems to represent the
intuitive counterpart of the singular indefinite article.

     There are also some differences, however, which limit the analogy and prevent
us from simply stating that bare nouns instantiate the phonetically "null" version
of the partitive article. An interesting peculiarity, for example, is that the number
specification (i.e., the semantic distinction between singular and plural) may some-
times be irrelevant. Number is obviously irrelevant in the case of mass nouns and
of negated existentials, which have null designation, but as convincingly argued by
Benincà (1980), bare plurals are often allowed to be neutral between the singular/
plural distinction. For example, consider the following paradigm:

(13)   a.   Ogni giorno mangia patate.
              every day   he eats potatoes
        b.   Ogni giorno mangia alcune/delle patate.
              every day   he eats some/PARTIT ART potatoes

In uttering (13b), which contains an overt existential quantifier or the partitive
indefinite article (here formed by contraction of *di* 'of' + *le* 'the' fem. pl.), we com-
mit ourselves to the claim that the person in question eats more than a single
potato per day, whereas in uttering (13a), we make no such plurality commitment.
Another peculiarity concerns scopal phenomena: unlike all overt existential

determiners, including the singular indefinite article and the partitive one, determinerless nominals in both English and Italian are subject to an obligatory narrow scope constraint. This applies with respect to negation, quantifiers, and intensional contexts (thus producing a necessarily opaque or *de dicto* reading), as discussed by Carlson (1977a,b) precisely in order to distinguish between the indefinite article and bare plurals.

Therefore, a plausible observational generalization appears to be that existential quantification can be expressed using a bare noun under certain special conditions. Now, is this existential interpretation of Italian bare nouns the consequence of an absolute lack of the category "Determiner" in these constructions or is it assigned as the default semantic option to an empty category syntactically present in the D position? One fact suggests the plausibility of the latter solution: the distribution of such bare nouns in Italian, as well as in other Romance languages, seems to be subject to a sort of lexical government requirement, similar to that constraining empty categories in general and empty functional heads in particular (e.g., empty Cs of finite clauses in English; see Stowell 1981). In other words, Romance bare nouns are usually excluded from preverbal subject position, but admitted in internal argument position and, to a certain extent, also as inverted subjects of unergative predicates.

(14) a. *Acqua viene giù   dalle colline.
         water   comes down from the hills
     b. Viene giù   acqua dalle colline.
         comes down water from the hills
     c. Ho preso acqua dalla sorgente.
         I took   water from the spring

(15) a. *In questo ufficio marocchini telefonano sempre.
         in this office   Moroccans always call up
     b. In questo ufficio telefonano sempre marocchini. (from Brugger 1990)
         in this office   always call up   Moroccans
     c. In questo ufficio incontro sempre marocchini.
         in this office   I always meet   Moroccans

Similarly impossible is a determinerless noun in another position analyzed at length and argued to be non-lexically governed in Longobardi 1980, [1994], that of postcopular argument expressions (also see Higgins 1973, Ruwet 1975, and Moro 1993).

(16) *La causa delle rivolte sono spesso marocchini.the cause of the riots are often Moroccans

On the contrary, no violation arises if a nonargument expression such as a predicative NP, even with a singular count head, occurs in a non-lexically governed position, as is shown by the acceptability of (17), whose relevance was originally pointed out by Luigi Burzio (personal communication).

(17)   Amico di Maria sembra essere Gianni.
       friend of Maria seems to be   Gianni

These observations may be taken to suggest that an empty category in need of some kind of lexical government is necessarily present in (12)–(16) but not in (17); if such a category is actually a head D, its presence may also suffice to explain the otherwise unmotivated restrictions to plural/mass nouns and to the existential reading, which do not arise in the case of sentences like (17) or other nonargument (e.g., vocative) usages. In fact, the empty D could instantiate some sort of existential operator and as such impose constraints regarding the count/mass interpretation of the head nouns it quantifies over (see section 5 for discussion): the behavior of certain overt existential determiners, such as the above-mentioned partitive article, is, after all, analogous.

On the basis of these observations, and abstracting away from the marginal and peculiar cases like (12c), we will make the following assumptions:

(18)   Empty determiners may occur at S-Structure in Italian only under the
       following conditions:
       a.   They are restricted to plural or mass head nouns like several other
            determiners.
       b.   They are subject to a lexical government requirement like other empty
            heads.
       c.   They receive an indefinite interpretation corresponding to an existen-
            tial quantifier unspecified for number and taking the narrowest
            possible scope (default existential).

Further support for the existence of an empty D is provided by the following considerations. According to the analysis in Longobardi 1980, [1994], the postcopular predicative position is always lexically governed, so it could in principle contain a predicate nominal introduced by an empty determiner. In fact, even if determiner-less predicative expressions are exempted from the need for a phonetically null D, there seems to be some evidence that they can be introduced by such an empty category at least with mass and plural heads.

(19)   a.   Gianni è medico.
            Gianni is doctor
       b.   Gianni è un medico.
            Gianni is a doctor
       c.   *Gianni è medico che si cura davvero dei suoi pazienti.
             Gianni is doctor   who really   cares for his patients
       d.   Gianni è un medico che si cura davvero dei suoi pazienti.
             Gianni is a doctor   who really cares for his patients
       e.   Noi siamo medici che ci curiamo davvero dei nostri pazienti.
             we are   doctors who really care for our patients
       f.   Noi siamo dei medici   che ci curiamo davvero dei nostri pazienti.
             we are   PARTIT ART doctors who really care for our patients

g.   Questa è acqua.
     this is   water
h.   Questa è dell'acqua.
     this is   PARTIT ART water
i.   Questa è acqua che è stata presa dalla sorgente.
     this is   water which was taken from the spring
j.   Questa è dell'acqua   che è stata presa dalla sorgente.
     this is   PARTIT ART water which was taken from the spring

The fact that relativization on a predicative head is only possible either with an overt determiner or with a plural/mass noun may suggest that the presence of a D position (subject to generalization (18a), if empty) is required in order to license a relative clause. If this line of reasoning is correct, (19c) will be ruled out since an empty D with a nonmass singular would violate (18a); (19e) and (19i) will certainly contain such a null determiner, and (19 g) will contain it optionally. Similar conclusions about the possibility of empty Ds with predicates can be drawn from the following sentences:

(20)   a.   Ritengo Mario *(un) bravo medico.
            I believe Mario   (a) good doctor
       b.   Ritengo Gianni e Mario   (dei) bravi medici.
            I believe Gianni and Mario (PARTIT ART) good doctors

    With certain adjectivally modified predicates headed by count nouns, an overt determiner is sometimes obligatory in the singular, but not in the plural, suggesting that a D category may always be required, remaining empty just in the plural, as expected given (18a).
    The assumptions made in (18) finally allow us to revise (11) as follows:

(21)   A "nominal expression" is an argument only if it is introduced by a
       category D.

It is obvious how (21) also overcomes the conceptual shortcoming of (11) noted earlier, by eliminating reference to the content of the D position.
    The crucial nature of the category D for argumenthood seems to be independently manifested by the fact that certain designation properties typical of arguments, such as the semantic import of grammatical number, lie precisely in the D position. Actually, we have already observed that an empty D (therefore, one morphologically unspecified for number) may yield semantic indeterminacy between singular and plural designation despite the plurality of the head noun. But stronger evidence for this point is provided by pairs of subject phrases like the following:

(22)   a.   La mia segretaria e tua collaboratrice   sta/*stanno uscendo.
            the my secretary and your collaborator is/are going out
       b.   La mia segretaria e la tua collaboratrice   stanno/*sta uscendo.
            the my secretary and the your collaborator are/is going out

In (22a) two morphologically singular nominal projections are coordinated, excluding the determiner, which remains unique and is also morphologically singular; here the whole subject argument of the clause is understood as designating a single individual, as is clarified by the verbal agreement. By contrast, in (22b) the coordination includes the determiners, one for each conjunct, and the designation of the argument is obligatorily understood as plural. In other words, irrespective of the cardinality of head nouns present, a single singular determiner is sufficient to impose singular designation on the entire nominal expression, whereas the sum of two singular determiners automatically imposes plural designation. When occurring as a predicate, however, the phrase in (22b) is not required to be interpreted as plural.

(23)   Maria è la mia segretaria e la tua collaboratrice.
        Maria is the my secretary and the your collaborator

Thus, arguments but not predicates crucially rely on the D position in order to define their meaning with respect to number.

3 Proper names

If it is really the D position that turns a nominal expression into an argument, an obvious question arises concerning those proper names (in particular, names of individuals, cities, certain "small" islands, companies, days, and months) that are allowed in Italian to occur freely in argument function without any determiner; it is rather clear that they cannot be introduced by an empty D, since the properties of the latter, as identified in (18), seem to be inapplicable in the case of proper names. For example, a proper name like *Gianni* in (24a) is not understood as designating a mass, is not plural, does not receive an indefinite interpretation, and may occur in a lexically ungoverned position.[16]

The theoretical framework so far defined provides a restrictive and almost inescapable answer to this problem: a D position introducing the subject argument must be syntactically present in a sentence like (24a) and cannot be empty; thus, the only possible element that can occupy such a D position is the proper name itself. Yet another way to formulate essentially the same problem involves recalling that several varieties of Romance display free or stylistically conditioned alternations between the presence and the absence of the article with proper (first or last) names of human beings.

(24)   a.   Gianni mi ha telefonato.
              Gianni called me up
        b.   Il Gianni   mi ha telefonato.
              the Gianni called me up

In some cases the alternation is also semantically conditioned; for instance, with last names of female human beings the article in standard Italian is virtually obligatory.

(25)   La Callas/*Callas ha cantato.
        the Callas/Callas sang

The natural question that arises here, although it has so far never been raised, is whether *Gianni* in (24a) occupies the same S-Structure position as *Gianni* in (24b) or rather the position of *il* in (24b). As noted earlier, the assumptions motivated in section 2 force us to adopt the latter hypothesis and suggest the existence of a transformational relation between the pair of sentences in (24), established through movement of *Gianni* in (24a). In fact, now it becomes necessary to assume that such Ns as those proper names that occur in argument function without any overt determiner have undergone raising from N^0 to D^0, in order for the structure to comply with (18) and (21). This is so because they must be base-generated in the N^0 position and optionally allowed to remain there, to account for those cases in which they occur introduced by an article. This hypothesis, put forth on theoretical grounds, turns out to receive straightforward empirical confirmation from a curious and subtle paradigm of certain varieties of Romance, which it contributes to explaining.

   [. . .]

First, notice that Italian adjectives, both possessive and nonpossessive, may occur in prenominal position between D and N, or in postnominal position, but never before D with either common or proper nouns.

(27)   a.   *mio il   Gianni
               my   the Gianni
        b.   *vecchio il   tavolo
               old   the table

Now consider the following paradigm:

(28)   a.   Il mio Gianni   ha finalmente telefonato.
               the my Gianni finally called up
        b.   *Mio Gianni ha finalmente telefonato.
               my Gianni   finally called up
        c.   Gianni mio ha finalmente telefonato.
               Gianni my   finally called up
        d.   Il Gianni mio ha finalmente telefonato.
               the Gianni my finally called up

The two surface order possibilities (A N and N A) are preserved when the proper name, here in a typical argument position, is introduced by the determiner, but an unexpected gap in the paradigm appears with articleless names: although many varieties, especially in central and southern Italy, accept (28c), none accepts the severely ungrammatical (28b). The generalization appears to be that the lack of the article forces an N-initial order. This otherwise surprising idiosyncrasy immediately becomes understandable assuming that the proper name needs to move from N^0 in order to fill in the empty D^0 position, thus crossing over the adjective

presumably lying in its specifier position. There is also an interesting piece of semantic evidence in favor of this hypothesis – in particular, in favor of the assumption that the possessive AP of (28c) does not follow the N^0 position and has become postnominal only as the result of an N-preposing process. Normal postnominal possessives tend to be strongly contrastive in Italian, as is the case for *mio* in (28d), which can only be interpreted with contrastive reference to the existence of another salient Gianni in the domain of discourse who is not "mine" (i.e., is related to someone-else [sic]). This interpretation is not required by prenominal possessives, like the one in (28a), which can be perfectly understood as an affective expression in an environment where no other Gianni's existence is presupposed. Now, the interpretation of *mio* in (28c) does not need to be contrastive, exactly like that in (28a) and contrary to that in (28d). This may be explained on the basis of the general fact that contrastiveness is uniformly required of posthead possessives but not of those in Spec NP and of the crucial hypothesis that it is *Gianni* that moved in (28c), crossing over *mio*.

[. . .]

The evidence of this section thus argues for the existence of N-movement to D, and consequently in favor of the structure (2), in Italian and probably in other Western Romance languages as well. Once we adopt (2), the natural way of reformulating the content of principle (21) becomes the following, as also pointed out by Stowell (1989):

(21′)   DP can be an argument, NP cannot.

Consider in this light the syntactic licensing of NPs. Having distinguished between NPs and DPs and having argued for the structure in (2), we must now provide for the licensing of such categories under Chomsky's ([1986]) principle of Full Interpretation. According to Chomsky, who follows Rothstein (1983) (also see Rothstein 1990), maximal projections, apart from operators, can be licensed either as arguments or as predicates. DP can certainly be licensed as an argument in most cases, as we have seen, or as a predicate in others (e.g., many copular or small clause constructions). By contrast, NP was shown not to be able to assume argument function unless introduced by an overt or empty determiner (i.e., unless the complement of a D position). This fact suggests that in a structured utterance (i.e., unless used in isolation, as in vocative and exclamatory expressions) NP can only be licensed through a predicative interpretation. I propose, then, to enlarge the notion of predication so that NP can be syntactically predicated of the head selecting it, namely, of a D.

Note

16 Another sharp semantic difference between bare (common) nouns and proper names arises in the domain of scope facts. I have briefly mentioned Carlson's observation, reproducible in Italian, that bare nouns are forced to take the narrowest possible scope, in particular with respect to negation and intensional contexts (i.e., in traditional terms, they are read *de dicto*).

(i) a. Non ho incontrato studenti.
         I did not meet students
         'There are no students such that I met them.'
         '*There are some students such that I did not meet them.'
     b. Vorrei incontrare studenti.
         I would like to meet students
         'I would like for there to be some students such that I could meet them.'
         '*There are some students such that I would like to meet them.'

In the same contexts, however, proper names do not obey any narrow scope restriction: actually, they always display a scope–insensitive (essentially *de re*) existential reading. The actual existence of Maria seems in fact to be implied by the utterance of either sentence of (ii).

(ii) a. Non ho incontrato Maria.
          I did not meet Maria.
      b. Vorrei incontrare Maria.
          I would like to meet Maria

For further remarks on this crucial property of proper names, see section 5 [of the full article].

21.3 Questions pertaining to Longobardi (1994)

1 How must Longobardi's proposal that N moves to D in certain languages be interpreted from a (subsequent) bare phrase structure perspective, that is to say, what is the proper interpretation of "moving to D"? What exactly excludes the "moving into" interpretation? (Extra credit: How might Longobardi try to adapt his proposal that proper names do move into D (by substitution) to Chomsky's bare phrase structure?)

2 Discuss the possible significance for Longobardi of Perlmutter's (1970) argument that definite and indefinite articles do not belong to the same category.

3 Longobardi proposes that Italian bare plurals/mass nouns contain an unpronounced determiner. In what way is Kayne's (2008b) proposal similar to and yet different from Longobardi's?

4 In the general case, Italian disallows bare singular (common) count nouns, a fact that Kayne (2008b, (112)) suggests relating in effect to the absence of a filled specifier. What might be the reason(s) for Brazilian Portuguese allowing such bare singulars, as discussed by Schmitt and Munn (2002)?

5 Longobardi has Italian proper names like *Gianni* raising via head movement to D in cases where no visible D is present, in order to account for such proper names being able to occur without a visible D in Italian much more readily than common nouns. Discuss the significance of the fact that the same is true in Italian of full proper names like *Gianni Bruni* that contain both a given name (*Gianni*) and a family name (*Bruni*).

6   Longobardi presents strong evidence for raising of proper names (across a prenominal possessive) in Central and Southern varieties of Italian, on the basis of phrases like *Gianni mio* ('John my'). How might he try to fit in the Northern varieties of Italian that do not admit such phrases?

7   A parallel derivation is suggested by Longobardi for phrases such as *Napoleone terzo* ('Napoleon third'), with the proper name *Napoleone* raising across the ordinal *terzo*. Discuss the implications of the fact that the corresponding English phrase obligatorily has the definite article, as in *Napoleon *(the) third*. (Extra credit: Bring into the discussion *last Thursday*, *?Thursday last* and **the Thursday last*.)

8   In many Northern Italian dialects, a feminine first name (but not a masculine one) is obligatorily preceded by the definite article (*la Maria* vs. *Gianni*). To what extent might that contrast be linkable to the following contrast from English: *John lives in (*the) Cuba* vs. *John lives in *(the) West Indies*?

9   In agreement with Stowell (1989), Longobardi suggests that DP can be an argument, while NP cannot be one. How does this idea interact with Rosenbaum's (1967) analysis of sentential complementation?

10 Longobardi takes the presence of the prenominal adjective in cases like *I love sweet France* to indicate that no head-raising to D has taken place. Discuss the feasibility of taking the adjective in such cases in English to have raised to D (or to Spec,DP). (Extra credit: Bring in Romanian and/or Bulgarian.)

11 Longobardi attributes the difference between *Nobody swims in the Hudson anymore* and **Nobody goes to the Paris anymore* to the difference between *the Hudson river* and **the Paris city*. This linkage seems correct. Yet it leaves open the question of the relevance of *the city of Paris*. How might one reconcile the impossibility of **the Paris* with the possibility of *the city of Paris*?

12 The contrast noted by Longobardi between *I met two Susans yesterday* and **I met two hers/shes yesterday* certainly indicates, as he argues, a difference in status between proper names and pronouns. Discuss the feasibility of relating this difference to that between *I met two people/women named Susan yesterday* and **I met two people/women named her/she yesterday*. (Extra credit: If extended to definite singular proper names, in the spirit of the Appendix of Kayne 2007, how might this kind of analysis support Russell's 1905 view of proper names as disguised descriptions?)

13 Longobardi, following Postal (1966, Ch. 1 of this volume), takes pronouns to be base-generated in D (i.e., in bare phrase structure terms, to be merged parallel to other Ds). Discuss the possible importance of the fact that Maori (cf. Harlow 2001: 35; Bauer 1993) has both pronouns (especially singular pronouns) and proper names preceded by a special "article."

14 Longobardi takes empty determiners to have default existential interpretation. Discuss the importance, in this context, of Déprez's (2005) discussion of bare nouns that in some languages can have a definite interpretation.

15 Longobardi suggests that English has N substitution into D at LF (in the case of proper names). In a very different set of cases, Kayne (1998) later argued that LF movement should be replaced by overt movement (cf. Chomsky 2001: 15). How damaging would it be for Longobardi's idea that proper

names are interpreted in D, if there is in fact no LF movement at all? Justify your answer.

16 If there is no LF movement, then Longobardi's account of English *The rich are becoming even richer* vs. **Rich are becoming even richer* (in terms of LF adjective raising not being available) must be replaced by another. Discuss the plausibility of an account based on Kayne (2006), whereby what is at issue in these examples is the licensing of the silent NP (PEOPLE), which requires the availability of a specifier position whose head (*the*) must now arguably not itself be silent.

17 According to Longobardi (and other authors he cites) English is unusual, even among the Germanic languages, in allowing sentences like *All of a sudden, a dog appeared on my lawn*. Discuss the plausibility of relating this property of English to the fact that colloquial English hardly allows *All of a sudden, there appeared a dog on my lawn*. (Extra credit: Bring in Kayne 2008a on *there*.)

18 (Extra credit) English sentences with plural generics, such as *Cats are cute*, disallow the definite article, that is to say, the following cannot be interpreted in the same generic way: *The cats are cute*. Yet *The damn things are cute* can be. Why might that be?

19 Longobardi takes the definite article in *la Maria* (cf. question 8 above) in Italian to be 'expletive'. Discuss the implications of the possible close similarity to the definite article in English *the woman (who is) named Mary*. (Extra credit: Bring in Longobardi's discussion of Catalan and Frisian, along with question 13 above.)

20 Longobardi mentions a possible correlation, citing Gerhard Brugger (personal communication), between the limited use in English of what he takes to be expletive definite articles (on proper names and generics, compared with other Germanic languages), on the one hand, and the lack of number and gender agreement on English definite articles, on the other. Test this correlation by looking into Afrikaans.

References

Abney, S. 1986. *Functional elements and licensing*. Talk presented at GLOW. Available at http://www.vinartus.net/spa/86b.pdf. [Accessed January 2013.]

Abney, S. 1987. *The English noun phrase in its sentential aspect*. Doctoral dissertation, MIT.

Battye, A. 1989. Putting some Romance into the DP-analysis. Ms., University of York.

Bauer, W. A. 1993. *Maori*. London: Routledge.

Benincà, P. 1980. Nomi senza articolo. *Rivista di grammatica generativa* 5: 51–63.

Bernstein, J. 1991a. Nominal enclitics in Romance. In *MIT Working Papers in Linguistics* 14: *Papers from the Third Student Conference in Linguistics*, Cambridge, MA: Department of Linguistics and Philosophy, MIT, pp. 51–66.

Bernstein, J. 1991b. DP's in French and Walloon: Evidence for parametric variation in nominal head movement. *Probus* 3: 101–126.

Bernstein, J. 1993. *Topics in the syntax of nominal structure across Romance*. Doctoral dissertation, CUNY.

Brito, A. M. 1990. Nominal specifiers in European Portuguese. Ms., University of Porto.

Burge, T. 1973. Referents and proper names. *The Journal of Philosophy* 70: 425–439.

Carlson, G. 1977a. *Reference to kinds in English*. Doctoral dissertation, University of Massachusetts, Amherst.

Carlson, G. 1977b. A unified analysis of English bare plurals. *Linguistics and Philosophy* 1: 413–457.

Chomsky, N. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.

Chomsky, N. 1993. A minimalist program for linguistic theory. In K. Hale and S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press, pp. 1–52.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, pp. 1–52.

Contreras, H. 1986. Spanish bare NPs and the ECP. In I. Bordelois, H. Contreras, and K. Zagona (eds), *Generative Studies in Spanish Syntax*, Dordrecht: Foris, pp. 25–49.

Delsing, L.-O. 1988. The Scandinavian noun phrase. *Working Papers in Scandinavian Syntax* 42: 57–80.

Delsing, L.-O. 1993. *The internal structure of noun phrases in Scandinavian languages*. Doctoral dissertation, University of Lund.

Déprez, V. 2005. Morphological number, semantic number and bare nouns. *Lingua* 115: 857–883.

Dobrovie-Sorin, C. 1987. À propos de la structure du groupe nominal en roumain. *Rivista di grammatica generativa* 12: 123–152.

Fassi Fehri, A. 1989. Generalised IP structure, Case and VS order. In *Working Papers in Linguistics 10: Functional Heads and Clause Structure*, Cambridge, MA: Department of Linguistics and Philosophy, MIT, pp. 75–113.

Giorgi, A., and G. Longobardi. 1991. *The Syntax of Noun Phrases*. Cambridge: Cambridge University Press.

Grosu, A. 1988. On the distribution of genitive phrases in Rumanian. *Linguistics* 26: 931–949.

Harlow, R. 2001. *A Maori Reference Grammar*. Auckland: Longman.

Higgins, R. 1973. *The pseudo-cleft construction in English*. Doctoral dissertation, MIT.

Holmberg, A. (ed.). 1992. *Papers from the Workshop on the Scandinavian Noun Phrase*. Report 32, Department of General Linguistics, University of Umeå.

Kayne, R. S. 1998. Overt vs. covert movement. *Syntax* 1: 128–191.

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2006. On parameters and on principles of pronunciation. In H. Broekhuis, N. Corver, R. Huibregts, U. Kleinhenz, and J. Koster (eds), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, Berlin

and New York: Mouton de Gruyter, pp. 289–299. [Reprinted in Kayne (2010, Ch. 4).]

Kayne, R. S. 2007. A note on the syntax of quantity in English. In J. Bayer, T. Bhattacharya, and M. T. H. Babu (eds), *Linguistic Theory and South Asian Languages: Essays in Honour of K. A. Jayaseelan. Linguistik Aktuell/Linguistics Today 102*, Amsterdam and Philadelphia: John Benjamins, pp. 73–105. [Reprinted in Kayne (2005, Ch. 8).]

Kayne, R. S. 2008a. Expletives, datives, and the tension between morphology and syntax. In T. Biberauer (ed.), *The Limits of Syntactic Variation*, Amsterdam and Philadelphia: John Benjamins, pp. 175–217. [Reprinted in Kayne (2010, Ch. 6).]

Kayne, R. S. 2008b. Some preliminary comparative remarks on French and Italian definite articles. In R. Freidin, C. P. Otero, and M. L. Zubizarreta (eds), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, Cambridge, MA: MIT Press, pp. 291–321. [Reprinted in Kayne (2010, Ch. 1).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Kripke, S. A. 1980. Preface to "Naming and necessity." In *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Longobardi, G. 1980. Modularity: From a single rule to several conditions without loss of generalization. The case of identificational predicates. Paper presented at the GLOW Colloquium, Nijmegen, April 1980.

Longobardi, G. 1994. Prolegomena to a "principles-and-parameters" theory of copular constructions. Ms., University of Venice.

Moro, A. 1993. *I predicati nominali e la struttura della frase*. Padua: Unipress.

Neale, S. 1990. *Descriptions*. Cambridge, MA: MIT Press.

Ouhalla, J. 1988. Movement in noun phrases. Paper presented at the LAGB Meeting, Durham, 1988.

Perlmutter, D. 1970. On the article in English. In M. Bierwisch and K. E. Heidolph (eds), *Progress in Linguistics*, The Hague: Mouton de Gruyter, pp. 233–248.

Postal, P. 1966. On so-called 'pronouns' in English. In F. P. Dinneen (ed.), *Report on the 17th annual Round Table Meeting on Linguistics and Language Studies*, Washington, DC: Georgetown University Press, pp. 177–184. [Reprinted as Postal (1969).]

Postal, P. 1969. On so-called 'pronouns' in English. In D. Reibel and S. Schane (eds), *Modern Studies in English*, Englewood Cliffs, NJ: Prentice-Hall, pp. 201–224.

Ritter, E. 1986. NSO noun phrase in a VSO language. Ms., MIT.

Ritter, E. 1989. A head-movement approach to construct-state noun phrases. *Linguistics* 26: 909–929.

Rosenbaum, P. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: MIT Press.

Rothstein, S. D. 1983. *The syntactic forms of predication*. Doctoral dissertation, MIT.

Rothstein, S. D. 1990. Subjects, predicates and saturation. Ms., Bar-Ilan University, Haifa.

Russell, B. 1905. On denoting. *Mind* 14: 479–493.

Ruwet, N. 1975. Les phrases copulatives en français. *Recherches linguistiques* 3: 143–191.

Schmitt, C., and A. Munn. 2002. The syntax and semantics of bare arguments in Brazilian Portuguese. In P. Pica and J. Rooryck (eds), *Linguistic Variation Yearbook* 2, Amsterdam and Philadelphia: John Benjamins, pp. 185–216.

Siloni, T. 1989. *Le syntagme nominal en hébreu*. Ms., University of Geneva.

Siloni, T. 1996. Hebrew noun phrases: Generalized noun raising. In A. Belletti and L. Rizzi (eds), *Parameters and Functional Heads: Essays in Comparative Syntax*, Oxford: Oxford University Press, pp. 239–267.

Stowell, T. 1981. *Origins of phrase structure*. Doctoral dissertation, MIT.

Stowell, T. 1989. Subjects, specifiers and X-Bar theory. In M. R. Baltin and A. S. Kroch (eds), *Alternative Conceptions of Phrase Structure*, Chicago and London: University of Chicago Press, pp. 232–262.

Szabolcsi, A. 1983/84. The possessor that ran away from home. *The Linguistic Review* 3: 89–102.

Szabolcsi, A. 1987. Functional categories in the noun phrase. In I. Kenesei (ed.), *Approaches to Hungarian 2*, Szeged: JATE, pp. 167–189.

Szabolcsi, A. 1994. The noun phrase. In F. Kiefer and K. É. Kiss (eds), *The Syntactic Structure of Hungarian. Syntax and Semantics 27*, San Diego: Academic Press, pp. 179–274.

Taraldsen, K. T. 1990. D-projections and N-projections in Norwegian. In J. Mascaró and M. Nespor (eds), *Grammar in Progress: GLOW Essays for Henk van Riemsdijk*, Dordrecht: Foris, pp. 419–431.

Torrego, E. 1988. Evidence for determiner phrases. Ms., University of Massachusetts, Boston.

22

The Noun Phrase

Anna Szabolcsi

1994

22.1 Introduction

Szabolcsi's work on the Hungarian noun phrase (Szabolcsi 1983/84, [1987b], 1994) has had a fundamental impact on the appreciation, in syntactic theory, of both Hungarian and the noun phrase. It constitutes a central pillar of the influential hypothesis that the structure of the noun phrase is parallel to that of the clause.

Part of the argument capitalizes on the empirically well-supported categorial distinction between the definite article and other determiners such as demonstratives and quantifiers. Szabolcsi adopts Abney's (1987) term DP, but diverges from that work in taking only the definite article (Hungarian *a(z)*) to correspond to D, while other determiners are in a lower head Det. D, she argues, is analogous to C, the functional category heading the clause: They are both subordinators that enable their complement to act as an argument. As such, they are predicted to be absent from (or empty in) nonembedded contexts, such as matrix clauses and vocative noun phrases, respectively. Languages with matrix complementizers, such as, for example, Korean, constitute an initial challenge for this prediction; but they do not represent a real counterexample, given that matrix complementizers in this language are markers of clause type. Some empirical challenges remain, such as the possibility of a definite article in vocatives in some languages, like Romanian, as Szabolcsi herself notes. But the hypothesis appears to be overall supported.

A central argument for the CD/DP parallelism comes from the syntax of possessor phrases. Possessors can occur in two positions within their possessee's DP in Hungarian. One position is between the definite article and the possessee nominal, the second (and higher) position is left-peripheral, that is, to the left of the definite article. Like Spec,CP in the clause, Spec,DP is an operator position:

Bare operator possessors must move there. The difference between the two positions is also reflected in the morphology: in the lower position the possessor is not overtly case-marked (lack of overt case marking is identified as nominative case in Hungarian), whereas in the higher position it exhibits dative morphology.

Furthermore, possessors can extract from the DP, in which case they are also dative marked. Szabolcsi argues that extraction proceeds via a left-peripheral escape hatch, analogous to subject extraction from the clause. The role of the article in possessor extraction is analogous to that of the complementizer in Rizzi's (1990) account of subject extraction. Once the possessor extracts from DP, it and the DP, Szabolcsi notes, behave like anaphorically related co-arguments of the verb (an observation that the reader might keep in mind when reading Hornstein 1999, Ch. 28 of this volume). In the context of the existential verb *van* 'be.' possessor extraction gives rise to the Hungarian equivalent of a *have*-sentence, in which the nominative noun phrase agrees in person and number with a dative co-argument.

In the second part of the paper, Szabolcsi investigates the parallelism between the verbal and the nominal domain with regard to argument structure, concluding that complex event nominals have argument structure in a way largely analogous to verbs.
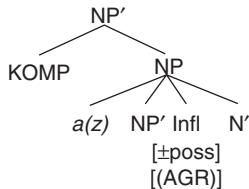
The hypothesis of the crosscategorial parallelism between the noun phrase and the clause has provoked much work (cf. *Theoretical Linguistics* 36.2/3), and remains a topic of fruitful discussion. Beyond its agenda-setting theoretical discussion, *The Noun Phrase* provides a rich yet clearly articulated pathway into the morphosyntax of the Hungarian noun phrase and a wealth of observations and ideas.

22.2   From "THE NOUN PHRASE: PART I, POSSESSORS AND DETERMINERS"
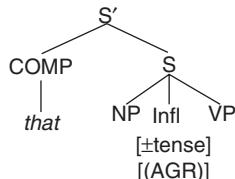
3   On the clausal analogy to be proposed

In this chapter I lay out a significant parallelism between the structures of Hungarian noun phrases and (configurational) clauses. The first version of this analysis was put forth in Szabolcsi (1981, 1983/84), where I proposed the structure in (21a), which is to be compared with the then-standard clausal structure (21b), cf. Chomsky (1981):
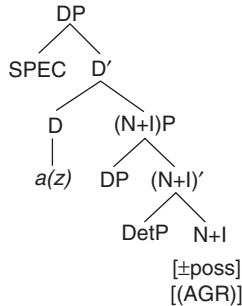
(21) a.   b.
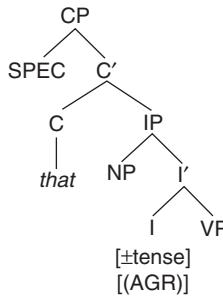
Both structures contain a lexical item not properly attended to. In the noun phrase structure, *a*(*z*) 'the' hangs unlabeled from NP; in the clausal structure, *that* is in

the same COMP position that serves as a landing site for WH-movement. Among other things, attention to such details was the critical factor in motivating a revision of both analyses. In Szabolcsi ([1987a, 1987b], 1989) I proposed the structure (22a), compare Chomsky's (1986) (22b).

(22) a.
```
            DP
           /  \
        SPEC   D′
              /  \
             D    (N+I)P
             |    /    \
           a(z)  DP    (N+I)′
                      /     \
                   DetP      N+I
                           [±poss]
                           [(AGR)]
```

b.
```
            CP
           /  \
        SPEC   C′
              /  \
             C    IP
             |   /  \
          that  NP   I′
                    /  \
                   I    VP
                [±tense]
                [(AGR)]
```

I label the full noun phrase DP in acknowledgement of Abney's (1986, 1987) proposal. A brief comparison of the main similarities and differences may be useful here.

(23) a. Szabolcsi: The (Hungarian) noun phrase has a sentence-like structure. It contains inflection. It is headed by a determiner.
Abney: The (English) noun phrase has a sentence-like structure. It contains inflection. It is headed by a determiner.

b. Szabolcsi: (Hungarian) inflection is "real" inflection, whereas the determiner that heads the noun phrase is an analog of the complementizer (C). That is, DP = CP.
Abney: (English) inflection (viz.,'*s*, Abney, 1986, or empty AGR, Abney, 1987) and the determiner that heads the noun phrase belong to the same category. That is, DP = IP.

c. Szabolcsi: (Hungarian) determiners fall into two distinct categories. Only the article belongs to the category D that heads the noun phrase.
Abney: All (English) determiners belong to the category D that heads the noun phrase.

Abney (1987) builds on my analysis of Hungarian in two ways. On the one hand, the observations concerning inflection and the co-occurrence of possessors and determiners are used to motivate details of his proposal concerning English. On the other hand, he reanalyzes some of the Hungarian facts along those lines. I briefly comment on his reanalysis in section 4 [of the full article]; given the focus of this volume, I will not attempt to work out a proposal for English. It may be important to bear in mind that some of the global differences between Abney's analysis and mine stem not from disagreement, but from the fact that he explicitly restricts his attention to the justification of an inflection-like head in the noun phrase, and thus a large portion of the arguments presented below go beyond the scope of his proposal.

[. . .]
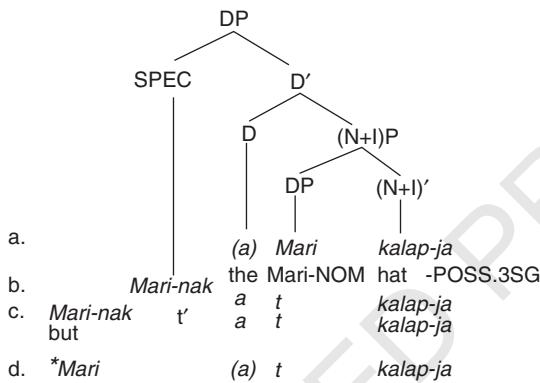
5 The structure of DP: Possessor extraction

The claim that the possessive construction has a sentence-like structure becomes interesting if further data can be insightfully analyzed in this light. Data concerning possessor extraction are one case in point.

In section 2 [in the full article] I mentioned that the possessor has an alternative, dative-marked variant in the noun phrase [. . .] This differs from the nominative variant in that it can be extracted. I will assimilate the mechanics of possessor extraction to that of subject extraction, arguing specifically for the analysis in (37).

(37)

	SPEC	D	DP	(N+I)′
a.		*(a)*	*Mari*	*kalap-ja*
		the	Mari-NOM	hat -POSS.3SG
b.	*Mari-nak*	*a*	*t*	*kalap-ja*
c.	*Mari-nak* t′	*a*	*t*	*kalap-ja*
	but			
d.	**Mari*	*(a)*	*t*	*kalap-ja*

I will point out that the interest of these data lies not only in the DP/CP parallelism, but also in the fact that they provide support for a specific analysis of subject extraction.

[. . .]

The main claims to be made in this section are as in (38)–(41). I refer to the "dative-marked possessor" as "-*nak* possessor."

(38) The -*nak* possessor is a constituent of the noun phrase, and it is not only morphologically, but also structurally, distinct from the nominative possessor.

(39) The specifier of DP has some characteristics of operator positions.

(40) The possessor cannot be extracted directly from its root position, but it can be if it proceeds through the specifier of DP.

(41) Once the possessor leaves DP, it and the DP behave as if they were independent, though anaphorically related, arguments of the verb.

It is to be noted that only the possessor can be extracted from DP; inherently case marked complements and adjuncts cannot.

5.1 On the distinctness of two possessor positions within DP

Let us begin with (38). Given that Hungarian has relatively free word order, and given that the -*nak* possessor need not be adjacent to the possessed noun, the suspicion may arise that they never really form a constituent. This can be refuted

straightforwardly by showing that they together undergo WH-movement and focus-movement, which affect only constituents. (Both target the same preverbal position, whose nature is discussed by É. Kiss, [1994])

(42) *Kati   ki- nek a kalap-já-   t látta?*
      Kati(-NOM) who-DAT the hat- POSS.3SG-ACC saw
      'Whose hat did Kati see?'

(43) *Kati (nem) csak Mari-nak a   kalap-já-   t   látta*
      Kati not   only Mari-DAT the hat- POSS.3SG-ACC saw
      'Kati saw (not) only Mari's hat (but . . .)'

They can also be conjoined.

(44) *Kati Mari-nak a kalap-já-   t és Péter-nek a   kabát-*
      Kati Mari-DAT the hat- POSS.3SG-ACC and Peter-DAT the coat-
        já-   t látta.
        POSS.SG- ACC saw
      'What Kati saw was Mari's hat and Peter's coat.'

Next, we need to show that the structural position of the *-nak* possessor is different from the nominative possessor's. The most important argument in favor of this is that the former precedes, and the latter follows, the article *a(z)*; thus it is crucial to show that the article that precedes the nominative possessor does not (need to) belong to this possessor but rather to the whole of the construction. The data are descriptively complicated but, to my mind, uncontroversial.

First, as was noted in connection with (17) [in the full work], personal pronoun possessors in present-day Hungarian are always preceded by *a(z)* (a significant case when they are not is discussed in section 6). But personal pronouns normally do not take an article (of course).

(45) a. *az én   kalap- om*   *az én   kalap- ja- i- m*
          the I(-NOM) hat- POSS.1SG   the I(-NOM) hat- POSS- PL- 1SG
          'my hat'   'my hats'
      b. * *én   kalap-om*   * *én   kalap- ja- i- m*
           I(-NOM) hat- POSS-1SG   I(-NOM) hat- POSS- PL- 1SG
(46) a. * *Az én   isz- om.*
           the I(-NOM) drink- 1SG
      b. *Én   isz- om.*
          I(-NOM) drink- 1SG
          'I drink.'

If only pronominal possessors exhibited this pattern, it might be possible to analyze (45)–(46) entirely differently from non-pronominal constructions. But the pattern of (45)–(46) is essentially replicated by names of persons in the Upper Tisza and the Debrecen dialects, documented in Simonyi (1914), Magda Szabó's novel *Freskó* (1958), and Magda Szabó (personal communication, 1988). These

dialects are unique in that names of persons do not take an article but as possessors they are preceded by one. The (*) in (47b) indicates that the article-less form is not ungrammatical but it is not the usual form in this dialect:

(47) a. *a János   kalap-ja*
          the Janos(-NOM) hat- POSS.3SG
          'Janos's hat'
      b. (*) *János   kalap-ja*
              Janos(-NOM) hat- POSS.3SG
(48) a. **A János   isz- ik*.
          the Janos(-NOM) drink-3SG
      b. *János   isz- ik*.
          Janos(-NOM) drink-3SG
          'Janos drinks.'

These contrasts mean that, at least in (45) and (47), $a(z)$ belongs to the whole possessive construction and is not part of the possessor. I assume that in the spirit of parametric variation it is legitimate to use dialectal data to support a unified analysis. Thus I conclude that in other cases, where the possessor itself has an overt or covert article, or some other determiner, there is an $a(z)$ missing (deleted, as I argue in section 6). As a result, in the statistically speaking typical case the determiner we actually see is in fact part of the nominative possessor. This is the case in both (49) and (50). (50) presents two dialects (neither Upper Tisza) that differ in whether proper names take an overt article in their own right. In the rest of the paper I use the (50b) variant unless otherwise indicated.

(49) a. *a /egy/minden fiú   kalap-ja*
          the/one/every   boy(-NOM) hat- POSS.3SG
          'the/a/every boy's hat'
(50) a. *a Mari   kalap-ja*   cf. *a Mari*
          the Mari(-NOM) hat- POSS.3SG   the Mari
          'Mari's hat'   'Mari'
      b. (**a*) *Mari   kalap-ja*   cf. (**a*) *Mari*
          the Mari(-NOM) hat- POSS.3SG   the Mari
          'Mari's hat'   'Mari'

These latter facts make the picture superficially complicated, but they do not seem to call into question that "possessor-independent $a(z)$" exists. It "(re)surfaces" in the examples in (51), which are synonymous with the ones above.[7]

(51) a. *János-nak a – kalap-ja*   cf. (47a)
          Janos-DAT the hat- POSS.3SG
          'Janos's hat'
      b. *a /egy/minden fiú-nak a – kalap- ja*   cf. (49)
          the/one/every boy-DAT the hat- POSS.3SG
          'the/a/every boy's hat'

c. *Mari-nak a – kalap-ja*   cf. (50)
    Mari-DAT the hat- POSS.3SG
    'Mari's hat'

5.2 On the operator character of SPEC of DP

It is the relation between the nominative and the -*nak* possessors that interests us in this section. I take the above data to mean that their structural positions are different. And since they are in complementary distribution, I argue that the possessor moves from the post-article position into the pre-article position.

This movement is analogous to the movement of the subject to the clause-initial position (formerly COMP, currently SPEC of CP). For ease of reference, from now on I call this position SPEC of DP, although its exact nature is only established in the subsequent sections.

(52) [$_{DP}$ [$_{SPEC}$ *Mari-nak*] [$_{D'}$ *a* [$_{(N + I)P}$ t *kalap-ja*]]]

One reason to believe that SPEC of DP is analogous to SPEC of CP is that it seems to be an operator position. First, in present-day Hungarian bare operator possessors must move here.

(53) a. **ki   kalap- ja*
              who(-NOM) hat- POSS.3SG
      b. *ki-   nek a* t *kalap-ja*
              who-DAT the hat- POSS.3SG
              'whose hat?'

Similarly for *aki* 'who (relative)', *melyik* 'which one', *mindenki* 'everyone', *senki* 'no one', and numerals with an empty head noun, *hány* 'how many [ones]', *három* 'three [ones]', and so on. Possessors consisting of a WH or quantificational determiner and a noun, and possessors that do not have any lexical operator features, move to SPEC of DP optionally.

(54) a. *hány   fiú   kalap-ja*
              how-many boy(-NOM) hat- POSS.3SG
              'how many boys' hats'
      b. *hány   fiú-nak a* t *kalap-ja*
              how-many boy-DAT the hat- POSS.3SG
              'how many boys' hats'

and similarly for *melyik fiú* 'which boy', *amelyik fiú* 'whichever boy', *minden fiú* 'every boy', *semelyik fiú* 'neither boy', *a fiú* 'the boy', *Mari*, and so on. Note that it is clearly bare operatorhood, rather than indefiniteness, that characterizes the class patterning with (53). For similar contrasts among operators, see Ambar, Lois, and Obenauer (1986).

Second, although non(-bare) operator possessors move to SPEC of DP optionally, it appears that, once there, they acquire an abstract operator feature (or, alternatively, possessors that in fact move may have always had one). This assumption, together with

pied piping, provides a simple explanation for the patterning of multi-layered posses-sive constructions. The important observation is that if *Mari* becomes a *-nak* possessor, the containing DP, *Mari-nak a barát-ja* 'Mari's friend', must also do so: (55d) is ungrammatical. I use article-less *Mari* for simplicity's sake. All versions mean the same.

(55) a.   *Mari   barát- ja   kalap-ja*
           Mari(-NOM) friend-POSS.3SG(-NOM) hat- POSS.3SG
           'Mari's friend's hat'
      b.   *Mari   barát- já-   nak a kalap-ja*
           Mari(-NOM) friend- POSS.3SG-DAT the hat- POSS.3SG
      c.   *Mari-nak a   barát-já-   nak a kalap-ja*
           Mari-DAT the friend-POSS.3SG-DAT the bat- POSS.3SG
      d. **Mari-nak a   barát- ja   kalap-ja*
           Mari-DAT the friend-POSS.3SG(-NOM) hat- POSS.3SG

This pattern is easily understood if we assume that the operator feature of the most deeply embedded possessor in SPEC percolates onto its containing DP, forcing it to move to SPEC itself. (*Where did he go, do you think?* may be analyzed as a similar, though more restricted, case of clausal pied piping.)

   To sum up, SPEC of DP is a non-thematic operator position. The question arises whether it is an argument (A) or non-argument (A-bar) position in current terms. The above would suggest it is A-bar, but the fact that the possessor acquires a *-nak* 'dative' morpheme here may be disturbing. It is assumed in Mahajan (1990) that case-marking is the defining property of A-positions. *-nak* is presumably not a real case marker here, however. On the one hand, the possessor moves into SPEC of DP from an already case-marked position. On the other hand, the *-nak* morpheme serves a variety of other un-case-like purposes, such as marking modifiers in left dislocation and in complex predicate constructions (*Boldog-nak boldog vagyok* 'Happy I am,' *Boldog-nak látszol* 'You seem happy,' etc.). I have nothing insightful to say about how *-nak* arises, but its source is certainly DP-internal (presumably, D itself). The reason is that *-nak* is available in vocatives, discussed in section 6, in which event it cannot have a DP–external source.

5.3 Possessor movement

With all this in mind, let us turn to possessor movement. Recall (37), abbreviated here.

(37) a.   [$_{DP}$ [$_{SPEC}$] [$_{D'}$ *a* [$_{(N + I)P}$ *Mari kalap-ja*]]]
      b.   [$_{DP}$ [$_{SPEC}$ *Mari- nak*] [$_{D'}$ *a* [$_{(N + I)P}$ t *kalap-ja*]]]
      c.   *Mari-nak* . . . [$_{DP}$ [$_{SPEC}$ t′] [$_{D'}$ *a* [$_{(N + I)P}$ t *kalap-ja*]]]
      d. **Mari*   . . . [$_{DP}$ [$_{SPEC}$] [$_{D'}$ *a* [$_{(N + I)P}$ t *kalap-ja*]]]

The basic observation is that the nominative possessor is confined to the adjacent-to-possessed position, whereas the *-nak* possessor surfaces either in the pre-article position or outside the noun phrase (in complementarity with the nominative possessor). Whichever version of generative syntax of the past two decades we are assuming, these facts indicate that the possessor is capable of moving first to a

peripheral position, and then out of the noun phrase, but not of moving out of the noun phrase in one swoop. This is exactly analogous to the procedure assumed for subject extraction in configurational languages like English. Note, though, that in the case of subjects this procedure had been established on more or less speculative grounds, which led proponents of other theories to deny the involvement of a peripheral position on the whole. What is particularly interesting about the possessor data is that they provide more solid theory-neutral evidence for the assumed procedure than any of the actual subject extraction data I am aware of in the literature. More precisely, they provide evidence for the existence of a peripheral position, distinct from the nominative one, and for the relatedness of the detached possessor to this peripheral position, as opposed to the nominative one. They of course do not provide evidence for the global claim that these relations are to be captured by movement (literal or metaphorical), as opposed to some other theoretical device, since such claims cannot be directly justified by data.

Naturally, not every version of generative syntax is equally successful in accounting for all the details of this procedure. Note in particular the presence of the article $a(z)$, which I argue to be analogous to the complementizer in clauses. Although some cross-linguistic variation was known, preoccupation with the blocking effect of *that* in English gave rise to theories that treated the complementizer as a mere nuisance for subject extraction, as in (56).

(56)   *Who do you think (*that) left?*

The article in the Hungarian noun phrase, on the other hand, does not harm possessor movement at all. This discrepancy was the main reason why in Szabolcsi (1981) I left $a(z)$ practically unanalyzed, compare (21a) above. Rizzi (1990), exploring the C-headed clausal structure proposed in Chomsky (1986), put forth the first detailed theory of extraction according to which the complementizer, overt or covert, is potentially a blessing, rather than a curse, for the movement of subjects. The Hungarian possessor movement data can now be seen as providing specific support for Rizzi's theory.

In Szabolcsi (1989) I provide a detailed account of possessor extraction in Rizzi's terms. I refrain from reproducing it here as it is rather technical in nature. Let me summarize it as follows. The possessor cannot extract in one swoop because its root position is not governed by an active head. By moving to SPEC of DP first and entering into abstract specifier–head agreement with D, it turns D into an active governor of its trace in the root position. In SPEC of DP it is governed by the verb, so it can move on.

Most languages do not exhibit possessor extraction; they typically lack even the first step of movement. In case the structure of their noun phrases is by and large similar to the one I assume for Hungarian, Rizzi's theory offers two basic options to explain the absence of possessor movement. One is that the possessor is unable to trigger specifier–head agreement with D (for instance, because it is not a plain noun phrase); the other is that D cannot carry even abstract agreement features. I have no account of why the requisite agreement relation is such a rare phenomenon.

[. . .]

6.1   On the co-occurrence of articles and "other" determiners

The fact that nouns may be doubly determined in Hungarian had not been observed in either descriptive or theoretical literature before Szabolcsi ([1987a]). The reason is that in the statistically speaking typical case the co-occurrence of determiners (DETs) is strictly prohibited. [Review (13) and (29) [in the complete work] for fuller lists.]

(65)   a. $\begin{bmatrix} az & \textit{'the'} \\ ezen \,/\, azon & \textit{'this / that'} \\ minden & \textit{'every'} \\ melyik & \textit{'which'} \end{bmatrix}$ *állítás* (om) *'claim* (-poss.1sg)*'*

   'the/this/every/which/ . . . claim (of mine)'

   b. * $\begin{bmatrix} aminden & \textit{'the every'} \\ minden\ a & \textit{'every the'} \\ minden\ ezen & \textit{'every this'} \\ \dots \end{bmatrix}$ *állítás* (om) *'claim* (-poss.1sg)*'*

These data create the impression that all the determiners compete for the same position. That this is false becomes clear, however, as soon as some string, to be notated as \$, manages to intervene between the two items. In that case determiners split into two groups. $a(z)$ 'the' appears preceding \$, and only there, whereas all others appear following \$, and only there. Moreover, they may co-occur, so that we cannot say that the article gets into its surface position by fronting, as was pointed out to me by M. Brody (personal communication, 1986). The pattern is summarized in (66).

(66)   D \$ Det N   where D = $\{a(z), \varnothing\}$
                                  Det = $\{minden, ezen, melyik, \dots\}$

What are the \$s that can intervene between D and Det? I am aware of two entirely independent possibilities; their independence is important as it indicates that pattern (66) is not the peculiarity of an individual construction. One \$ may be an overt possessor in the nominative (whether pronominal or not). That is, data like (45) and (29) [in the full article] can be combined.

(67)   *az én   minden állítás-om*
       the I(-NOM) every   claim-POSS.1SG
       'my every claim'

The other relevant construction is the one with a prenominal participial modifier, as in *tőled kapott* 'received from you'. This may either immediately precede the head noun (here *levél* 'letter'), as in (68a), or it may be separated from it by a Det (here *valamennyi* 'each'), as in (68b). In the latter case, however, an article appears obligatorily.

(68) a. *Valamennyi* [*től– ed   kapott*] *levél*   *rövid volt*.
          each   from–2SG received letter(–NOM) short was
          'Each letter received from you was short.'
      b. *A* [*től-ed   kapott*]   *valamennyi levél*   *rövid volt*.
          the from–2SG received   each   letter(–NOM) short was
          'Each letter received from you was short.'
      c. **A*   valamennyi* [*től– ed kapott*] *levél*   *rövid volt*.
          the each   from–2SG received letter(–NOM) short was
      d. **[*Tőled   kapott*] *valamennyi levél rövid volt*.
          from–2SG received each   letter short was

The crucial observation to be made now is that whereas the presence of the article
is required in one set of the examples and prohibited in the other, this makes no
difference for interpretation. That is, (68a,b) are synonymous and, similarly, the
interpretation of (67) differs from that of (65a) with *-om* only in that it contains
a stressed rather than a dropped pronoun. (Both *minden* 'every' and *valamennyi*
'each' are distributive quantifiers, exactly like their English counterparts.) It is
therefore convenient to assume that an article is underlyingly present in all cases,
but its surface realization is restricted. The emergent generalization is as in (69).

(69) Haplology:
      a. The co-occurrence of D and Det is grammatical if they are linearly
          separated by some intervener.
      b. Contiguous strings of the type D Det, or D D, are ungrammatical.
          Ungrammaticality can be eliminated either by deleting *a*(*z*) of D in
          phonetic form, or by moving the constituent that contains Det or the
   second D.

    [. . .]

6.2 Articles as subordinators

Recall now that we are working toward the full justification of a proposal under
which the article is analogous to the complementizer.

(22) a.   b.

Important preliminary conclusions have been reached: [C] is a functional head that plays a similar role in possessor extraction to what D plays in subject extraction and, being distinct from real determiners, it needs some function within the noun phrase. The critical question left to be answered is this: Can we attribute analogous functions to D and C? Following Szabolcsi ([1987a, 1987b]) the suggestion is as in (80) (the proposal is refined below).

(80) a. Only phrases in the canonical argument format can function as arguments of θ-role assigning heads.
      b. Both the complementizer and the article are subordinators in the sense that they enable the clause or noun phrase to act as arguments.

There are general syntactic considerations that lend some plausibility to the claim that arguments in the above sense come with a subordinator. First consider what categories, besides noun phrases, are assigned thematic roles. In terms of Chomsky (1981), they are embedded finite clauses, infinitival clauses, and small clauses. Embedded finite and infinitival clauses are standardly assumed to have a complementizer, whether overt or phonetically null. Small clauses are a misfit because they do not have a complementizer but require a thematic role. However, Stowell (1991) argues that they in fact undergo restructuring, at s-structure or at LF. He proposes that this is forced by a principle like (81a) or (81b).

(81) a. A predicative category may not function as an argument.
      b. Only a referential category may function as an argument.

On this proposal, small clauses no longer constitute an exception.

Next, consider categories that do not act as arguments. Matrix clauses are a case in point, and we know that in most languages they may not have a complementizer (I return to exceptions below).

(82) *That John left.

Let us now look for an analog of matrix clauses in the domain of noun phrases. Vocatives suggest themselves, since they are quite obviously not arguments of any predicate. The question is, May vocatives contain an article?

It is well known that in many languages/dialects names of persons take an article. Such is the case, for instance, in various dialects of German and Hungarian (but never in languages with a word-final article, e.g., Scandinavian, Macedonian, and Bulgarian; F. Kiefer, personal communication).

(83) *Der Peter kommt.*
      the Peter comes
      'Peter is coming.'
(84) *Jön   a   Péter.*
      comes the Peter
      'Peter is coming.'

It is clear that in (83)–(84) the presence of the article is a purely formal require-
ment: it does not change the meaning of the name in any usual sense. Nonetheless,
even in those dialects the article is impossible in the vocative.

(85) *Peter, komm!*   versus   **Der Peter, komm!*
      'Peter, come'

(86) *Péter, gyere!*   versus   **A Péter, gyere!*
      Peter come-IMP.2SG   the Peter come-IMP.2SG
      'Peter, come'

Similarly, I noted that possessive constructions whose nominative possessor is a
personal pronoun are invariably introduced by an article [cf. (17) [in the full
article] and (45)]. Vocatives are the one exception.

(87) *Én   barát- om,   gyere!* versus
      I(-NOM) friend-POSS.1SG . . .
      'My friend, come'

      *Az én   barát- om,   gyere!
      the I (-NOM) friend-POSS.1SG

The absence of an article from vocatives is precisely what my proposals concerning
the D/C parallelism and subordination predict.

   It may be interesting to point out that D is not absent from vocatives, only genu-
inely empty. The arguments for this are similar to the arguments concerning the
presence of C in matrix clauses. For instance, the possessor of the vocative may be
in the dative, making use of the SPEC of D position, just as the SPEC of C can be
filled in matrix questions.

(88) a. *Péter   barát- ja!*
          Peter(-NOM) friend-POSS.3SG
          'Peter's friend'
      b. *Péter- nek barát- ja!*
          Peter-DAT   friend-POSS.3SG
          'Peter's friend'

Related facts are discussed in Longobardi (1990), who proposes that vocative
Gianni mio 'John my' differs from argumental *il mio Gianni* 'the my John' in that
Gianni underwent N-to-D movement. For some reason, no similar movement into
D is possible in Hungarian.

(89) a. *Kicsi János!* versus b. **János kicsi!*
          little Janos   Janos little
          'Little Janos'

Incidentally, C and D raise the intriguing theoretical question of how it is possible
for a head category to project when it is genuinely empty of lexical content. I have
no answer to offer.

22.3   Questions pertaining to Szabolcsi (1994)

1   An important part of Szabolcsi's paper concerns the syntax of dative possessors in Hungarian. To what extent can her analysis of these dative possessors be transposed to genitive possessors in Turkish?

2   To what extent can her analysis of possessive sentences be transposed to languages of the Mayan or Quechua families?

3   For Szabolcsi, dative possessors pick up their dative Case as the result of movement to the specifier position of the larger DP containing them. This aspect of her analysis is contested by den Dikken (1999). Which approach would fit better into a strongly uniformist perspective according to which possessive DPs are crosslinguistically virtually identical? Give your reasons.

4   Szabolcsi notes that English "*my train* need not be one that I built or one that I own: it may be one that I ride to work, one that I just missed, one that I like to watch passing by at dusk," etc. Against the background of such a broad range of possible interpretations in English (and Hungarian), pick three languages and test to see if they have a range of possessive interpretations equal to English and to each other. If they do not, what might be the implications?

5   In Hungarian possessive DPs, whether with a dative or a nominative possessor, the noun agrees with that possessor (which is prenominal). For example one has *az én kalap-ja-i-m* ('my hats'), in which *én* is the prenominal nominative possessor and *-m* the agreement morpheme. (*Az* is the definite article, *kalap* ('hat') the noun, *-ja-* a possessive morpheme, and *-i-* the plural morpheme.) Taking into account Bernstein and Tortora (2005), discuss the question whether all languages should be taken to have (possibly silent) possessor agreement.

6   In Hungarian, possessor agreement is also found in possessive sentences in which a dative possessor has been extracted from within the larger DP, e.g., *Nekem van kalapom* ('I have a hat.'), in which *nekem* is a stressed (complex) first person singular dative, *van* is the existential verb used in possessive sentences, and the *-m* of *kalapom* is the agreement morpheme. Unlike Hungarian, English possessive sentences have a nominative possessor and a verb *have* (rather than existential *be*). These English–Hungarian differences leave open the question whether English has a counterpart of Hungarian possessive agreement. Discuss the pros and cons of taking pronominal possessors like *his, our, her* in sentences like *He has a mind of his own, We have our faults, too, Mary has her (preferred) way of doing things* to be those counterparts.

7   Szabolcsi points out that the existential *van* mentioned in the previous question differs both from the *van* of locative sentences and from copula *van* in the following way. In the third person singular indicative the copula must be silent (and the locative verb may be silent), but the existential one must be pronounced. Why would 3sg. be singled out in this way? And why must the existential verb always be pronounced, even in the 3sg.? Bring in as many other languages as you can.

8   Does the possessive AGR illustrated in questions 5 and 6 project, as it would
    if Pollock (1989) is transposed to these cases? Give your reasons, while
    bringing in Rizzi (1982, chapter 4) and Kayne (2008a). To what extent does it
    matter here whether or not person and number agreement are considered
    distinct?

9   Szabolcsi mentions that in many languages/dialects names of persons take a
    definite article, for example in various dialects of German and Hungarian.
    She cites a suggestion by F. Kiefer to the effect that this co-occurrence of
    definite article and proper name is never found in languages with a word-final
    article (e.g., Scandinavian, Macedonian, and Bulgarian). Test Kiefer's
    hypothesis using as many languages as possible.

10 Szabolcsi makes the following conjecture: "Only in languages that have no
    overt articles do nonspecific direct objects fail to be accusative marked (or, in
    general, fail to be marked in the same way as specific direct objects)." Test this
    conjecture using as many languages as possible.

11 As shown in this paper, Hungarian requires extraction of the possessor when
    the containing DP is a nonspecific indefinite, a fact that Szabolcsi links to the
    well-known definiteness effect seen in existential sentences, as well as in the
    contrast, in the simple possessive sense of *have*, between *John has a sister* and
    *John has the sister. Yet the latter itself contrasts with *John has the sister he
    deserves, Mary has the makings of a first-rate linguist*. How might one express
    these contrasts from the perspective of Kayne's (2008b) idea that the definite-
    ness effect is to be traced back to a prohibition against extracting (a silent
    counterpart of) *there* across a definite article (in a way recalling Fiengo and
    Higginbotham 1981)? Bring in Leu (2008).

12 How might the facts mentioned in the previous question be related to the fact
    that Hungarian requires the definite article *a(z)* to be pronounced in the
    presence of a prenumeral modifier? How might they be related to English *We
    want the ones *(that are not expensive)*?

13 Hungarian differs from English in that Hungarian lacks a transitive verb like
    have for expressing possession. Mahajan (1994) argues that "head-final"
    languages lack a transitive *have* for principled reasons. What precise notion of
    "head-final" might overcome the problem posed for his specific formulation
    by German and Basque? Show how that notion of "head-final" might account
    for Hungarian lacking transitive *have*.

14 The Hungarian possessive morpheme *ja* mentioned in question 5 recalls to
    some extent English '*s*. How might one account for their difference in
    position? How would you integrate such possessive morphemes into Cinque's
    (2005) analysis of DP-structure?

15 The complex dative form *nekem* of question 6 shows the first singular
    agreement morpheme *-m* that one finds in possessive DPs, as part of a more
    general (partial) resemblance between postpositional phrases and possessive
    DPs in Hungarian. How might this internal complexity of dative *nekem* be
    made compatible with Szabolcsi's proposal that dative Case is picked up in
    Spec,DP via movement from a lower (nominative) position? To what extent
    might Kayne (2004) be relevant?

16 Szabolcsi emphasizes a parallelism between complementizers and (definite) articles. To what extent is that compatible with Rosenbaum's (1967) idea that sentential arguments are always accompanied by an *it* (which may be deleted/silent)? How exactly would Kayne's (2008a; to appear) proposal that sentential complements are relative clauses affect that possible parallelism? Which of these perspectives is most compatible (and which is least compatible) with Chomsky's (2001) proposal that CP is a phase?

References

Abney, S. 1986. *Functional elements and licensing*. Talk presented at GLOW. Available at http://www.vinartus.net/spa/86b.pdf. [Accessed January 2013.]

Abney, S. 1987. *The English noun phrase in its sentential aspect*. Doctoral dissertation, MIT.

Ambar, M., X. Lois, and H. Obenauer. 1986. Move Infl and sentential arguments. *GLOW Newsletter* 16: 17–18.

Bernstein, J., and C. Tortora. 2005. Two types of possessive forms in English. *Lingua* 115: 1221–1242.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, pp. 1–52.

Cinque, G. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36: 315–332.

den Dikken, M. 1999. On the structural representation of possession and agreement. The case of (anti-)agreement in Hungarian possessed nominal phrases. In I. Kenesei (ed.), *Crossing Boundaries: Theoretical Advances in Central and Eastern European Languages*, Amsterdam and Philadelphia: John Benjamins, pp. 137–178.

Di Sciullo, A. M., and C. Boeckx (eds), 2011. *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford Studies in Biolinguistics. Oxford and New York: Oxford University Press.

É. Kiss, K. 1994. Sentence structure and word order. In F. Kiefer and K. É. Kiss (eds), *The Syntactic Structure of Hungarian. Syntax and Semantics* 27, San Diego: Academic Press, pp. 1–90.

Fiengo, R., and J. Higginbotham. 1981. Opacity in NP. *Linguistic Analysis* 7: 395–421.

Hornstein, N. 1999. Movement and control. *Linguistic Inquiry* 30: 69–96.

Kayne, R. S. 2004. Prepositions as probes. In A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures* 3, New York and Oxford: Oxford University Press, pp. 192–212.

Kayne, R. S. 2008a. Antisymmetry and the lexicon. In J. Rooryck and J. van Craenenbroeck (eds), *Linguistic Variation Yearbook 8*, Amsterdam and Philadelphia: John Benjamins, pp. 1–31. [Reprinted in Kayne (2010, Ch. 9) and in Di Sciullo and Boeckx (2011, Ch. 14).]

Kayne, R. S. 2008b. Expletives, datives, and the tension between morphology and syntax. In T. Biberauer (ed.), *The Limits of Syntactic Variation*, Amsterdam and Philadelphia: John Benjamins, pp. 175–217. [Reprinted in Kayne (2010, Ch. 6).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Kayne, R. S. to appear. Why isn't *this* a complementizer? In P.Svenonius (ed.), *Functional Structure from Top to Toe: A Festschrift for Tarald Taraldsen*, New York and Oxford: Oxford University Press. [Reprinted in Kayne (2010, Ch. 10).]

Leu, T. 2008. *The internal syntax of determiners*. Doctoral dissertation, New York University.

Longobardi, G. 1990. N-movement in syntax and in LF. Ms., University of Venice.

Mahajan, A. 1990. *The A/A-bar distinction and movement theory*. Doctoral dissertation, MIT.

Mahajan, A. 1994. The ergativity parameter: *have-be* alternation, word order and split ergativity. In M. Gonzàlez (ed.), *Proceedings of NELS 24*, Amherst: GLSA, University of Massachusetts, vol. 1, pp. 317–331.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Rizzi, L. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Rizzi, L. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Rosenbaum, P. 1967. The Grammar of English Predicate Complement Constructions. Cambridge, MA: MIT Press.

Simonyi, Zs. 1914. *A jelzők mondattana* (The syntax of nominal modifiers). Budapest: MTA.

Stowell, T. 1991. *Small clause restructuring,* In R. Freidin (ed.), *Principles and Parameters in Comparative Grammar*, Cambridge, MA: MIT Press, pp. 182–218.

Szabolcsi, A. 1981. The possessive construction in Hungarian. *Acta Linguistica Scientiarum Academiae Hungaricae* 31, 261–289.

Szabolcsi, A. 1983/84. The possessor that ran away from home. *The Linguistic Review* 3: 89–102.

Szabolcsi, A. 1987a. *A birtokos szerkezet és az egzisztenciális mondatok* (The Possessive Construction and Existential Sentences). Doctoral dissertation, Hungarian Academy of Sciences. [Revised version published 1992, Budapest, by Akadémiai Kiadó.]

Szabolcsi, A. 1987b. Functional categories in the noun phrase. In I. Kenesei (ed.), *Approaches to Hungarian 2*, Szeged: JATE, pp. 167–190.

Szabolcsi, A. 1989. Noun phrases and clauses: Is DP analogous to IP or CP? In J. Payne (ed.), *The Structure of Noun Phrases,* to appear, Berlin: Mouton de Gruyter.

Szabolcsi, A. 1994. The noun phrase. In F. Kiefer and K. É. Kiss (eds), *The Syntactic Structure of Hungarian. Syntax and Semantics* 27, San Diego: Academic Press, pp. 179–274.

23

Distributivity and Negation. The Syntax of *Each* and *Every*

Filippo Beghelli and Tom Stowell

1997

23.1   Introduction

Noun phrases come in several types, each with different morphological, syntactic, and semantic properties. Some consist of a pronoun, or of a proper name, or of a lexical noun in co-occurrence with a demonstrative or a determiner, and denote individuals. Other noun phrases do not denote individuals; in English, they may consist of a single word (e.g. *nothing, everything, something*), or of a lexical noun introduced by a quantifying determiner (e.g. *no, every, each, few, most*). We call them quantificational noun phrases, or quantifier phrases (QPs). One of their distinguishing properties is that they might give rise to so-called scope ambiguities, that is, to more than one interpretation for the sentence in which they occur. This can be easily seen in sentences with multiple quantifiers, as in (1):

(1)   Every student admires some professor. (May 1985)

This sentence has two readings. In one, it means that every student admires a professor, but not necessarily the same one. In the second reading, there is a particular professor that every student admires; here the QP in object position takes scope over the one in subject position ("inverse reading"). May (1977) proposed that these readings arise from the fact that all QPs undergo movement (even when the sentence contains only one QP, as in 'John saw everyone'). This movement, labeled QUANTIFIER RAISING (QR), does not affect the word order because it takes place at the level of Logical Form (LF), a level of syntactic representation that feeds the semantic but not the phonological component. For May, the ambiguity of (1) arises as follows: both QPs raise to a position outside the clause; if *every student* c-commands *some professor*, the first reading arises; if *some professor* c-commands

every student, we get the inverse reading. For May (1977, 1985) QR is a movement operation available to (and obligatory for) all QPs.

In this article, Beghelli and Stowell make the important empirical observation that different types of quantifiers have different scope possibilities. Hence QR does not apply uniformly to all types of quantifiers, and is not obligatory for all of them. Building on insights from Szabolcsi (1994, 1997), they distinguish five major classes of QPs: interrogative QPs (like *what, which student*), negative QPs (like *nobody, no student*), distributive-universal QPs (those headed by *every* and *each*), counting QPs (introduced by *few, fewer than five, more than five*, . . .) and group-denoting QPs (headed by *some, several*, bare numerals). Each class has different logical functions and specific scope-taking possibilities.

Beghelli and Stowell also make a precise proposal aimed to capture the scope possibilities of each of these classes. Building on May's idea that the scope of a QP corresponds to the structure that the QP c-commands, they propose that the members of the different classes of QPs take scope from different structural positions, and therefore have different c-command domains. To implement this proposal, they postulate the existence of a series of functional heads, with a fixed position in the functional sequence, each bearing a particular syntactic feature that matches that of a certain class of QPs (in some cases this feature has a morphological reflex, as, for example, the *n-* morpheme in negative quantifiers). In this view, a QP moves to the specifier of one of these functional heads to check that feature in a configuration of specifier-head agreement.

The architecture of scope-positions postulated in this article receives strong support from Hungarian, where (preverbal) quantifiers occur in their scope position in surface syntax (that is, their scopal order matches their left-to-right linear order). As shown in Szabolcsi (1997), the positions in which these QPs occur closely match the positions of the functional categories independently proposed by Beghelli and Stowell.

23.2   From "DISTRIBUTIVITY AND NEGATION: THE SYNTAX OF *EACH* AND *EVERY*"

1   Introduction

This paper is concerned with the syntax and semantics of quantifier scope construal, focussing on the distributive quantifiers *every* and *each*, and their interaction with negation. Our discussion is based on the theory of the syntax of quantifier scope developed more fully in Beghelli and Stowell (1994) and in Beghelli (1995).

The quantifier *every* has traditionally been analyzed in natural language semantics as the quantifier ∀, familiar from classical logic. We will show that *every* is more complex than this; a number of observations on its logico-semantic behavior lend plausibility to the view that *every* exhibits a kind of quantificational variability characteristic of licensed and bound elements. The quantifier *each* has been analyzed as a wide-scope variant of *every*, which is supposedly used in order to disambiguate between pairs of possible scope construals. We will show that the

distinction between *every* and *each* is more properly characterized in terms of an intrinsic distinction between optional and obligatory distributivity. The effects of this distinction are often masked, however, by the effects of the syntactic mechanisms by which these notions are expressed in the grammar of natural languages, as we will see.

The paper is organized as follows. In Section 2, we introduce the general theory of scope and quantifier types on which the rest of the paper is based. In Section 3, we discuss the syntax of distributivity, concentrating on the distinctive behavior of QPs headed by *every* and *each*, which we refer to as Distributive-Universal QPs (DQPs). In Section 4, we examine the scopal interactions of DQPs with negation, bringing to light certain distinctive properties of these QPs, and highlighting some surprising differences between *every* and *each*. In Section 5, we discuss other differences between *every* and *each*, which we will use to explain the differential behavior that they exhibit with respect to negation.

2   Target scope positions for QP-types

2.1   Scope uniformity

Our analysis of *every* and *each* is formulated within the overall theory of quantifier scope developed in Beghelli and Stowell (1994) and in Beghelli (1995). We present here a sketch of that proposal; the reader is referred to those works for further discussion. We adopt two central assumptions of the standard theory of quantifier scope in generative grammar. First, quantifier scope is determined by c-command relations holding at the level of Logical Form (LF); second, Quantifier Phrases (QPs) are assigned scope by undergoing movement to their scope positions in the derivation of the LF representations.

However, we reject one central assumption that has guided virtually all previous work on scope, namely that all QPs have the same scope possibilities. This can be stated in terms of QUANTIFIER RAISING (QR), as in (1):

(1)   The Uniformity of Quantifier Scope Assignment (Scope Uniformity)
        Quantifier Raising (QR) applies uniformly to all QPs. Neither QR nor any
        particular QP is landing-site selective; in principle, any QP can be adjoined
        to any (non-argument) XP.

In this respect, we depart from the standard account in May (1977, 1985), as well as from refinements of it in Aoun and Li (1989, 1993), and Hornstein (1995).

The reason why Scope Uniformity cannot be maintained is empirical: different QP-types have correspondingly different scope possibilities. Some of the evidence for this conclusion is reviewed below.

May (1977, 1985) assumes that pairs of subject and object QPs are typically scopally ambiguous, and concludes that all QPs normally undergo movement from their (S-structure) Case positions to distinct scope positions at LF. In other words, he assumes that Case positions never serve as scope positions for QPs. On the other hand, Hornstein (1995) proposes that every link in the A-chain of a given QP

is a possible scope position for that QP – including both the Case position occupied by the QP at Spell-Out and its θ-position.

In this study, we propose a hybrid theory, incorporating aspects of both May's and Hornstein's approaches. The central innovative aspect of the system developed here is that it draws distinctions among various QP-types; whereas *certain* QP-types may take scope in their Case positions (remaining *in situ* at LF), other QP-types must move to distinct LF scope positions reserved for them. Moreover, there are further distinctions among those QP-types that must undergo movement, in the sense that each type has a designated LF scope position defined in the hierarchical phrase structure of the clause.

2.2   QP types

Although it is possible, *a priori*, to draw many distinctions among various QP-types, we believe that – in a first approximation – the syntax of quantifier scope can be adequately captured by recognizing five major classes of QP-types. Our classification incorporates insights of Szabolcsi (1994, [1997]). The reader is especially referred to the latter paper, where the relation with our proposal is discussed at length.

QP-Types

a.   Interrogative QPs (WhQPs). These are familiar Wh-phrases such as *what, which man*, etc. We adopt the standard convention of attributing a [+Wh] feature to these QPs, encoding their interrogative force.

b.   Negative QPs (NQPs). These are QPs such as *nobody, no man*, etc. (In this group belong also French n-words such as *personne* 'nobody,' and possibly Italian/Spanish n-words such as *nessuno/nadie* 'nobody,' which sometimes require an overt negative element to license them.) We assume that these QPs bear a feature [+Neg].

c.   Distributive-Universal QPs (DQPs). These are QPs headed by *every* and *each*, which occur only with singular nouns. We attribute to them, in a first approximation, a distributive feature [+Dist(ributive)] (we will revise this assumption in Section 5, where we will attribute to *each* an intrinsic feature of distributivity [+Dist], leaving *every* underspecified for [Dist] and specified merely for universality [+Univ]). Both *each*-QPs and *every*-QPs are usually interpreted as both universal and distributive.

d.   Counting QPs (CQPs). These include decreasing QPs with determiners like *few, fewer than five, at most six*, . . . and generally cardinality expressions built by modified numerals (e.g., *more than five, between six and nine, more (students) than (teachers)*, . . .). The characteristic semantic property of these QPs is that they count individuals with a given property, have very local scope (take scope essentially *in situ*) and resist specific interpretations.

e.   Group-Denoting QPs (GQPs). To this large class belong indefinite QPs headed by *a, some, several*, bare-numeral QPs like *one student, three students*, . . ., and definite QPs like *the students*. The fundamental property of GQPs is that they denote *groups*, including plural individuals. Even leaving aside their ref-

erential reading (the type of epistemic specificity discussed first by Fodor and Sag 1982), GQPs can easily be construed as taking widest scope within their clause, though they might be c-commanded by other scopal elements. We maintain that this capacity for wide scope derives from their ability to introduce group referents. (Another property of GQPs that derives from this is that they support collective interpretations in contexts where DQPs require a distributive construal.) Indefinite and Bare-numeral GQPs can also support readings where they have very local scope, behaving like CQPs. We factor out such readings (exhibited by some of the members of this class) in terms of an ambiguity between a GQP and CQP reading.

2.3   Logical functions associated with QP-types

On the basis of this typology, we identify the following logical functions and relative LF positions where they are satisfied.
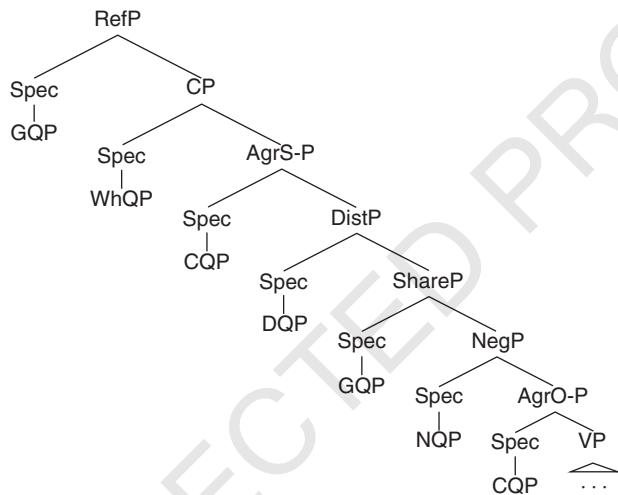
Scope positions for QP types

a.   WhQPs take scope in the Spec of CP, where they assume their interrogative force by virtue of their [+Wh] feature being checked via Spec-Head agreement with the question operator Q.

b.   NQPs take scope in the Spec of NegP, where their [+Neg] feature is checked via Spec-Head agreement with the (silent) Neg^0 head, as in Zanuttini (1991) and Moritz and Valois (1994). Clausal negation with *not*, which we assume involves negative quantification over eventualities or situations, is licensed in the same way.

c.   DQPs headed by *each* and *every* normally move to the Spec position of the Distributive-Universal category DistP, where they undergo Spec-head agreement with the Distributive-Universal head $Dist^0$, resulting in their characteristic interpretation. We will also suggest, however, that *every* can occur in other LF–positions as well, under certain circumstances; details are given in Section 4 and 5.

d.   GQPs may select one of several distinct scope positions, resulting in the different interpretations that they receive:

(i)   GQPs that are referentially independent normally occupy the Spec of RefP position (located above CP), where they fulfill the function of (logical) subject of predication, and are interpreted with widest scope relative to other scope-bearing elements in their clause.

(ii) A lower LF position, accessible by GQPs headed by an indefinite or a bare numeral, as well as QPs containing an externally bound variable, is the Spec of ShareP, which we locate just below DistP. GQPs scoping in this position are interpreted with "dependent" specific reference, in the particular sense of specificity developed by Diesing (1990, 1992), i.e. ranging over individuals whose existence is presupposed. (This allows for a kind of narrow-scope specific reading, discussed below.) Whereas specific *indefinite*

GQPs can occupy either the Spec of ShareP or the Spec of RefP position, specific *definite* GQPs must normally take scope in the Spec of RefP of that clause, and are scopally independent within it.

(iii) Indefinite or bare-numeral GQPs may also take scope in their Case positions (i.e. *in-situ*), where they are interpreted non-specifically, like CQPs.

(e) CQPs cannot ordinarily be interpreted as specific. Therefore they are interpreted in their Case positions and take scope *in-situ*. For a discussion of the properties of CQPs, the reader is referred to Szabolcsi ([1997]).

The relative scope positions of our five QP-types, based on their location in the functional structure of the clause, are given in (2):

Given the well-known lack of island effects with definite and specific indefinite GQPs – which, like indexical pronouns and names, can have a *de re* construal even when they are embedded within islands – it has often been suggested that a wide-scope referential (*de re*) construal does not depend on movement. We will not be concerned here with the issue of how referential readings (cf. Fodor and Sag 1982) of indefinite QPs should be generated. We refer the reader to Kratzer (1995) for a recent proposal.

We assume that true GQPs become associated with an existential operator over a restricted variable, ranging over witness sets of the GQP. This proposal seems to us essentially similar to that contained in Reinhart (1995), where the existential operator ranges over choice functions (cf. Abusch 1994, Beghelli 1993, Beghelli 1995, Ruys 1993, etc. for further discussion).

[. . .]

2.4 Scope and feature-checking

In the system that we propose, the movement of DQPs and GQPs to their scope position is driven by the need to check features that are associated to their QP-types. We will therefore refer to our proposal as a checking theory of scope assignment. We will

return later on in this paper to the precise characterization of some of these features (in particular, to the different featural specification of *every* vs. *each*). Here we simply wish to present the overall picture, and evaluate some of its consequences.

Membership in any of the QP-types listed in Section 2.2 is indicated by a number of syntactic properties, some of which have been mentioned there. These properties are morphologically encoded in the determiner position of the DP or QP: this is obvious in the case of WhQPs and NQPs, as they bear *Wh-* and *n*-markings, but it arguably holds for other QP-types as well.

Thus, the determiners of DQPs (*each*, *every*) have what we may call *e*-morphology. Morphological markings (the presence of un-modified numerals, (in)definite article, etc.) distinguish the various subtypes of GQPs, and CQPs are characterized by the presence of modified numerals. These morphological specifications are not inherently different from the usual ones (agreement, case marking, etc.). We propose that they represent the syntactic encoding of logico-semantic features.

What is special with these, we propose, is that they carry logico-semantic features. WhQPs check their [+Wh] features through Spec-Head agreement with a Wh-operator hosted in C^0, and NQPs check [+Neg] in Spec of NegP, under agreement with the Neg-operator in Neg^0. Let us assume that a similar process obtains with the other QP-types. Feature-checking may appear to be more complex with the latter than it is with the former, but we are interested in pursuing the hypothesis that the process is essentially the same.

Our basic assumption is that DQPs need to check their [+Dist] features under agreement with a distributive operator (which we can indicate as ∀) hosted in $Dist^0$, whereas GQPs need to check group reference ([+group ref]) with an existential operator-head (∃). Existential operator-heads occur in both $Share^0$ and Ref^0. The hierarchy in (2) thus corresponds to a hierarchy of operators. We claim that one of the basic roles served by the functional hierarchy of the clause is to encode the structural order in which semantic information is processed.

This gives the basic idea of what we think is going on in the process of scope assignment: scope is simply the by-product of agreement processes. Within this overall scenario, individual sub-types of QPs (and possibly individual quantifiers) realize additional features. GQPs are not, as a class, assigned a unique landing site: though definites typically take scope in Spec of RefP, numerals and indefinites can move to either RefP or ShareP. Extending the logic of our analysis, we suggest that when a GQP is endowed with an extra feature that marks it as the logical subject of predication, it will be driven to move up to (Spec of) RefP; otherwise it will remain in ShareP. If an indefinite GQP lacks the feature [+group ref] altogether, it behaves like a CQP, i.e. it goes no further than its Case position at LF). Unlike DQPs and GQPs, we assume CQPs do not have syntactically relevant features to check.

On a somewhat more technical level, we assume that scope positions can be reached either directly, through (leftward/upward) movement, or by (rightward/downward) reconstruction to a lower link in the chain of the QP. There is no principled difference between movement and reconstruction: each QP-chain is associated with one scope position, defined as the unique link which is compatible with the featural specification of the QP.

[. . .]

2.6   Empirical justification

We have stressed that the fundamental motivation for our approach is empirical. We will now review some of the empirical justification for the rich structural representation that we hypothesize. We concentrate on interactions between clausemate QPs surfacing in subject and object positions, where one of the QPs is an indefinite GQP. We present only some of the relevant data in this section; further data will be considered in later parts of this paper. Scopal interactions between DQPs and negation (including both clausal negation and NQPs) are considered in Section 3; scopal interactions involving WhQPs are discussed extensively in Beghelli ([1997]). Furthermore, we will make only passing references, in discussing the predictions of our theory, to the scopal behavior of CQPs, since they bear only tangentially on the focus of the present paper; the reader is referred to Beghelli and Stowell (1994) and Beghelli (1995).

2.6.1   Clause-internal scopal asymmetries

We begin our empirical discussion by enumerating [four] predictions implied by the hierarchy of positions in (2):

(3) a. A WhQP should always take wide scope with respect to any other QP in their clause, other than GQPs when these are assigned scope in Spec of RefP.

  b. A GQP should be scopally ambiguous with respect to a clausemate DQP, depending on whether the GQP moves to Spec of RefP or to Spec of ShareP.

  c. A GQP object should be scopally higher than clausal negation, owing to the fact that it takes scope in Spec of ShareP or Spec of RefP – except in the case mentioned above where an indefinite or bare-numeral GQP remains in its Case position (Spec of AgrO-P) and receives a counting interpretation; cf. (diii) in 2.3. A GQP subject should always take wide scope with respect to clausal negation and/or a clausemate NQP.

  d.   A CQP in object position should never be able to take inverse scope over a GQP or DQP occurring in subject position.

Let us now see the empirical status of these predictions, and how they follow from our assumptions. Some of the predictions in (3) are, of course, familiar facts from the literature. For instance, (3a) – that WH-QPs take widest scope – is widely assumed, and we are essentially following a long tradition here. Prediction (3b) – that clausemate GQP/DQP pairs are scopally ambiguous – is also a familiar fact, exemplified in paradigms such as (4):

4 a. Every/Each student read two books.
  b. Two students read every/each book.

In each case, the indefinite GQP headed by *two* can be construed either inside or outside the scope of the DQP headed by *every/each*.

Our account of (4) does not differ empirically from the classical QR-based the-ory advanced by May (1977), although it derives the scopal ambiguity in a different way. The classical theory of May (1977) captures the ambiguity as a result of QR being free to apply sequentially, in either order, to both QPs. Either QP may adjoin to S, creating a higher S-node; then the other QP will adjoin to the higher S-node, taking wider scope than the QP that moved first. Since either QP can be the first to move, two LF-configurations are possible, resulting in the ambiguity. (This analysis could be translated into a Minimalist framework, by allowing both QPs to adjoin to AgrS-P, or by allowing one to adjoin to AgrS-P, and the other to adjoin to some other functional category, such as TP.)

In contrast, the Checking Theory of scope that we are advocating here must claim that the DQP will always end up in the same LF scope position, namely in the Specifier position of the Distributive Phrase (Spec of DistP). Hence the scopal ambiguity must arise in some other way. We suggest that it arises because indefinite GQPs have an ambiguous quantifier type, making more than one LF position available to them; in fact, we suggest that they have *four* possible LF landing sites. One of these – Spec of RefP – is superior to the DQP's position in Spec of DistP; another – Spec of ShareP – is inferior to it. The other two positions are both Case positions (Spec of AgrS-P and Spec of AgrO-P, for subjects and objects, respectively); of these, the latter is inferior to the LF position of the DQP, while the former is superior to it.

Consider now (4b), where an indefinite QP occurs in the subject Case position (Spec of AgrS-P) at Spell-Out, and a DQP occurs in the object position. Since the DQP must move to the Spec of DistP position, which is inferior to the Case position of the subject, a narrow scope construal of the subject will be possible only if the subject reconstructs to a scope position *lower* than Spec of DistP. For the GQP subject in (4b), a narrow scope construal of the subject must involve its reconstructing to the Spec of ShareP position, since it cannot reconstruct to the Spec of AgrO-P. (The possibility of its reconstructing to its θ-position is discussed below.)

The reader may wonder how the Checking Theory of scope can account for sentences containing two DQPs, such as *Each boy read every book* or *Every professor gave every student an A*. If DQPs headed by *each* and *every* have a unique LF landing site, then one might expect that a given sentence could contain only one of them. The analytical problem posed by such examples is no different in principle from that posed by multiple Wh-questions or by sentences containing multiple NQPs, e.g., in languages exhibiting "negative harmony" such as Spanish. For such cases, we follow a long tradition in assuming that the Spec positions of scopal categories can be multiply filled, either because there may be more than one specifier for the same projection, or through a process of absorption applying to quantifiers of the same logical type.

The first prediction in (3c) – that *indefinite GQP objects can take inverse scope over negation* – is also a familiar fact, based on examples like (5a, b):

(5) a. The students didn't read two/some books.
     b. No student read two/some books.

The second prediction in (3c) – the possibility of a *narrow-scope* construal for an indefinite GQP object, as in (5a), follows from our proposal that some (e.g., bare-numeral) GQPs can be interpreted as CQPs and remain in their Case positions at LF, as in 2.2 and 2.3.

Empirical support for the third prediction in in (3c) – that *indefinite GQP subjects must take scope over negation* – is less widely recognized, though it is supported by (6a, b):

(6)　a.　Two/some students didn't read this book.
　　　b.　Two/some students read no books.

Assuming that the LF scope position of both clausal negation and NQPs is located at the NegP level, the *possibility* of a wide-scope construal of indefinite GQP subjects and objects is expected, given that indefinite GQPs have two possible LF landing sites above NegP in (2) – Spec of ShareP and Spec of RefP. (The distinction between these two positions is not obvious in examples like (5) and (6), and may appear at this stage to be an artifact of our account of (4); however, we will provide justification for this shortly.)

However, the GQP subjects in (6) apparently *must* take wide scope relative to negation, suggesting that there is no position below the scope domain of the negative operator (in Spec of NegP) that these subject GQPs can reconstruct to. Our hierarchical arrangement of scope positions provides an account of this, in the spirit of Hornstein (1995). Unlike an object GQP, whose Case position (Spec of AgrO-P) lies *within* the scope of negation, a subject GQP would have to reconstruct to a position within VP in order to derive a narrow-scope construal relative to negation, since the subject Case position (AgrS-P) is too high up. (Reconstruction to the Spec of ShareP can derive a narrow scope construal relative to a distributive operator in DistP, but it is not low enough to produce a narrow scope construal relative to negation.)

Thus, there is only one way in which a narrow-scope construal of a subject GQP relative to negation might be derived: by reconstruction of the subject GQP to its original θ-position below NegP. Evidently this option must be excluded. A natural way of deriving this result would be to assume that every quantifier phrase must syntactically bind a trace as a variable in the LF representation. (Though the semantic basis for such an assumption is not obvious, we will assume nevertheless that such a condition holds, on LF representations, at least.) Then reconstruction of a GQP – or another quantifier phrase – to its original θ-position would be excluded, since there would be no trace in a lower position for the GQP to bind.

Simple indefinites (singular indefinites with the article *a/an* and bare plurals) in subject position do seem to be capable of reconstructing below NegP, however, as in (7):

(7)　a.　A student didn't write this book.
　　　b.　Students didn't write this book.

Furthermore, as is well known, simple indefinites and bare plurals can routinely be bound by generic operators and adverbs of quantification, whereas numerals and *some* do not show this type of variability. We can provide an explanation for the

difference between (6) and (7) if we follow much recent work in assuming that simple indefinites and bare plurals are actually restricted variables which can be unselectively bound by a variety of external quantifiers, including negative quantifiers. This will allow them to reconstruct into a θ-position because, being variables and not quantifier phrases, they do not need to bind variables themselves. Nor do they need to be checked with an operator-head in Spec,ShareP or Spec,RefP for existential quantification, because they are unselectively bound. Hence the contrast between (6) and (7).

Lastly, we should point out that the introduction of a special type for CQPs is motivated by a basic asymmetry in subject-object scope interactions. Whereas both DQPs and GQPs can, when in object position, take wide scope over a subject GQP (though not in the same way – cf. Section 3), CQPs are not able to take inverse scope:

(8) a. Some/one of the students visited more than two girls.
     b. Some/one of the students visited few(er than three) girls.
     c. Every student visited more/fewer than three girls.

In neither of (8a, b, c) can the object QP take scope over the subject (at least if normal intonation is employed). For example, we cannot construe (8a) to mean that for more than two girls, it is the case that some student, or one of the students, visited her.

This is derived directly under our analysis, since an object CQP cannot scope higher than Spec of AgrO-P, and a subject GQP, as seen above, cannot reconstruct lower than Spec of ShareP. (Nor can a subject DQP reconstruct below Spec of DistP.) Our assumptions about the local scope of CQPs are further confirmed by the observation that these QPs only support a *de dicto* reading when they are complements of intensional predicates:

(9) Someone wanted to visit more than two professors.

2.6.2 Cross-linguistic evidence

As a second argument for the Checking Theory of scope, we cite empirical evidence from surface constituent order in a number of languages, supporting our contention that there are distinctive scope positions defined in the phrase structure of the clause for DQPs and (particular construals of) GQPs. The paradigmatic case of one-to-one correlation between surface order and scope seems to be Hungarian, a language known to 'wear LF on its sleeve.' Szabolcsi ([1997]) presents striking evidence in support of the Checking Theory, by showing that, in Hungarian, a hierarchy of positions essentially similar to (2) governs the surface order of QPs. In this language, GQPs, DQPs, and CQPs move in the overt syntax to their specified scope positions in the hierarchy of functional projections in (2).

With respect to DQPs, Kinyalolo (1990) has shown that, in the Bantu language KiLega, universally quantified noun phrases that are obligatorily distributive must undergo overt leftward movement in the visible syntax. We interpret this as evidence that KiLega requires DQPs to be spelled out in Spec of DistP, just as

English requires (most) WhQPs to be spelled out in Spec of CP. Similarly, Khalaily (1995) shows that the Palestinian Arabic counterparts of our DQPs must undergo leftward movement in the overt syntax in a parallel fashion; he argues that Palestinian Arabic exhibits an overt counterpart to our LF movement to Spec of DistP, a conclusion that we concur with.

Further cross-linguistic evidence comes from the recent literature on scrambling in Hindi (Mahajan 1990) and various Germanic languages (cf. Kratzer 1988 and Diesing 1990, among others). A number of proposals have suggested that specific construals of indefinites are necessarily associated with (overt) leftward movement out of VP. Though the exact location of the landing site of scrambling is still being debated, we believe that the position that we identify as Spec of ShareP is a common landing site for scrambling. We will not develop this point here, however, since this would take us too far afield.

23.3   Questions pertaining to Beghelli and Stowell (1997)

1   The notion of (non-)specificity plays a role in Beghelli and Stowell's paper. Discuss the merits of the following proposal: The semantic distinction concerning specificity has a syntactic counterpart, to the effect that all specific indefinite DPs must contain either overt *certain* or silent CERTAIN.

2   In distinguishing "strong distributivity" (found with *each*) from "pseudo-distributivity" (found with *all*), Beghelli and Stowell use the idea that when *all* participates in distributivity it does so via the presence of a silent counterpart of *each*. What might be the licensing conditions on such a silent counterpart of *each*? Bring in Heim et al. (1991) and Kayne (2003).

3   To what extent does *all* have the same behavior as *all three, all four*, etc.?

4   As Beghelli and Stowell discuss, *every* often acts like a distributive element, as shown by *?Every student in the class left together*, where it is pretty much incompatible with collective *together*. Yet *Everybody in the class left together* is appreciably more acceptable. Why might that be? (Bring in Leu 2005 and references cited there.)

5   English distinguishes *everybody* from *eachbody* and *allbody*. Is the restriction against *eachbody* the same as the restriction against *allbody?* Bring to bear evidence from other languages. (Extra credit: Why is *everyhow* (much) less widely accepted than *somehow, anyhow*, and *nohow?*)

6   Discuss differences and similarities between *all* and *whole*. (Extra credit: Find differences between *whole* and *entire*.)

7   Sentences like *Our top three students have published eleven papers this year* are usually said to be ambiguous between a distributive reading (in which thirty-three papers are published) and a nondistributive reading (in which just eleven papers are published). On the distributive reading, Beghelli and Stowell would have such sentences contain a silent EACH. Discuss the merits and demerits of having such sentences, on their nondistributive reading, contain a silent TOTAL, for example, as in *our top three students have published A TOTAL OF eleven papers this year*.

8   How exactly might Beghelli and Stowell try to accommodate sentences like *Philosophy students are taking fewer/less than six of our courses this semester*, which seems to have a natural interpretation according to which the object (a counting QP, in their terms) scopes over the bare plural subject?

9   Beghelli and Stowell account for the inverse scope reading of sentences like *A different student called every professor* by having the phrase *every professor* move to Spec,DistP at LF. This approach leaves open the contrast between *A different student called up every professor*, which has a similar natural inverse scope reading, and *A different student called every professor up*, which does not. How might they try to account for this effect, which seems to be due to the post-DP particle? Evaluate the relative merits of such an account as compared with that of Kayne (1998), in which *every professor*, in the inverse reading, moves overtly rather than at LF.

10 Beghelli and Stowell propose a similar account for the reading of *Every boy didn't read one book* in which *one book* scopes over negation. *One book* moves at LF to Spec,ShareP, a position that c-commands negation. The following contrast (in the relevant reading) suggests, though, that in such readings the indefinite *one book* must move to its Spec,ShareP overtly (rather than at LF): *Every student didn't look up one/some word* vs. ?*Every student didn't look one/some word up*. How might they attempt to integrate these facts?

11 Pinpoint the similarities and differences between this paper and Rizzi (1997).

12 Postma and Rooryck (1996) take English phrases like *her every thought* to have *ever-* originate within a reduced relative whose head is *thought*, with *-y* possibly being a distributive quantifier. Discuss the similarities and differences between their paper and Jayaseelan (2011).

13 Postma and Rooryck (1996) do not extend their reduced relative analysis of *her every thought* to apparently simpler phrases such as *every car*. Try to do so by combining their ideas with those of Koopman (2003, 2005).

14 A bit like *her every thought* in English is *his all* as in *He gave it his all*. Discuss the pros and cons of taking *his all* to contain a silent noun. How might that allow integration with Postma and Rooryck (1996)? Discuss, in addition, the question of the (in)compatibility of Postma and Rooryck's paper with this paper by Beghelli and Stowell.

15 Although Beghelli and Stowell take sentences like *Each student read two books* to be scopally ambiguous, the object wide scope reading is for many speakers quite marginal. Better, though still not entirely natural, is the inverse scope reading of sentences like *Each student read two of these books*, with an overt definite contained within the object. What might be the reason for this improvement?

References

Abusch, D. 1994. The scope of indefinites. *Natural Language Semantics* 2: 83–135.
Aoun, J., and A. Li. 1989. Scope and constituency. *Linguistic Inquiry* 20: 141–172.
Aoun, J., and A. Li. 1993. *The Syntax of Scope*. Cambridge, MA: MIT Press.

Beghelli, F. 1993. A minimalist approach to quantifier scope. In A. J. Schafer (ed.), *Proceedings of NELS 23*, Amherst: GLSA Publications, University of Massachusetts, pp. 65–80.

Beghelli, F. 1995. *The phrase structure of quantifier scope*. Doctoral dissertation, UCLA.

Beghelli, F. 1997. The syntax of distributivity and pair–list questions. In A. Szabolcsi (ed.), *Ways of Scope Taking*, Dordrecht: Kluwer, pp. 349–408.

Beghelli, F., and T. Stowell. 1994. The direction of quantifier movement. *GLOW Newsletter* 32: 56–57.

Diesing, M. 1990. *The syntactic roots of semantic partition*. Doctoral dissertation, University of Massachusetts, Amherst.

Diesing, M. 1992. *Indefinites*. Cambridge, MA: MIT Press.

Fodor, J., and I. Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5: 355–398.

Heim, I., H. Lasnik, and R. May. 1991. Reciprocity and plurality. *Linguistic Inquiry* 22: 63–101.

Hornstein, N. 1995. *The Grammar of Logical Form from GB to Minimalism*. Oxford: Blackwell.

Jayaseelan, K. A. 2011. Comparative morphology and quantifiers. *Lingua* 121: 269–286.

Kayne, R. S. 1998. Overt vs. covert movement. *Syntax* 1: 128–191.

Kayne, R. S. 2003. Person morphemes and reflexives in Italian, French and related languages. In C. Tortora (ed.), *The Syntax of Italian Dialects*, New York and Oxford: Oxford University Press, pp. 102–136.

Khalaily, S. 1995. QR and the minimalist theory of syntax: The case of universally and negative quantified expressions in Palestinian Arabic. Ms., University of Leiden.

Kinyalolo, K. 1990. *Syntactic dependencies and the spec–head agreement hypothesis in KiLega*. Doctoral dissertation, UCLA.

Koopman, H. 2003. Inside the noun in Maasai. In A. Mahajan (ed.), *Syntax at Sunset 3: Head Movement and Syntactic Theory. UCLA Working Papers in Linguistics 10*, Los Angeles: University of California, Los Angeles, Linguistics Department, pp. 77–115.

Koopman, H. 2005. On the parallelism of DPs and clauses: Evidence from Kisongo Maasai. In A. Carnie, H. Harley, and S. A. Dooley (eds), *Verb First: On the Syntax of Verb–initial Languages, Linguistik Aktuell/Linguistics Today 73*, Amsterdam and Philadelphia: John Benjamins, pp. 281–302.

Kratzer, A. 1988. Stage–level and individual–level predicates. In M. Krifka (ed.), *Genericity in Natural Language*, Tübingen: University of Tübingen, pp. 247–284.

Kratzer, A. 1995. Scope or pseudoscope? Are there wide-scope indefinites? In S. Rothstein (ed.), *Events and Grammar*. Dordrecht: Kluwer.

Leu, T. 2005. Something invisible in English. In S. Arunachalam, T. Scheffler, S. Sundaresan, and J. Tauberer (eds), *Penn Working Papers in Linguistics: Proceedings of PLC 28*, vol. 11, pp. 143–154.

Mahajan, A. 1990. *The A/A–bar distinction and movement theory*. Doctoral dissertation, MIT.

May, R. 1977. *The grammar of quantification*. Doctoral dissertation, MIT.

May, R. 1985. *Logical Form*. Cambridge, MA: MIT Press.

Moritz, L., and D. Valois. 1994. LF pied–piping and Specifier–Head agreement. *Linguistic Inquiry* 25: 667–707.

Postma, G., and J. Rooryck. 1996. Modality and possession in NPs. In K. Kusumoto (ed.), *Proceedings of NELS 26*, Amherst: GLSA, University of Massachusetts, pp. 273–287.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 281–337.

Ruys, E. 1993. *The scope of indefinites*. Doctoral dissertation, Research Institute for Language and Speech, Utrecht.

Szabolcsi, A. 1994. Quantifiers in pair-list readings and the non-uniformity of quantification. In P. Dekker and M. Stokhof (eds), *Proceedings of the Ninth Amsterdam Colloquium*, Amsterdam: ILLC/Department of Philosophy, University of Amsterdam. pp. 646–664.

Szabolcsi, A. 1997. Strategies for scope taking. In A. Szabolcsi (ed.), *Ways of Scope Taking*, Dordrecht: Kluwer, pp. 109–154.

Zanuttini, R. 1991. *Syntactic properties of sentential negation: A comparative study of Romance languages*. Doctoral dissertation, University of Pennsylvania.

24

The Fine Structure of the Left Periphery

Luigi Rizzi

1997

24.1 Introduction
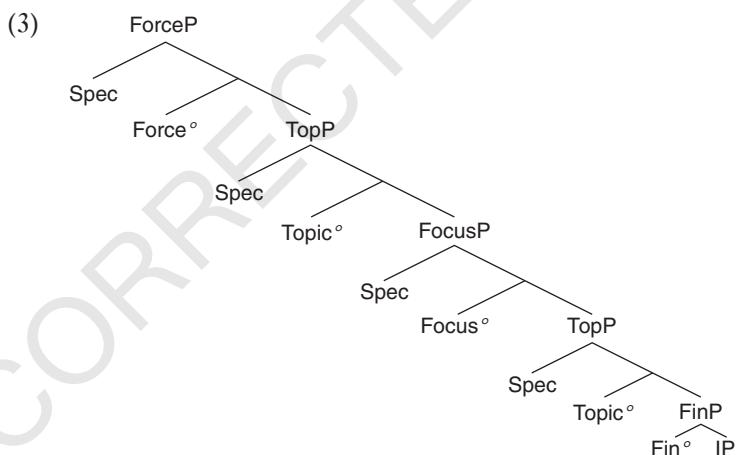
We typically think of clauses as consisting of the predicate, its arguments, and possibly a number of modifiers. But syntactic analysis since the early 1970s has been suggesting that, in thinking about the structure of a clause, we must also allow for an area in its "left periphery" to which elements move. This area has been viewed as consisting of structure that is higher than the elements that make up a clause (i.e., higher than the predicate, its arguments, and any modifier). In English, for example, where the canonical word order is SVO, the left periphery is taken to be an area structurally higher than the subject, where we can find an overt complementizer (e.g. *that* in (1a)), a finite auxiliary or modal in questions (e.g. *has* in (1b)), or a constituent that is being contrasted (e.g. *Mary* in (1c)). Sometimes we find two elements in the left periphery, a phrase and a head, as in the case of *wh*-questions in English (2a) or so-called Verb Second (V2) clauses in other Germanic languages (2b):

(1) a. I'm sure that she left.
     b. Has Mary already left?
     c. MARY I like, but not her parents.
(2) a. What was she doing there?
     b. Dieses Buch hat sie noch nicht gelesen. (German)
         this   book has she still not   read
         'She hasn't read this book yet.'

Within the X-bar conception of phrase structure, these observations were accommodated by proposing that the left periphery consists of a functional projection, CP

(Complementizer Phrase, Chomsky 1986). The head of CP hosts the complementizer and is the position to which a finite verb might move (e.g. in questions, or in V2 clauses); the specifier of CP is the position to which a variety of phrasal constituents might move (e.g., *wh*-phrases, constituents that express the topic or the focus of the clause).

Rizzi's *The Fine Structure of the Left Periphery* builds on these insights and goes far beyond them, providing a detailed and insightful analysis of the left periphery of the clause. Rizzi argues for an articulated CP-layer with a number of strictly ordered functional projections (outlined in (3)), each defined in terms of a certain feature. Elements from within the clause move to these positions to satisfy so-called *criterial* requirements (cf. Rizzi 1996; Haegeman 1995), that is, the need for a head bearing a certain feature to have in its specifier a maximal projection that bears that same feature, and, conversely, the need for an operator with a certain feature to be in the specifier of a head bearing that same feature (cf. also, in this volume, Haegeman and Zanuttini 1991, Ch. 17, and Beghelli and Stowell 1997, Ch. 23). Rizzi proposes that the highest projection of the CP-layer specifies force (here a notion akin to that of clause typing, distinguishing, for example, declarative from interrogative clauses) and is accessible to selection from above. The lowest projection specifies finiteness (whether the embedded clause bears finite or nonfinite tense) and connects the CP-layer to the IP-layer that it dominates (the clause). The ForceP and the FinP are the core components of the complementizer system. Sandwiched in between are other functional projections, to which constituents that are the topic of the focus of the sentence can move:

(3)

The empirical basis of Rizzi's article comes mainly from observations concerning the relative order of constituents in the left periphery in Italian, French, and English. The proposal that the left periphery of the clause contains one projection for a focus phrase, but more than one for topics, is based on sentences of the type in (4), where the focused demonstrative is both preceded and followed by a topic:

(4) A Gianni, QUESTO, domani, gli   dovrete   dire.
     to Gianni, THIS,   tomorrow, him should.2sg say
   'THIS you should tell Gianni, tomorrow.'

Benincà and Poletto (2004) refine this particular aspect of Rizzi's proposal, examining evidence from standard Italian and other Romance varieties spoken in Italy. They propose that the structure of the left periphery between Rizzi's ForceP and FinP in fact consists of two "subfields," each containing a finite set of functional projections that have different semantic properties. The projections in the lower subfield exhibit characteristic properties of operator elements, and host different kinds of focused elements; in contrast, those in the higher subfield do not exhibit properties of operators and host different kinds of topic elements.

The Fine Structure of the Left Periphery is one of the cornerstones of what is now known as the *cartographic* approach to phrase structure, a research program that has been implemented in work on a large and diverse set of languages and in different domains of syntax, and that has brought to light highly articulated universal sequences of elements. Particularly noteworthy is Cinque's (1999) work on the order of adverbs and corresponding functional heads. Representative of much other work in the cartographic tradition are the papers published in the dedicated cartography volumes on the DP and the IP (Cinque 2002), the CP and the IP (Rizzi 2004), and the PP (Cinque and Rizzi 2010).

24.2   From "THE FINE STRUCTURE OF THE LEFT PERIPHERY"

1   Introduction

Under current assumptions, the structural representation of a clause consists of three kinds of structural layers, each layer an instantiation of the X-bar schema:

(1)   The lexical layer, headed by the verb, the structural layer in which theta assignment takes place.
(2)   The inflectional layer, headed by functional heads corresponding to concrete or abstract morphological specifications on the verb, and responsible for the licensing of argumental features such as case and agreement.
(3)   The complementizer layer, typically headed by a free functional morpheme, and hosting topics and various operator-like elements such as interrogative and relative pronouns, focalized elements, etc.

In the mid eighties, each layer was identified with a single X-bar projection (VP, IP, CP), but this assumption quickly turned out to be too simplistic. Under the impact of Pollock's (1989) influential analysis of verb movement, IP dissolved into a series of functional projections, each corresponding to a single feature specification overtly or abstractly expressed on the verbal system (Agr, T, Asp, . . .). Kayne's (1984) binary branching hypothesis naturally led to the postulation of multiple VP layers for multi-argument verbs, e.g. along the lines of Larson (1988) and much related work.

Various proposals in the recent literature indicate that the complementizer layer should share the same fate: much more than a single X-bar schema seems to constitute the left (pre-IP) periphery of the clause.

In this article, I would like to explore some aspects of the fine structure of the left periphery. The first part (sections 2–6) is devoted to the identification of the basic configurational structure. Four kinds of elements typically occurring in the left periphery will be taken into account: interrogative and relative pronouns, topics and focalized elements. Studying the interactions between these elements, we will be led to postulate an articulated array of X-bar projections which will be assumed to constitute the complementizer system. The second part (sections 7–12) concerns a number of adjacency and anti-adjacency effects involving elements of the C system and different kinds of fillers of the subject position (overt DP, PRO, trace) which are amenable to an explanation in terms of the assumed structure of the C system. The core of the empirical material to be discussed is drawn from Italian, French and English, with occasional comparative extensions to other Romance and Germanic languages.

A preliminary word on the theoretical framework adopted in this work is necessary. An idea borrowed from the system presented in Chomsky (1993) will play a crucial role: syntactic movement (or, more neutrally, the formation of non-trivial chains in syntax) is "last resort" in the precise sense that it must be triggered by the satisfaction of certain quasi-morphological requirements of heads. As I will be concerned with the A′ system, I will phrase such requirements in the style of the Criteria (Rizzi 1991, Haegeman 1995 and much related work), rather than as feature checking, the main reason for this choice being that such features have an interpretive import (Wh, Neg, Top, Foc, . . .): they determine the interpretation of the category bearing them and of its immediate constituents (e.g., see section 3), function as scope markers for phrases with the relevant quantificational force in a local configuration, etc. so that their role cannot simply be to trigger movement and disappear from representations. Independently from the particular style of presentation, the "last resort" intuition provides the conceptual justification for postulating a rich and articulated structure to host the different kinds of phrases moved to the left periphery: no free preposing and adjunction to IP is permissible, all kinds of movements to the left periphery must be motivated by the satisfaction of some criterion, hence by the presence of a head entering into the required Spec-head configuration with the preposed phrase. So, the "last resort" guideline will be critical for drawing the map of the left periphery; the presence and action of the system of heads involved will be independently detected by the various adjacency and anti-adjacency effects that we will focus on in the second part. A restrictive theory of adjunction (à la Kayne (1994) and related work) is also instrumental for this endeavor.

On the other hand, in the following discussion I will continue to assume that Relativized Minimality (RM) is a representational principle, and that one of the core structural relations allowed by UG is head government, as in Rizzi (1990) and contra Chomsky (1993). As for the second point, head government continues to be needed, as far as I can see, for optimally simple accounts of various familiar subject-object asymmetries of the *that*-t kind, as well as for many cases in which a head enters into some kind of "action at a distance" with the specifier of its complement (for Case assignment/checking or the licensing of different kinds of ec's). A number of examples of this sort are analyzed in what follows; we will adopt approaches based on head government and will occasionally allude to

properties of possible alternatives not referring to head government, even though no systematic comparison will be attempted.

[. . .]

2 The force-finiteness system

One important question to be asked at the outset of a study on the complementizer system is: what is the role of the complementizer in the clausal structure?

We can think of the complementizer system as the interface between a propositional content (expressed by the IP) and the superordinate structure (a higher clause or, possibly, the articulation of discourse, if we consider a root clause). As such, we expect the C system to express at least two kinds of information, one facing the outside and the other facing the inside.

Consider first the information looking at the higher structure. Complementizers express the fact that a sentence is a question, a declarative, an exclamative, a relative, a comparative, an adverbial of a certain kind, etc., and can be selected as such by a higher selector. This information is sometimes called the clausal Type (Cheng 1991), or the specification of Force (Chomsky 1995).

[. . .]

The second kind of information expressed by the C system faces the inside, the content of the IP embedded under it. It is a traditional observation that the choice of the complementizer reflects certain properties of the verbal system of the clause, an observation formalized, e.g., by "agreement" rules between C and I, responsible for the co-occurrence of *that* and a tensed verb, of *for* and an infinitive in English (Chomsky and Lasnik 1977), etc. A straightforward manner to account for these dependencies would be to assume that C contains a tense specification which matches the one expressed on the lower inflectional system (an idea which goes back at least to Den Besten (1977)). On the other hand, the "temporal" properties encoded by C are very rudimentary. For instance, in Italian the form *che* co-occurs with present, past and future indicative, with present and past subjunctive and past conditional, thus distinguishing these forms from infinitival, gerundival and participial clauses, a situation which is quite general in Romance and Germanic. So, it appears that, at least in these language families, C expresses a distinction related to tense but more rudimentary than tense and other inflectional specifications on the verbal system: finiteness.

I will assume here that the finiteness distinction is a valid linguistic one, even though its morphological realization can vary somewhat from language to language. Languages tend to split verbal paradigms into two classes of forms. Finite forms can manifest mood distinctions (indicative, subjunctive, conditional and/or other distinctions of the realis/irrealis type), manifest tense and subject (person) agreement, cooccur with overt nominative subjects. Non-finite forms do not manifest mood distinctions, in the core case they do not express person agreement, and do not co-occur with nominative subjects, they have a more rudimentary system of tense distinctions (e.g., in many languages non-finite forms do not have a morphological present/future distinction, can express past only through the periphrastic

form aux+past participle, etc.). The first class of forms co-occurs with complementizers of the *that* kind, the second does not. Various dissociations from these core clusters are apparently tolerated, but a split along these lines is robustly attested cross-linguistically.

Following much recent work (e.g., Holmberg and Platzack 1988), I will then assume that the C system expresses a specification of finiteness, which in turn selects an IP system with the familiar characteristics of finiteness: mood distinctions, subject agreement licensing nominative case, overt tense distinctions (these specifications being subjected to some cross-linguistic variation, as we have seen).

Again, we should think of finiteness as the core IP-related characteristics that the complementizer system expresses; languages can vary in the extent to which additional IP information is replicated in the complementizer system: some languages replicate mood distinctions (special subjunctive complementizers in Polish, etc.), some replicate subject agreement (different Germanic varieties; Haegeman 1992, Bayer 1984, Shlonsky 1994), some seem to express genuine tense distinctions (Irish, Cottell 1994), negation (Latin, Celtic), etc.

How does the CP system relate to the rest of the clausal structure? Recent proposals consider the IP system an extension of the V system: the different inflectional heads are V-related in that they attract the verb (overtly or covertly) to check its morphological specification (Chomsky 1993), so that the whole IP system can be seen as an extension of the verbal projection (an "extended projection", in Grimshaw's (1991) sense). Should the CP system be considered an analogous extension of the IP system, hence ultimately of the VP? I believe there is a substantial difference between the two cases. Whatever "inflectional" properties C reflects, they are not encoded in the form of verbal morphology, in the general case: they are expressed on free functional morphemes (*that*, *que*, etc.) which, if anything, look nominal more than verb–like, as they often resemble demonstrative pronouns, *wh* elements, certain kinds of nouns ("fact", etc.), etc. So, I will continue to assume that the C system is fundamentally distinct from the I system, the latter but not the former being V-related in the general case.

3   The topic-focus system

If the force-finiteness system expresses the selectional relations between a C system and the immediately higher and lower structural systems, the C system can have other functions which are by and large independent from selectional constraints.

A traditional articulation of the clause that typically involves the left periphery is the articulation in topic and comment, as expressed by the English construction referred to as Topicalization:

(1)   Your book, you should give t to Paul (not to Bill)

The topic is a preposed element characteristically set off from the rest of the clause by "comma intonation" and normally expressing old information, somehow available

and salient in previous discourse; the comment is a kind of complex predicate, an open sentence predicated of the topic and introducing new information.

Formally similar but interpretively very different is the focus-presupposition articulation:

(2)   YOUR BOOK you should give t to Paul (not mine)

Here the preposed element, bearing focal stress, introduces new information, whereas the open sentence expresses contextually given information, knowledge that the speaker presupposes to be shared with the hearer (see below for further refinements). If the interpretive relation of the preposed element to the open sentence is very different, virtually the opposite in the two cases, the form of the two articulations appears to be constant in English (even though significant differences emerge at a more refined analysis: see Culicover's (1992) discussion, based in part on Gundel's (1974) earlier analysis, and, on focus, Rochemont and Culicover (1990)).

Other languages sharply distinguish the form of the two articulations as well. We will briefly analyze here two Italian constructions which illustrate the point. In Italian, and more generally in Romance, the topic-comment articulation is typically expressed by the construction that Cinque (1990) has called Clitic Left Dislocation (CLLD), involving a resumptive clitic coreferential to the topic (this construction differs from left dislocation in languages which do not possess clitic forms in a number of respects, so that the English gloss, involving a non-clitic resumptive pronoun, is somewhat misleading: see Cinque (1990: 57–60) for relevant discussion; see also Cecchetto (1994), Iatridou (1991)):

(3)   Il tuo libro, lo ho letto
   "Your book, I have read it"
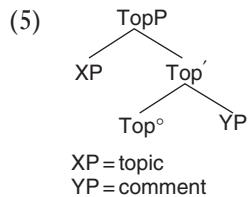
The focus-presupposition articulation can be expressed in Italian by preposing the focal element (focalization) and assigning it special focal stress:

(4)   IL TUO LIBRO ho letto (, non il suo)
      "Your book I read (, not his)"

In Italian this structural option is restricted to contrastive focus, i.e., (4) presupposes that you believe that I have read something different from your book, and corrects this belief. It could not be felicitously uttered as conveying non-contrastive new information, i.e. as an answer to the question "What did you read?". Other languages use the clause initial focus position for non-contrastive focus as well (Hungarian: Kiss (1987), Horvath (1985), Brody (1990, [1995]), Puskas (1992) and references quoted there; Albanian: Turano (1995), Greek: Tsimpli ([1994])). Some other languages (e.g. French) do not seem to use a structural focus position, at least in the overt syntax (Spanish seems to have a focus construction similar to the Italian one: Laka (1990)).
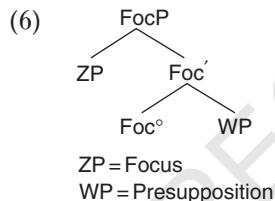
I will assume here that these two articulations–are expressed by the usual building block of syntactic representations: the X-bar schema (whether the schema is a

primitive, or can be derived from more elementary principles (Kayne 1994, Chomsky 1995) is irrelevent [sic] for our purposes). I.e., topic-comment has the following structure:

(5)

XP = topic
YP = comment

A Topº head, a functional head belonging to the complementizer system, projects its own X-bar schema with the following functional interpretation: its specifier is the topic, its complement is the comment. Topº defines a kind of "higher predication", a predication within the Comp system; its function is thus analogous to the function of AgrS within the IP system, which also configurationally connects a subject and a predicate. The most basic difference between higher and lower predication is that the former involves a specifier which is an A′ position.

Analogously, a Focº head takes the focus as its specifier and the presupposition as its complement:

(6)

ZP = Focus
WP = Presupposition

Here too Italian seems to possess a lower focalization, involving focal stress (possibly contrastive, but not necessarily so) on an element *in situ* (see Antinucci and Cinque (1977), Belletti and Shlonsky (1995), Calabrese (1982), Cinque (1993)):

(7)   Ho letto IL TUO LIBRO (, non il suo)
       "I read YOUR BOOK, not his"

But it is conceivable that at LF (7) will have a representation involving (6) if the focal element must be moved to a peripheral position, as Chomsky's (1976) classical analysis of Weak Cross-over implies.

While Topº and Focº are phonetically null in Italian, they may be pronounced in other languages. For instance, Aboh (1995) argues that the focus particle *wè* in Gungbe should be analyzed as Focº, an analysis immediately plausible for many other cases of such markers found across languages (we will not analyze here other constructions involving focalization such as clefts and inverse copular sentences (Moro 1995)).

   [. . .]

How is the topic–focus system integrated into the force-finiteness system? We think of the latter as the essential part of the C system, so we assume it to be present in all non-truncated clausal structures (i.e., except in ECM and other "S′ deletion" contexts). On the other hand, it is reasonable to assume that the topic-focus system is present in a structure only if "needed", i.e. when a constituent bears topic or focus features to be sanctioned by a Spec-head criterion. If the topic-focus field is activated, it will inevitably be "sandwiched" in between force and finiteness, as these two specifications must terminate the C system upward and downward, in order to meet the different selectional requirements and properly insert the C system in the structure. So, we should have:

(8)   . . . Force . . . (Topic) . . . (Focus) . . . Fin IP

We will see later on that this positional property of the topic-focus system is instrumental for the explanation of several adjacency and anti-adjacency effects. For the time being we can simply observe two straightforward empirical reflexes of the theory of C that is taking shape.

In Italian, and more generally in Romance, prepositional elements introducing infinitives such as *di* in (9)b are generally considered the non-finite counterparts of the finite complementizer *che* of (9)a (see Kayne 1984, Rizzi 1982 for relevant evidence); still *che* always precedes and *di* always follows a left-dislocated phrase (examples like (11)b are slightly marked if compared to the corresponding cases of CLLD with finite embedded sentences, but the contrast with (11)a is very sharp):

(9)   a.   Credo che loro apprezzerebbero molto il tuo libro
             "I believe that they would appreciate your book very much"
       b.   Credo di apprezzare molto il tuo libro
             "I believe 'of' to appreciate your book very much"
(10)   a.   Credo che il tuo libro, loro lo apprezzerebbero molto
             "I believe that your book, they would appreciate it a lot"
       b.   *Credo, il tuo libro, che loro lo apprezzerebbero molto
             "I believe, your book, that they would appreciate it a lot"
(11)   a.   *Credo di il tuo libro, apprezzarlo molto
             "I believe 'of' your book to appreciate it a lot"
       b.   Credo, il tuo libro, di apprezzarlo molto
             "I believe, your book, 'of' to appreciate it a lot"

This distribution is hardly consistent with a theory assuming a unique C position, while it can be immediately expressed within the current articulated theory of C by assuming that *che* manifests the force position, while *di* manifests the finiteness position, hence they show up on opposite sides of the topic. We will come back to this peculiar distribution in section 6.

A similar type of argument is provided by the distribution of different kinds of operators hosted by the C-system. In Italian, relative operators must precede

topics, while question operators must follow topics in main questions and can follow or (slightly marginally) precede them in embedded questions:

(12) a. Un uomo a cui, il premio Nobel, lo daranno senz'altro
          "A man to whom, the Nobel Prize, they will give it undoubtedly"
      b. *Un uomo, il premio Nobel, a cui lo daranno senz'altro
          "A man, the Nobel Prize, to whom they will give it undoubtedly"
(13) a. *A chi, il premio Nobel, lo daranno?
          "To whom, the Nobel prize, will they give it?"
      b. Il premio Nobel, a chi lo daranno?
          "The Nobel prize, to whom will they give it?"
(14) a. Mi domando, il premio Nobel, a chi lo potrebbero dare
          "I wonder, the Nobel Prize, to whom they could give it"
      b. ? Mi domando a chi, il premio Nobel, lo potrebbero dare
          "I wonder to whom, the Nobel Prize, they could give it"

This distribution suggests that relative operators occupy the highest specifier position, the Spec of Force, while question operators can occupy a lower position within the Topic/Focus field (the ordering in (13)a being blocked by the fact that I to C movement is compulsory in main questions (Rizzi 1991)). See below for more detailed discussion of these positional properties. The crucial point here is again that a theory involving a unique C head and projection does not seem equipped to deal with such simple distributional constraints.

   [. . .]

9   Adjacency effects on traces

If Topicalized elements involve an independent X-bar projection in the C system, we expect that the presence of a topicalized element will interfere with subject extraction, under standard assumptions on the licensing of traces. In fact, we find two opposite, almost contradictory, kinds of interactions: some preposed elements block subject extraction; other preposed elements alleviate *that*-trace violations and make subject extraction possible. We shall call these two effects adjacency and anti-adjacency effects on traces, respectively. Let us concentrate on the former in this section.

First of all, intervening CLLD phrases induce subject-object asymmetries in French:

(66) a. ? Je ne sais pas à qui, ton livre, je pourrais le donner t
          "I don't know to whom, your book, I could give it t"
      b. *? Je ne sais pas qui, ton livre, t pourrait l'acheter
          "I don't know who, your book, t could buy it"
(67) a. ? Un homme à qui, ton livre, je pourrais le donner t
          "A man to whom, your book, I could give it t"
      b. *? Un homme qui, ton livre, t pourrait l'acheter
          "A man who, your book, t could buy it"

A verbal complement can be moved across a Topic with slightly marginal results in questions and relatives; movement of a subject across a Topic determines a clear decrease of acceptability. These asymmetries are obviously reminiscent of the familiar subject-complement extraction asymmetries (*that*-trace effects, which produce somewhat sharper contrasts; on this, see below):

(68) a. A qui crois-tu que Marie va parler t?
          "To whom do you believe that Marie is going to speak t?"
      b. *Qui crois-tu que t va parler à Marie?
          "Who do you believe that t is going to speak to Marie?"

Comparative evidence supports the hypothesis that (66)–(67) are parallel to (68). Italian, a language which does not show subject-complement asymmetries of the kind illustrated in (68) (ultimately as a function of the positive fixation of the Null Subject Parameter, see Rizzi 1982, ch. 4, 1990: 62–65 for discussion), also allows subject and complement extraction across a Topic at the same level of acceptability:

(69) a. Un uomo a cui, il tuo libro, lo potremmo dare
          "A man to whom, your book, we could give it"
      b. Un uomo che, il tuo libro, lo potrebbe comprare
          "A man who, your book, could buy it"

  [. . .]

Consider now the structural representation of such examples as (66)b, etc. under the X-bar analysis of Topics (C here is whatever head of the complementizer system has the Wh element in its Spec in indirect questions):

(73) Je ne sais pas [qui C° [ton livre Top° . . . [t pourrait . . .]]]

Here, even if C is turned into a governor via agreement, it is too far away to license the subject trace, due to the intervening head Top°, a standard case of Relativized Minimality effect. If Fin, lower than Top, is endowed with Agr features, things do not change: the trace in subject position t would be well-formed, but the subject should move through the Spec of Fin to license the Agr features on this head, and the trace in the Spec of Fin, t′, would now be the offending trace:

(74) Je ne sais pas [qui C° [ton livre Top° [t′ Fin+Agr [t pourrait . . .]]]]

(66)b with representation (74) is close enough to (68)b to make it possible to appeal to the same structural explanation for both cases of subject/non-subject asymmetries; on the other hand, it is different enough to leave room for an account of the different strength of the effect with respect to ordinary *that*-trace effects.

  [. . .]

We have seen in (69) that Italian differs from French in that CLLD does not induce subject-object asymmetries in A′-chains, a fact that is amenable to other similar contrasts between the two languages as a consequence of the different fixation of the Null Subject Parameter. On the other hand, adjacency effects are detectable in Italian if we look at A-chains. Here the relevant contrast is between raising and control: control infinitives are compatible with CLLD (with the dislocated element preceding the infinitival complementizer, as we have seen), while raising infinitives are not: compare, in particular, the different behavior of the control and raising use of *sembrare* (seem):

(78)   a.   Gianni pensa, il tuo libro, di PRO conoscerlo bene
              "Gianni thinks, you [sic] book, of to+know it well"
        b.   Mi sembra, il tuo libro, di PRO conoscerlo bene
              "It seems to me, your book, of to+know it well"
        c.   *? Gianni sembra, il tuo libro, t conoscerlo bene
              "Gianni seems, your book, to know it well"

If raising infinitives must involve a bare IP in order to allow the subject trace to be properly governed by the main V, there is no room for a TopP to occur in such structures; on the other hand, control infinitives can (and must) involve a CP system, so that they are compatible with a TopP.

A more subtle case of incompatibility with a dislocated phrase is provided by the special Romance construction involving Wh extraction of the subject from the infinitival complement of an epistemic verb, a complement which does not allow an overt subject *in situ* (Kayne 1984, Rizzi 1982, 1990):

(79)   Un uomo che ritengo (*a Gianni,) potergli parlare
        "A man whom I believe (to Gianni) to be able to talk to him"

Here a C structure (say, minimally, a –Fin head) is needed to ensure, on the one hand, the fact that the subject trace satisfies the ECP, and, on the other hand, the case-licensing of the chain of the subject by the higher verb:

(80)   Un uomo che ritengo [t′ –Fin [t potergli parlare]]

But then, if a TopP occurs, it will make t′ inaccessible to V for case licensing and satisfaction of the ECP, so that the structure will be ill-formed. Again, no such effect is found with the control structure, normally possible with epistemic verbs in Italian:

(81)   Ritengo, a Gianni, di potergli parlare
        "I believe, to Gianni, to be able to speak to him"

Here, no special relation must be established between the main V and (the chain of) PRO, hence a TopP can occur in the C system.

10 Anti-adjacency effects

[. . .]

Consider a typical *that*-t effect, as in (82)a below. In this context, argument topicalization and adverb preposing differ sharply. If embedded argument topicalization applies (with comma intonation and the pragmatics of topic-comment), as in (82)b, the effect is not alleviated (on the other hand, if the preposed element bears focal stress, the acceptability improves; see below on this effect):

(82) a. *A man who I think that t knows this book very well
      b. *A man who I think that, this book, t knows t very well

On the other hand, Bresnan (1977: 194) observed that an adverb interpolating between *that* and the subject trace renders the structure clearly more acceptable (thanks to Kinsuke Hasegawa for bringing Bresnan's observation to my attention in the context of his comment paper to Rizzi (1993), Tokyo, November 1992); these facts have been analyzed independently in Culicover (1992, 1993), Fukui (1993)). Consider the following examples from Bresnan's article:

(83) a. *An amendment which they say that t will be law next year
      b.   An amendment which they say that, next year, t will be law
(84) a. *Which doctor did you tell me that t had had a heart attack during an operation?
      b.   Which doctor did you tell me that, during an operation, t had had a heart attack?

[. . .]

Suppose that the force-finiteness system can be expressed by a single item drawn from the functional lexicon. In English, for embedded finite declaratives we have the alternation *that/0*; I will continue to assume that the latter, but not the former, is consistent with Agr:

(87)   That = +Decl, +fin
       0   = +Decl, +fin, (+Agr)

The analysis of the simple cases of subject extraction then proceeds as in Rizzi (1990). If the form *that* is selected, the trace in subject position remains non properly governed and ECP is violated. If *0* is selected, it is turned into a governor by the Agr specification (which, in turn, is sanctioned by the passage of the subject through its specifier, where it leaves t′), and it properly governs the subject trace t (in turn, t′ is properly governed by the higher verb):

(88) a. *Who do you think [t′ that [t will win the prize]]?
      b.   Who do you think [t′ 0 [t will win the prize]]?

Suppose now that the Topic-Focus field is activated in the C system. Then, the force-finiteness system cannot be realized on a single C head any more because either one or the other specification would not be adjacent to its selecting or selected domain. The force-finiteness system must then split into two heads which sandwich the topic-focus field. So, in examples like the following:

(89)   a.   I think that next year John will win the prize
        b.   Bill said that your book, he really liked

the force specification, which interfaces the C system (and the whole clause) with its selector (the higher V) must be manifested by *that* above the topic; on the other hand, finiteness, which interfaces the C system with the IP, must be manifested by a zero C head (Fin) under the topic. So, we should revise (87) in the following way:

(90)   That = +decl, (+fin)
        0   = (+decl), +fin, (+Agr)

That expresses declarative force and may optionally express finiteness; *0* expresses finiteness, and may optionally express declarative force (as well as agreement).

        [. . .]

    The question which remains to be answered is: [. . .] why couldn't one always violate the *that*-t constraint by separating force and finiteness, hence have a lower agreeing *0* finiteness head licensing the subject trace and cooccurring with a higher *that* (with t′ licensed by head movement of Finº+Agr to *that*)?

(97)   Who do you think [that [t′ Finº+Agr [t will win the prize]]]

This representation must be barred, otherwise we would have free violations of *that*-t. So, the descriptive generalization appears to be that we can have the split between Force and Finiteness (and the consequent salvaging of the subject trace) only if the split is forced by the activation of the topic-focus field. This state of affairs has an obvious "last resort" flavor, and as such is reminiscent of much discussed economy constraints (Chomsky 1991, 1993, 1995, etc.). I will assume the following economy principle to constrain the structure-building process:

(98)   Avoid structure

much in the line of analogous proposals by Safir (1992), Grimshaw (1993), Speas (1994), Giorgi and Pianesi (1994), Crisma (1992) and other recent work (the principle has no exact equivalent in Chomsky's system, but is akin to his Economy of representations).

    The effect of principle (98) in the case at issue is intuitively clear: as the grammar of English has the option of expressing Force and Finiteness in a single head, this option wins over the option of selecting two separate heads (which would

imply two X-bar projections); the latter becomes permissible only if the former is not available because of the activation of the topic-focus field, which forces the split (otherwise, selectional constraints would be violated).

[. . .]

This is rather straightforward intuitively, but the question arises as to how (98) may work formally. Let us assume the basic idea of Chomsky's (1995) approach: economy is computed by comparing derivations within a given reference set, and selecting the simplest. The question then reduces to how the reference set is defined. [. . .] Consider, for instance, the basic distributional constraint on *do* support: *do* can occur only when it is needed (Grimshaw 1993). It is natural to try to express this constraint in terms of an economy principle like (98) (Rizzi 1995), but this is not possible if the reference set is restricted by the numeration: structures with and without *do* would always have distinct numerations. The same problem may be raised, e.g., by the distributional constraints on certain kinds of expletives (in German, Icelandic, etc.), which are limited to positions in which they are needed to satisfy the V-2 constraint. So, our case seems to belong to a larger family of cases having this structure: functional element X can occur only if it is needed to satisfy some structure-building principle. It is natural to try to explain these constraints through principle (98), but this requires a less strict definition of the reference set. A simple modification which achieves the desired result here is that we define the reference set exclusively on the basis of the *lexical* elements of the numeration: functional elements do not define the reference set, rather their occurrence is limited by principle (98) (this is very similar to the approach, expressed within Optimality Theory, by Grimshaw (1995)).

[. . .]

In fact, there is another familiar case in which I to C movement, otherwise obligatory, does not apply in connection with movement of the subject. This happens with main questions on the subject:

(106)   a.   Who did you see t?
        b.   *Who you saw t?
(107)   a.   *Who did see you?
        b.   Who saw you?

I will reproduce here the basic elements of the analysis of Rizzi (1991): I to C movement is compulsory in (106) in order to carry the Wh feature, generated under T, to C, as is required to fulfill the Wh Criterion at S-structure (or before Spell-out); in fact, if I to C does not apply, as in (106)b, the structure is ill-formed; on the other hand, I to C movement cannot apply in the case of a subject question (107)a because the subject trace does not satisfy the ECP in that environment [. . .] nevertheless, the Wh Criterion is satisfied: as the subject has been moved from its base position in the VP to the Spec of C through the Specs of T and AgrS, we obtain the following representation:

(108) $[Who_i\ C^o_i\ [t_i\ AgrS^o_i\ [t_i\ T^o_i\ [t\ V^o\ .\ .\ .]]]]$
                        +wh

C, AgrS and T have specifiers belonging to the same chain, so that, assuming Spec-head coindexation, they share the same index. As they are in the appropriate local relation (no other head intervenes), they can form a representational chain which possesses the Wh feature (still sitting under T); if we define the Wh criterion on chains (a Wh operator must be in a Spec head configuration with a head whose chain possesses the Wh feature), we achieve the desired result: I to C is not required to fulfill the Wh Criterion just in case the questioned element is the subject.

[. . .]

24.3   Questions pertaining to Rizzi (1997)

1   How exactly does Kayne (1994) support Rizzi's claim that there is no adjunction to IP?

2   Pick one instance in which Rizzi uses Relativized Minimality and explore how one might recast the analysis in derivational terms.

3   Discuss to what extent Kayne's (2008; to appear) claim that sentential complements are all relative clause structures impinges on the role of Rizzi's ForceP.

4   Rizzi speaks of an "agreement" rule between C and I that would express the fact that English *that* can introduce finite clauses, but not nonfinite clauses. How might this property of *that* be linked to the contrast between *We're looking for a chair on which to sit* and **We're looking for a chair which to sit on* (in turn contrasting with *We're looking for a chair which we can sit on*)? (Extra credit: What effect would all this have on a general theory of agreement?)

5   Conversely, English *for* can introduce an infinitival complement, as in *For there to be further discussion would be a good idea*, but not a finite complement, as shown by **For there will be further discussion would be a good idea*. How would an agreement approach to these facts compare with an alternative approach based on the idea that what's wrong with the finite example is that *for* has nothing to (Case-)license?

6   Thinking of the similarity in English between *That book I've already read twice* and *That book is one that I've already read twice*, discuss the plausibility and possible consequences of taking Rizzi's proposed Topic[0] to be a copula.

7   Rizzi notes that in Italian a focused phrase in the left periphery necessarily instantiates contrastive focus. Discuss the plausibility and possible consequences of taking sentences containing such preposed contrastively focused phrases to essentially have the structure of cleft sentences.

8   Rizzi, following Chomsky (1977), takes English topics to be merged directly in Spec, Top, rather than being moved into that position from some lower position within IP. Discuss the implications for this proposal of the relative acceptability of sentences like *These pictures of each other, they would rather keep hidden, That large a picture of his children, no man could possibly fit into his wallet*, bringing in chapter 3 of Chomsky (1995).

9   On the basis of the fact that relative pronouns precede topics, Rizzi concludes that relative pronouns must occupy Spec,Force, which is the highest specifier position in the left periphery. How exactly might this follow from the

raising analysis of relatives developed by Vergnaud (1974), Kayne (1994), and Bianchi (1999) and others? (Extra credit: How could one fit in headless relatives?)

10 Rizzi suggests, following Chomsky and Lasnik (1993), that the distribution of controlled PRO is regulated by null Case, which Rizzi takes to be licensed in turn by –fin (via government), that is, to require the presence of (at least part of) the left periphery. Discuss the implications for this proposal of English sentences like *They wrote their paper while under the influence of drugs*, *They left the party happy* and *They acknowledged having cheated on the exam*, or their counterparts in other languages.

11 Rizzi excludes English **John believes to have solved the problem* by attributing to such infinitives the absence of a left periphery. French, though, allows a direct counterpart of this unacceptable English example, namely *Jean croit avoir résolu le problème*. How might Rizzi try to relate this English–French contrast to the ECM facts discussed by Kayne (1981)?

12 Rizzi notes that in Italian a topic cannot intervene between a raising verb like *sembrare* ('to seem') and a following infinitive, as shown by **?Gianni sembra, il tuo libro, conoscerlo bene* ('John seems, the your book, (to-)know-it well'), whereas in control structures such intervention in Italian is fully acceptable. What would Hornstein (1999) have to say to make such contrasts between raising and control compatible with his proposal that (obligatory) control involves movement? (Extra credit: To what extent does the movement approach to control implicit in Kayne 2002 face the same challenge?)

13 (Extra credit) Compare Rizzi's approach to *that*-trace effects in English and French with the more recent one of Koopman and Sportiche (to appear). What are the advantages of each?

14 Rizzi suggests that the impossibility of English sentences like (i) **John did see them only yesterday*, with unstressed *do*, derives from an economy effect of the "avoid structure" sort. Alternatively, thinking of Pollock's (1989, p. 5) idea that the *do* of *do*-support is a kind of pronominal counterpart of the main verb, one could take the unacceptability of sentences like (i) to result from a Condition C-like violation, with *do* illicitly c-commanding its antecedent *see*. What would the constituent structure implications then be for sentences like *Who did John see yesterday?* and *John didn't see us yesterday*? (Extra credit: What kinds of derivations would one then be led to envision?)

15 How would the suggestion of the preceding question affect Rizzi's discussion of the unacceptability of English sentences like **Who did leave first?* (with unstressed did)?

References

Aboh, E. 1995. Notes sur la focalisation en Gungbe. Ms., University of Geneva.

Antinucci, F., and G. Cinque. 1977. Sull'ordine delle parole in italiano: l'emarginazione. *Studi di grammatica italiana* 6: 121–146.

Bayer, J. 1984. COMP in Bavarian. *The Linguistic Review* 3: 209–274.

Beghelli, F., and T. Stowell. 1997. Distributivity and negation: The syntax of *each* and *every*. In A. Szabolcsi (ed.), *Ways of Scope Taking*, Dordrecht: Kluwer, pp. 71–109.

Belletti, A. and U. Shlonsky. 1995. The order of verbal complements: A comparative study. *NLLT* 13: 489–526.

Benincà, P., and C. Poletto. 2004. Topic, focus and V2: Defining the CP sublayers. In L. Rizzi (ed.), *The Structure of CP and IP: The Cartography of Syntactic Structures* 2, New York and Oxford: Oxford University Press, pp. 52–75.

den Besten, H. 1977. On the interaction of root transformations and lexical deletive rules. Available at http://gagl.eldoc.ub.rug.nl/root/1981-20/01/ [accessed January 2013]. Extended version in W. Abraham (ed.),1983, *On the Formal Syntax of the Westgermania. Linguistik Aktuell/Linguistics Today* 3, Amsterdam and Philadelphia: John Benjamins, pp. 47–131.

Bianchi, V. 1999. *Consequences of Antisymmetry: Headed Relative Clauses*. Berlin: Mouton de Gruyter.

Bresnan, J. 1977. Variables in the theory of transformations. In P. Culicover, T. Wasow, and A. Akmajian (eds), *Formal Syntax*, New York: Academic Press, pp. 157–196.

Brody, M. 1990. Some remarks on the focus field in Hungarian. *UCL Working Papers* 2, London: University College London.

Brody, M. 1995. Focus and checking theory. In I. Kenesei (ed.), *Levels and Structures. Approaches to Hungarian 5*, Szeged: JATE, pp. 30–43.

Calabrese, A. 1982. Alcune ipotesi sulla struttura informazionale delle frase in italiano e sul suo rapporto con la struttura fonologica. *Rivista di Grammatica Generativa* 7: 3–78.

Cecchetto, C. 1994. Clitic left dislocation and scrambling: Towards a unified analysis. Ms., DIPSCO, Fondazione San Raffaele, Milan.

Cheng, L. 1991. *On the typology of Wh questions*. Doctoral dissertation, MIT.

Chomsky, N. 1976. Conditions on rules of grammar. *Linguistic Analysis* 2: 303–351. [Reprinted in R. Cole (ed.). 1977. *Current Issues in Linguistic Theory*. Bloomington: Indiana University Press.

Chomsky, N. 1977. On Wh-movement. In P. Culicover, T. Wasow, and A. Akmajian (eds), *Formal Syntax*, New York: Academic Press, pp. 71–132.

Chomsky, N. 1986. Knowledge of Language: Its Nature, Origin, and Use. New York: Praeger.

Chomsky, N. 1991. Some notes on economy of derivation and representation. In R. Freidin (ed.), *Principles and Parameters in Comparative Grammar*, Cambridge, MA: MIT Press, pp. 417–454.

Chomsky, N. 1993. A Minimalist program for linguistic theory. In K. Hale and S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press, pp. 1–52.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N., and H. Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8: 425–504.

Chomsky, N., and H. Lasnik. 1993. The theory of principles and parameters. In J. Jacobs, A. von Stechow, W. Sternefeld, and T. Vennemann (eds), *Syntax: An International Handbook of Contemporary Research*, Berlin: de Gruyter, pp. 506–569. [Reprinted in Chomsky (1995, Ch. 1).]

Cinque, G. 1990. *Types of A′ Dependencies*. Cambridge, MA: MIT Press.

Cinque, G. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24: 239–298.

Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York and Oxford: Oxford University Press.

Cinque, G. 2002. *Functional structure in the DP and IP: The Cartography of Syntactic Structures* 1. Oxford Studies in Comparative Syntax, New York and Oxford: Oxford University Press.

Cinque, G., and L. Rizzi (eds). 2010. *The Cartography of PPs*. New York and Oxford: Oxford University Press.

Cottell, S. 1994. The representation of tense in modern Irish. Ms., University of Geneva.

Crisma, P. 1992. On the acquisition of Wh questions in French. *GenGenP* 0(1–2): 115–122.

Culicover, P. 1992. Topicalisation, inversion and complementizers in English. In D. Delfitto, M. Everaert, A. Evers, and F. Stuurman (eds), *Going Romance, and Beyond*, Utrecht: OTS Working Papers, University of Utrecht, pp. 1–43.

Culicover, P. 1993. The adverb effect: Evidence against ECP accounts of the that-t effects. In A. Schafer (ed.), *Proceedings of NELS 23*, Amherst: GLSA, University of Massachusetts, pp. 97–110.

Di Sciullo, A. M., and C. Boeckx (eds). 2011. *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford Studies in Biolinguistics. Oxford and New York: Oxford University Press.

Fukui, N. 1993. A note on improper movement. *The Linguistic Review* 10: 111–126.

Giorgi, A., and F. Pianesi. 1994. Extraction from subjunctive clauses and clausal architecture. Talk presented at the University of Geneva, November 1994.

Grimshaw, J. 1991. Extended projection. Ms., Rutgers University.

Grimshaw, J. 1993. Minimal projections, heads and optimality. Ms., Rutgers University.

Grimshaw, J. 1995. Projection, heads and optimality. Ms., Rutgers University.

Gundel, J. 1974. *The Role of Topic and Comment in Linguistic Theory*. Bloomington, Indiana: Indiana University Linguistics Club.

Haegeman, L. 1992. *Theory and Description in Generative Syntax*. Cambridge: Cambridge University Press.

Haegeman, L. 1995. *The Syntax of Negation. Cambridge Studies in Linguistics* 75. Cambridge: Cambridge University Press.

Haegeman, L., and R. Zanuttini. 1991. Negative heads and the NEG criterion. *The Linguistic Review* 8: 233–251.

Holmberg, A., and C. Platzack. 1988. The role of inflection in Scandinavian syntax. *Working Papers in Scandinavian Syntax* 42: 25–43.

Hornstein, N. 1999. Movement and control. *Linguistic Inquiry* 30: 69–96.

Horvath, J. 1985. *Focus in the Theory of Grammar and the Syntax of Hungarian*. Dordrecht: Foris.

Iatridou, S. 1991. Clitics and island effects. Ms., MIT.

Kayne, R. S. 1981. On certain differences between French and English. *Linguistic Inquiry* 12: 349–371.

Kayne, R. S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. S. 2002. Pronouns and their antecedents. In S. D. Epstein and T. D. Seely (eds), *Derivation and Explanation in the Minimalist Program*, Oxford: Blackwell, pp. 133–166. [Reprinted in Kayne (2005, Ch. 6).]

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2008. Antisymmetry and the lexicon. In J. Rooryck and J. van Craenenbroeck (eds), *Linguistic Variation Yearbook 8*, Amsterdam and Philadelphia: John Benjamins, pp. 1–31. [Reprinted in Kayne (2010, Ch. 9) and in Di Sciullo and Boeckx (2011, Ch. 14).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Kayne, R. S. to appear. Why isn't *this* a complementizer? In P. Svenonius (ed.), *Functional Structure from Top to Toe: A Festschrift for Tarald Taraldsen*. New York and Oxford: Oxford University Press. [Reprinted in Kayne (2010, Ch. 10).]

Kiss, K. 1987. *Configurationality in Hungarian*. Dordrecht: Reidel.

Koopman, H., and D. Sportiche. to appear. The *que/qui* alternation: New analytical directions. In P. Svenonius (ed.), *Functional Structure from Top to Toe: A Festschrift for Tarald Taraldsen*. New York and Oxford: Oxford University Press.

Laka, I. 1990. *Negation in syntax: On the nature of functional categories and projections*. Doctoral dissertation, MIT.

Larson, R. 1988. On the double object construction. *Linguistic Inquiry* 19: 335–391.

Moro, A. 1995. *The Raising of Predicates*. Cambridge University Press.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Puskas, G. 1992. The Wh criterion in Hungarian. *Rivista di Grammatica Generativa* 17: 141–186.

Rizzi, L. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Rizzi, L. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Rizzi, L. 1991. Residual verb second and the Wh criterion. *Technical Reports in Formal and Computational Linguistics 2*, University of Geneva.

Rizzi, L. 1993. A parametric approach to comparative syntax: Properties of the pronominal system. *English Linguistics* (Tokyo) 10: 1–27.

Rizzi, L. 1995. A note on do support. Ms., University of Geneva.

Rizzi, L. 1996. Residual verb second and the Wh-criterion. In A. Belletti and L. Rizzi (eds), *Parameters and Functional Heads*, Oxford: Oxford University Press, pp. 63–90. [Reprinted in Rizzi (2000, Ch. 9).]

Rizzi, L. 2004. The structure of CP and IP: The Cartography of Syntactic Structures 2. New York and Oxford: Oxford University Press.

Rochemont, M., and P. Culicover. 1990. *English Focus Constructions and the Theory of Grammar*. Cambridge: Cambridge University Press.

Safir, K. 1992. Structural economy. Ms., Rutgers University.

Shlonsky, U. 1994. Agreement in comp. *The Linguistic Review* 11: 351–375.

Speas, P. 1994. Null arguments in a theory of economy of projections. In E. Benedicto and J. Runner (eds), *Functional Projections*, UMOP 17: 179–208.

Tsimpli, I. M. 1994. Focussing in Modern Greek. In K. É. Kiss (ed.), *Discourse-Configurational Languages*, Oxford: Oxford University Press, pp. 176–206.

Turano, G. 1995. *Dipendenze sintattiche in Albanese*. Padua: Unipress.

Vergnaud, J.-R. 1974. *French relative clauses*. Doctoral dissertation, MIT. [Revised version in Vergnaud (1985).]

Vergnaud, J.-R. 1985. *Dépendances et niveaux de représentation en syntaxe*. Amsterdam: John Benjamins.

The Typology of Structural Deficiency: A Case Study of the Three Classes of Pronouns

Anna Cardinaletti and Michal Starke

1999

25.1   Introduction

That pronouns come in different variants has been known for a long time. For example, Wackernagel (1892) noted a correlation between the phonological property of being unable to bear stress and a special syntactic position in a certain class of elements, the so-called *clitic* pronouns. The correlation of prosodic deficiency and special syntactic position is not a logical necessity, nor is it descriptively general. Therefore linguists sometimes make a distinction (introduced in Zwicky 1977) between *special clitics*, which exhibit both properties, and *simple clitics*, which only exhibit prosodic deficiency (cf. Anderson 2005 for discussion). Cardinaletti and Starke's article further contributes to our understanding of the different types of pronouns that are found in natural language.

The article challenges the traditional bipartition of pronominal elements into either strong or clitic pronouns. It argues instead for a more refined, tripartite classification of pronouns into strong, weak, and clitic elements. The three-way classification is motivated by the observation that, in certain syntactic positions, some personal pronouns must refer to a human entity. Cardinaletti and Starke claim that this property characterizes a particular class of pronouns, the *strong* pronouns. Strong pronouns can occur both in their base-position and in peripheral positions (including in isolation), can be coordinated, and can be modified by a certain class of adverbs (e.g., a high modifier like *only*). Pronouns that are not *strong* are *deficient*. Among the deficient pronouns, two classes need to be distinguished: *weak* pronouns and *clitic* pronouns. Weak pronouns are maximal projections and occur in positions where other nominal maximal projections occur, whereas clitic pronouns are heads (in terms of X-bar theory, and occur in head

positions (e.g., adjoined to a verbal head). Furthermore, weak pronouns can bear stress while clitic pronouns cannot.

The article also makes the novel claim that the syntactic, semantic, and prosodic distinctions between the three types of pronouns derive from differences in their internal syntactic structure. The idea that these contrasts are structural (i.e., syntactic rather than lexical) is motivated by the observation that the same classes of correlating properties are found across languages; that is to say, they are not idiosyncratic properties of certain function words. Cardinaletti and Starke propose that the syntactic structure of a clitic pronoun is contained in the structure of its corresponding weak pronoun, which in turn is contained in the structure of the corresponding strong pronoun. In other words, a weak pronoun is like a strong pronoun that lacks its highest projection. This view is supported by considerations based on the morphological shape of pronouns, in particular the observation that if the corresponding pronouns from different classes are morphologically distinct, the deficient pronouns are systematically reduced relative to the strong pronouns.

The idea that function words are not (necessarily) heads but may have internal syntactic structure has had considerable impact on the field. Extending the idea beyond pronouns, Leu (2008, 2010) argues that quantifiers and determiners have layers of syntactic structure and are derived via syntactic operations. The general idea of representing the properties of functional elements in terms of an internal syntactic structure even at the level of the morpheme provides the foundation of what has come to be known as *nanosyntax* (Svenonius et al. 2009; Starke 2011).

25.2 From "THE TYPOLOGY OF STRUCTURAL DEFICIENCY: A CASE STUDY OF THE THREE CLASSES OF PRONOUNS"

1 On the study of pronouns

1.1 The notion of "classes of pronouns"

It is a general property of language that words fall into classes. Among the many relevant oppositions (verbs/adjectives, transitives/ergatives, etc.), one distinguishes itself from all others: that instantiated by the opposition between different classes of pronouns.

This opposition is unique in regularly contrasting *synonymous pairs*; in cutting across all components of grammar; in having no systematic correlation with any interpretive characteristic (semantic or phonetic); in determining a large set of (apparently) absolute universals; and in cutting across lexical classes, §1.1.1–5.

The fundamental goal of the present inquiry is to uncover the primitive underlying these exceptional classes.

1.1.1 What appears to be one pronoun (semantically/functionally defined) falls into distinct classes. The Italian third person plural feminine nominative pronoun,

for instance, divides into two distinct classes with respect to coordination and reference:

$$\langle +\text{hum}\rangle \, \langle -\text{hum}\rangle$$

(1) a. *Esse*   *(*e   quelle accanto) sono troppo alte.*   ✓   ✓
     b. *Loro*   *(e   quelle accanto) sono troppo alte.*   ✓   *
         3PL,FEM,NOM (and those besides) are   too   tall/high

One class of pronouns ("class 1", as in [1b]) may be coordinated, but it is limited to human referents, while the other ("class 2", as in [1a]) cannot be coordinated and may refer to both human and non-human entities. In many cases, the two classes are not only semantically/functionally but also phonetically non-distinct: The French translation of (1), for instance, reproduces exactly the same pattern without morphological variation. As shown in (2), the non-human reading vanishes in coordination:

(2) a. *Elles*   *sont trop grandes.*   ✓   ✓
     b. *Elles et celles d'à côté sont trop grandes.*   ✓   *

The mystery of this correlation between coordination and interpretation diminishes if the formal parallelism between (1) and (2) is taken into account: Despite phonetic identity, (2) features both classes of pronouns. As before, the class which may be coordinated can only refer to human entities:

(3)   occurs in coordination   only human referents
     class 1 *loro, elles$_1$*   +   +
     class 2 *esse, elles$_2$*   −   −

A single semantically/functionally defined pronoun (third person plural feminine nominative unstressed) is the surface reflex of two distinct underlying grammatical elements. The existence of regular synonymous (and often homophonous) pairs, is a rare, if not unique, characteristic of the class 1/class 2 distinction.
     [. . .]

1.1.4   This unique abstract and pervasive distinction also seems to be an absolute universal. It is for example always true that a coordinated personal pronoun cannot refer to a non-human entity. As an example of the cross-linguistic invariance of class 1 and class 2, the following languages all have an asymmetry identical to that in (1)–(2):

$$\langle +\text{human}\rangle \qquad \langle -\text{human}\rangle$$

(4)   German ($\in$ Germanic)
     a. *Sie*   *sind groß.*   ✓   ✓
     b. *Sie und die   daneben sind groß.*   ✓   *
         they and those besides are tall/big

(5) Slovak (∈ Slavic)

		⟨+h⟩	⟨−h⟩
a.	*Vidiel som ich.*	✓	✓
b.	*Vidiel som ich a tých druhých.*	✓	✳
	seen I-am them and these others		

(6) Hungarian (∈ Finno-Ugric)

		⟨+h⟩	⟨−h⟩
a.	*Láttam öket.*	✓	✓
b.	*Láttam öketés a mellettük levöket.*	✓	✳
	I-saw them and those besides		

(7) Hebrew (∈ Semitic)

		⟨+h⟩	⟨−h⟩
a.	*Hi gvoha.*	✓	✓
b.	*Hi ve-zot le-yad-a gvohot.*	✓	✳
	she and-that.one to-side-her tall/big		

(8) Gun (∈ Kwa)

		⟨+h⟩	⟨−h⟩
a.	*Yélè yon wankpè.*	✓	✓
b.	*Yélè kpo yélè kpo yon wankpè.*	✓	✳
	she and she and know(s) beauty		

[. . .]

1.1.6 The conjunction of such exceptional properties (regular synonymy, (homophony,) link between all components of grammar, no link to any interpretive characteristics, absolute universal) makes this distinction one of the most profound and mysterious properties of human grammar.

The goal of this study is to uncover the source of these asymmetries, that which makes a pronoun be a class 1/class 2 pronoun:

> What is γ, the underlying (universal) trigger of (1) which provokes a wide array of distributional, semantic, prosodic and morphological asymmetries between two forms of one and the same pronoun?

2 On being deficient

2.1 Morphology

When (2) is transposed to a masculine subject, two morphologically distinct, though related, pronouns appear, as in the French examples in (10). The same obtains with objects, here illustrated for Italian and Slovak, (11) and (12):

						⟨+h⟩	⟨−h⟩
(10)	a.	*Il*		*est*	*beau.*	✓	✓
	b.	**Il et celui de Jean sont beaux*				✳	✳
	c.	*Lui*		*est*	*beau.*	✓	✳
	d.	*Lui et celui de Jean sont beaux.*				✓	✳
		he and the.one of John is/are pretty					
(11)	a.	*Non metterò mai loro il cappuccio.*				✓	✓
	b.	**Non metterò mai loro e loro il cappuccio*				✳	✳

							⟨+h⟩	⟨−h⟩
	c.	*Non metterò	mai	il	cappuccio	a loro.*	✓	*
	d.	*Non metterò	mai	il	cappuccio	a loro*	✓	*

not I.will.put never the cap/pen-top to them
e   a   quelle altre.
and to those others

(12)	a.	*Vidím ho*				✓	✓
	b.	**Vidím ho*	*a*	*tých*	*druhých*	*	*
	c.	*Vidím jeho.*				✓	*
	d.	*Vidím jeho*	*a*	*tých*	*druhých.*	✓	*

I.see   it/him and these others

Minimally, the fact that the morphological differences exactly correlate with coordination possibilities and with possibilities with respect to human reference, supports the class 1/class 2 distinction. But morphology not only confirms the existence of an abstract γ, it also reveals another property associated with it: The morphological difference is asymmetric. If transparently distinct, class 2 personal pronouns are systematically reduced with respect to class 1 personal pronouns:

(13) *Morphological asymmetry*
Morphology (class 2) ≤ morphology (class 1).

Terminology: The abstractness of the two classes is no impediment to more intuitive terminology. Drawing on the clear orientation of the morphological asymmetry, class 2 elements will be called "deficient", and class 1 elements "strong".

2.2 Distribution

When the initial paradigm, (2) is embedded under *trouver* 'find', strong and deficient personal pronouns surface in different positions:

									⟨+h⟩	⟨−h⟩
(14)	a.	*Jean les*					*trouve*	*belles.*	✓	✓
	b.	**Jean les*		*et*	*celles d'à côté trouve*			*belles*	*	*
	c.	*Jean*					*trouve elles belles.*		✓	*
	d.	*Jean*					*trouve elles*		✓	*

John them(FEM) and those besides finds   them(FEM)
et   celles d'à côté belles.
and those besides pretty

Again, this asymmetry strictly correlates with those discussed above (coordination, human referents, morphology) and such a perfect correspondance [sic] of four properties legitimates the postulation of two abstract classes.

But again, not only is there a *difference* between the two classes, but there is an *asymmetric* difference: One class has an impoverished distribution with respect to the other. While strong pronouns have the distributional liberty of a corresponding noun-phrase (a full noun phrase must occur in post-verbal position in [14]), there

are three types of positions a deficient pronoun cannot occupy (cf. Kayne 1975 for an early systematization of the distributional properties of pairs such as *les/elles* in French).

[. . .]

2.2.3 C–Modification/Coordination

Noun–phrase internal modifiers cannot modify strong personal pronouns, (17a). Adverbs that modify the whole noun phrase (c-modifiers) may however do so, (17b, c). Deficient pronouns cannot be modified by any type of modifiers (17a′, b′, c′):

(17) a.   *{*beau; rapide; . . .*} lui*
     a′. *{*beau; rapide; . . .*} il*
     b.   ✓{*vraiment; seulement; . . .*} lui*
     b′. *{*vraiment; seulement; . . .*} il*
     c.   ✓*lui {seul; aussi; . . .}*
     c′. **il {seul; aussi; . . .}*

The ban on c-modification and coordination holds even if the complex occupies an otherwise licit position:

(18) a.   *Anche/Solo*   {*$*essa_D$; lei_S;   *Maria} è   bella*
                                                                   (c-modification)
     b.   *Lei e(d)*   {*$*essa_D$; lei_S;   *Maria} sono   belle.*
                                                                   (coordination)
          also/only/she and {$(3,SG,F)_D$; $(3,SG,F)_S$; Mary} is/are pretty

[. . .]

(20) *Syntactic Asymmetry*
     A deficient, but not a strong, personal pronoun:
     a.   must occur in a special derived position
     b.   is incompatible with c-modification, coordination.

[. . .]

10   Summary and conclusion

10.1   Prerequisites

10.1.1   The central thrust of the present proposal is that an adequate theory of clitic pronouns, i.e. oppositions between clitic and strong pronouns, such as:

(144) a.   *Gianni la vede   con piacere.*
      b.   *Gianni   vede lei con piacere.*
           Gianni her sees her with pleasure

should be a theory of a considerably enlarged paradigm.

10.1.2   It should be a theory of tripartitions (not bipartitions) of clitic, weak and strong elements, tripartition into which pronominal systems consistently divide, across languages:

(145)   a.   *Non gli   dirò   mai   *gli   tutto   *gli.
        b.   *Non *loro   dirò   mai   loro   tutto   *loro.
        c.   *Non *a lui   dirò   mai   *a lui tutto   a lui.
             not to.him/to.them I.will–say never   everything

Cross-linguistically, each class shares the same properties, which oppose it to both other classes, with a regularity that indicates the presence of three abstract underlying classes, rather than idiosyncratic lexical accidents.

Most notably, both weak and strong elements cross-linguistically occupy XP positions at S-structure, contrary to clitics found only in X^0 positions, while, on the other hand, clitic and weak are both *deficient*, i.e. restricted with respect to a large set of constructions, among which coordination (neither of them is coordinable, while strong elements are).

These two properties further illustrate the intermediate status of weak elements (identical to strong with respect to X-bar, but like clitics with respect to coordination (deficiency)), resulting in a typical *clitic < weak < strong* relationship across the three classes. This is most strikingly illustrated by the fact that all properties differentiating weak elements from strong elements also differentiate clitic elements from strong elements. The deficient characteristics of weak elements are a *proper subset* of the deficient characteristics of clitic elements, again *clitic < weak < strong*.

10.1.3   A theory of clitic pronouns should also handle morphological, as well as distributional, semantic, prosodic, and phonological contrasts. The rich net of asymmetries distinguishing the three classes cuts across all these components: morphology (clitic ≤ weak ≤ strong), distribution (clitic and weak pronouns must be in a derived position, contrary to strong ones; clitics are heads at S-structure, contrary to weak and strong pronouns, etc.), semantics (clitic and weak pronouns lack range, strong pronouns always have one), prosody (clitic and weak pronouns restructure prosodically, contrary to strong ones; weak and strong pronouns may have word-accent, contrary to clitics) and phonology (liaison and contraction rules are restricted to clitic and weak elements).

Surprisingly, while these asymmetries seem to be universal, none of the interpretive asymmetries is systematic: It is not the case that there is a strict covariation between being of one class, and having one type of semantic/phonetic interpretation. The interpretational characteristics are asymmetric but overlapping: The three classes are purely abstract (for instance, both deficient and strong elements can refer to human entities and to prominent discourse referents, although an asymmetry holds with respect to non-human entities and non-prominent referents).

10.1.4   Finally, a theory of clitic elements should be applicable across lexical classes: Just as personal pronouns may be either clitic, weak or strong, all of adverbs, adjectives, quantifiers, Wh-pronouns, nouns, etc. are found in all three formats. Furthermore, the characteristics of clitic, weak and strong elements are largely

identical across categories. A clitic pronoun differs from a strong pronoun in the same way as a clitic adverb differs from a strong adverb.

10.2   Summary
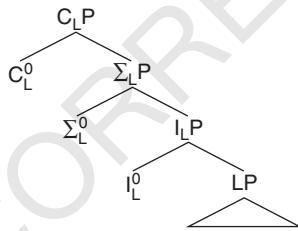
10.2.1 The morphological asymmetry between the three classes (clitic ≤ weak ≤ strong), together with the *Principle and Parameters* framework as it stands, indicates a simple analysis complying with all the above prerequisites. Since deficient elements are systematically morphologically reduced with respect to the strong elements, and since morphemes are syntactic terminals, deficient elements realise less syntactic structure than strong elements. This is particularly clear in transparent morphology, where one class is a morphological subset of the other.

10.2.2   Based on surface morphological forms (which are taken as indicators of the underlying trigger and not as actual triggers), the missing structure is systematically identified as a high functional morpheme: While strong pronouns are full nominal projections, weak pronouns lack the highest functional layer, and clitic pronouns further lack both of the two highest functional layers.
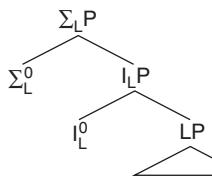
   The uniformity of these layers across classes has then led to the hypothesis that there is one and only one format for all syntactic structure, across languages, constructions and lexical items. Deviation from this basic format, an extremely rare fact, leads to *deficiency*, triggering strong consequences for the deficient element.

   Based on the nature of the interpretive properties involved, and on the most widely accepted labelling, that attributed to verbal (extended) projections, the labels adopted are (where IP is a cover term for a set of functional projections, and subscript L refers to any lexical category):
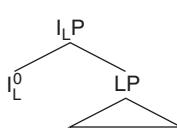
(146)   a.   Strong pronouns

         b.   Weak pronouns   c.   Clitic pronouns

From this, most aspects of deficiency directly follow: Morphological reduction is a direct reflex of lack of structure, impossibility of modification follows from the observation that those modifiers that can modify strong but not deficient elements only modify full CPs, the choice preference follows from the diverse structures in

combination with a general principle Minimise α, the prosodic asymmetries comply with the observation that "major constituents" (i.e. CPs) are treated differently from non-major constituents with respect to prosodic processes. Other distributional and semantic asymmetries follow not from the sheer absence of structure, but from the absence of features contained in those structures: case features in C^0 (and consequently, referential information) and polarity and prosodic features in Σ^0.

10.2.3   An attentive observation of the choice preference shows that strong elements have logical *priority* over deficient elements: A deficient element must be chosen but only if it is associated with the same features as *those which would have been contained in the strong counterpart*.

This primacy, together with the whole general theory of deficiency, may be implemented through three general assumptions, two of which are hardly more than expressions of what is generally assumed:

(147)   a.   Minimise α (Economy Principle)
        b.   Information of level R must be present at $R + n$

                                                          (Projection Principle)
        c.   All entries of the syntactic lexicon realise a fixed array of (underspecified) features, $\alpha_1 \ldots \alpha_n$.

The identity of all (extended) projections now follows from (147c): Features $\alpha_1 \ldots \alpha_n$ invariably project onto what becomes CP, IP, etc. It also follows that only strong pronouns are generable. The (generalised) economy principle then forces to reduce structure as much as possible, α being in this case structure (strictly speaking, *Minimise* α is forced to operate as Erase α in this case), thus deriving both the choice asymmetry and the primacy of strong elements.

Finally, the Projection Principle forces recovery of features erased by *Minimise* α. This recovery is possible only through a local relation between the deficient element and an adequate head at S-structure (assuming there to be no displacement at PF), thus deriving the distributional asymmetries.

10.2.4   The relevant set of properties now all follow, independently from the nature of the lexical head (across verbs, nouns, adjectives, etc.), with the desired morphological, semantic, prosodic and syntactic consequences:

(148)   a.   From the sheer absence of the highest projections in deficient
             elements (clitic and weak), it follows that:
             –   the more an element is deficient, the more it tends to be morphologically reduced;
             –   deficient elements cannot be modified by modifiers of CPs;
             –   deficient elements are not "major constituents", a central notion in prosody;
             –   the most deficient element possible is preferred. (by Minimise α)
        b.   From the absence of C-features in deficient elements, it follows that:
             –   deficient elements never have their own range (and are thus always either expletive or co-referent);

– deficient elements must be displaced to recover missing (case-) features.   (Projection Principle)

c.   From the absence of Σ–features in clitic elements, it follows that:
– clitic elements do not have word-accent;
– clitic-elements must be displaced to recover missing (prosodic) features. To not destroy the effect of the recovery of the C-feature, a X^0–chain must be created.   (Projection Principle)

10.3   Conclusion

The present investigation, we hope, illustrates the interplay of abstract theoretical constructs and empirical generalisations. The first part seeks to establish what is to be explained by a theory of simple oppositions between clitic and strong pronouns, arriving at several new conclusions. Most prominently, that the relevant opposition is among three distinct classes: clitic, weak and strong; but also that these classes are separated by a regular range of semantic (referential) oppositions. The global picture then becomes uniform: Clitic pronouns are deficient with respect to weak pronouns which are in turn deficient with respect to strong pronouns, both distributionally, morphologically, semantically and prosodically.

This generalisation, (75) [in the full article], then, indicates a simple abstract primitive: *Structural Deficiency*. Some pronouns are deficient in that they have a deficient syntactic structure. For the first time, to our knowledge, this opens a (tentative) road towards a unified derivation of the whole range of syntactic, morphological, semantic and prosodic effects involved, but also of the similar properties of pronominal, adverbial, adjectival, etc. clitic, weak and strong elements.

The postulation of diverse structures then entails a set of constraints which regulate the generation and derivation of syntactic structure, further constraining the general model of grammar upon which it is based. Choice patterns among distinct classes of pronouns indicate that only full, i.e. strong, structures are generable. The existence of deficient structure must then be attributed to a reduction process in syntax, traced down to a general *Minimise Structure* principle, subsumed under a global economy principle *Minimise* α (cf. Chomsky's 1993 economy guidelines). Finally, this entails a split lexicon, with post-syntactic access to morpho-phonological information (cf. Halle and Marantz 1993, Jackendoff 1994).

The "theory of clitics" thus developed is a general theory of arguments and adjuncts, and of their syntactic structure, thereby defining a novel set of central questions, which we hope to be productive avenues of research.

25.3   Questions pertaining to Cardinaletti and Starke (1999)

1   Pick two languages and discuss how possessive pronouns in those languages fit in to Cardinaletti and Starke's tripartite characterization of pronouns.
2   Prenominal possessive pronouns in Spanish fail to agree in gender when the pronoun is singular or third person plural (*mi(*a) casa, tu(*a) casa, su(*a)*

casa). Yet there is agreement in gender when the pronoun is first or second plural (*nuestra casa, vuestra casa*). How might this be related to Spanish having, as strong pronouns, *nosotros/nosotras* and *vosotros/vosotras*, in which the pronominal part *nos/vos* is followed by what looks like the Spanish counterpart of *other*, namely *otros/otras*?

3 Cardinaletti and Starke correctly note the impossibility of English *it* being used in a way parallel to *that*, e.g., *Give us that and that* vs. **Give us it and it*. On the other hand, they don't discuss the fact that *it* can be used in coordination in cases involving a contrast between person and thing or animal, e.g., (speaking of John and his dog) *Neither he nor it will be able to cross that street* or (speaking of John and his bicycle) *Both he and it have seen better days*. How might they try to integrate these facts? (Hint: In note 2, they begin to bring in demonstratives.)

4 Similarly, English allows fairly well (speaking of John and his latest paper) *We don't think very highly of him, we think highly only of it*. How directly does the answer to the previous question extend to this fact?

5 Cardinaletti and Starke mention the question of the number of syllables that a pronoun may or may not have. They argue that number of syllables is unlikely to lie at the heart of the phenomena they discuss. Their analysis is couched, rather, in terms of structural differences among subtypes of pronouns, which seems correct. Yet clitics, unlike the Italian weak pronoun *loro*, never seem to be bisyllabic. Test this hypothesis against three languages from families other than Romance, Germanic, or Slavic. If you find no counterexamples (i.e., no bisyllabic or multisyllabic clitics), suggest why such clitics might be prohibited.

6 What are the most salient similarities and the most salient differences between Cardinaletti and Starke's approach to pronouns and that pursued by Déchaine and Wiltschko (2002)?

7 Cardinaletti and Starke take clitics to be heads, as opposed to phrases. How might they defend this position against the objection that many clitics are bimorphemic or even multimorphemic (e.g., Spanish *los/las*)? (Extra credit: To what extent is their position compatible with Kayne's (2008) analysis of Paduan *ghe* and other "expletive" clitics?)

8 Why exactly can a clitic double a weak pronoun (or a lexical DP), without a weak pronoun being able to double a lexical DP?

9 French allows both *Marie court après lui* ('M runs after him'), with what looks like a strong pronoun, and *Marie lui court après*, with what looks like a clitic. (For additional relevant data, see Kayne (1975, Ch. 1 and 2).) How might this double possibility be reconciled with Cardinaletti and Starke's "Minimize Structure" principle?

10 What is the significance for Cardinaletti and Starke of French sentences like *Jean a presque tout oublié* ('J has almost all/everything forgotten') or English sentences like *You may perfectly well be right*, with a high, yet modifiable adverb?

11 Discuss how one might fit the data concerning English pronouns from Klima (1964) into Cardinaletti and Starke's analysis.

12 In French, negative *rien* ('nothing') moves to a relatively high position, as in *Jean n'a rien dit* ('J neg has nothing said'), whereas negative phrases containing a lexical noun do not, as exemplified by *Jean n'a lu aucun livre* ('J neg has read

no book') vs. *Jean n'a aucun livre lu. In Icelandic, even negative phrases that contain a lexical N move high. Which language fits better into Cardinaletti and Starke's approach? Discuss how the other might yet be integrated.

13 How might Cardinaletti and Starke's notion of deficient pronoun be generalized so as to encompass English OV compounds, as in *an avid magazine-reader*? What exactly are the similarities and differences between *magazine* in such compounds (cf. also Baker 1988 on noun-incorporation) and weak pronouns of the sort discussed by Cardinaletti and Starke?

14 Proper names in English sometimes share behavior with pronouns more than with full DPs, e.g., *Isn't that the car of ?your friend/*John/**him?* Find at least two comparable examples in other languages. How might Cardinaletti and Starke integrate such facts into their approach?

15 Discuss the tension between Cardinaletti and Starke's approach to prepositions such as French *à* ('to') or Italian *a* ('to') and Kayne's (2004) proposal.

16 Cardinaletti and Starke take both the *a/à*-type of preposition and the *di/de*('of')-type to be complementizer-like. To what extent is that compatible or not with Caha (2010)?

17 The general impossibility of modifying clitics or weak pronouns is taken by Cardinaletti and Starke to be closely related to the general impossibility of coordinating clitics or weak pronouns, which is in turn taken to be related to the difficulty of coordinating complements of prepositions like *a/à*. Yet for some speakers of English there is a contrast between ??*Mary showed her paper to only her sister* and *Mary showed her paper to both her brother and her sister*. How might Cardinaletti and Starke try to reconcile these English facts with their proposal?

18 How well would idiomatic object clitics in Romance languages be expected to fit in to Cardinaletti and Starke's analysis?

19 Cardinaletti and Starke argue that morphology is relevant to syntax insofar as their notion of "deficiency" is concerned. How similar or dissimilar is that to Pollock's (1989, Ch. 15 of this volume) discussion of morphology-syntax interactions?

20 Kayne (1975, Ch. 1 and 2) discusses instances in French of floating dative *tous* that are not accompanied by the preposition *à*. To what extent are those French facts similar to those concerning Italian *loro* as discussed by Cardinaletti and Starke?

References

Anderson, S. R. 2005. *Aspects of the Theory of Clitics*. Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press.

Baker, M. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Caha, P. 2010. The German locative-directional alternation: A peeling account. *Journal of Comparative Germanic Linguistics* 13: 179–223.

Chomsky, N. 1993. A Minimalist program for linguistic theory. In K. Hale and S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press, pp. 1–52.

Déchaine, R.-M., and M. Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33: 409–442.

Halle, M., and A. Marantz. 1993. Distributed morphology and the pieces of inflection. In K. Hale and S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press, pp. 111–176.

Jackendoff, R. 1994. Lexical insertion in a post-minimalist theory of grammar. Ms., Brandeis University.

Kayne, R. S. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

Kayne, R. S. 2004. Prepositions as probes. In A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures* 3, New York and Oxford: Oxford University Press, pp. 192–212.

Kayne, R. S. 2008. Expletives, datives, and the tension between morphology and syntax. In T. Biberauer (ed.), *The Limits of Syntactic Variation*, Amsterdam and Philadelphia: John Benjamins, pp. 175–217. [Reprinted in Kayne (2010, Ch. 6).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Klima, E. S. 1964. Negation in English. In J. A. Fodor and J. J. Katz (eds), *The Structure of Language*, Englewood Cliffs, NJ: Prentice-Hall, pp. 246–323.

Leu, T. 2008. *The internal syntax of determiners*. Doctoral dissertation, New York University.

Leu, T. 2010. The internal syntax of *jeder* 'every'. In J. van Craenenbroeck and J. Rooryck (eds), *Linguistic Variation Yearbook 9*, Amsterdam and Philadelphia: John Benjamins, pp. 153–204.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Starke, M. 2011. Towards elegant parameters: Language variation reduces to the size of lexically stored trees. Ms., Tromsø. [Available at http://ling.auf.net/lingBuzz/001183; accessed Jan 2013.]

Svenonius, P., G. Ramchand, M. Starke, and K. T. Taraldsen. 2009. Special issue on nanosyntax. *Nordlyd* 36.

Wackernagel, J. 1892. Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1: 333–436.

Zwicky, A. M. 1977. *On Clitics*. Bloomington, Indiana: Indiana University Linguistics Club.

26

Bare and Not-So-Bare Nouns
and the Structure of NP

Lisa Lai-Shen Cheng and Rint Sybesma

1999

26.1   Introduction

One of the challenges of research in linguistics, as in any other scientific field, is that of finding explanations that can be generalized from the case under investigation to the more general case. In linguistics, we want our explanations not only to further our understanding of the language (or languages) that we are studying, but also to extend to other, potentially very different languages. Take as an example two of the proposals made in Longobardi (1994, chapter 21 of this volume): that argument noun phrases must be DPs, and that N can move to D. These proposals were based on a detailed study of Italian and a comparison with English. How do they fare when applied to languages that lack overt determiners entirely? Does the distinction between the categories NP and DP extend to such languages? Is there any evidence of N-movement?

In this article, Cheng and Sybesma discuss the formal, distributional, and interpretive properties of noun phrases in two Chinese languages, Mandarin and Cantonese, whose noun phrases consist either of a bare noun, or a noun introduced by a classifier (optionally accompanied by an overt numeral or demonstrative). Classifiers, they argue, are one way in which the syntax indicates that something can be counted (another way, employed by English, is via the category number): they either create or name a unit of measure in which the entity denoted by the noun can be partitioned and counted. In Chinese, they are not restricted to mass nouns, but also occur with count nouns. They have some of the same properties as the category D in other languages, like that of making an NP able to refer and function as an argument.

Bare nouns can be interpreted as indefinite, generic, or definite in Mandarin, and as indefinite or generic in Cantonese. Interestingly, in both languages the distribution of bare nouns with indefinite interpretation is restricted: they can only occur as the complement of a lexical head (basically, in postverbal position):
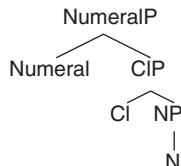
(1)   a.   Wufei   heoi   maai   *syu*. (Cantonese)
           Wufei   go   buy   book
           'Wufei went to buy a book/books.'
      b.   **Gau*   soeng   gwo   maalou.
           dog   want   cross road
           Intended meaning: 'A dog wants to cross the road.'

The distribution of bare nouns with generic interpretation, in contrast, is not restricted; as shown in (2), they can occur in both pre- and postverbal position:

(2)   a.   Ngo   zungji   *gau*. (Cantonese)
           I   like   dog
           'I like dogs.'
      b.   *Gau*   zungji   sik   juk.
           dog   like   eat   meat
           'Dogs love to eat meat.'

Cheng and Sybesma account for this pattern by arguing that argument bare nouns are never simply NPs: they are minimally ClassifierPs (ClPs), that is, NPs that are the complement of a classifier. Bare nouns with a generic interpretation (possible in both Mandarin and Cantonese) and with a definite interpretation (possible in Mandarin) are ClPs, in which the N has raised to Cl; because Cl has lexical content, their distribution is not restricted, as seen in (2). In contrast, bare nouns with an indefinite interpretation, like those in (1), are NumeralPs in both Mandarin and Cantonese, that is, phrases headed by a null Numeral head that takes the ClP as its complement:

(3)   Indefinite noun phrase:

In these noun phrases, N does not raise, and Num and Cl are phonetically null. Hence, like the Italian noun phrases with a null D described by Longobardi, these bare nouns can only occur as the complement of a lexical head, a restriction that is characteristic of certain null categories.

This article offers a particularly clear example of how ideas developed from the observation of a certain empirical domain can be successfully applied, with the right modifications, to different empirical domains. In particular, the hypotheses

that argument noun phrases have an articulated syntactic structure involving a functional element, D, and in some cases exhibit movement of N to D are transposed, with appropriate modifications, from languages that have determiners to those that appear to lack them entirely and employ (overt and null) classifiers and numerals to express countability and plurality.

26.2 From "BARE AND NOT-SO-BARE NOUNS AND THE STRUCTURE OF NP"

1 Introduction

It is well known that Chinese languages allow bare nouns to appear as arguments. In this article we discuss in detail the distribution and interpretation of bare nouns in Mandarin and Cantonese. We also examine [classifier + noun] ([Cl + N]) phrases in these two Chinese languages because they display interesting similarities to, and differences from, bare nouns. More specifically, we focus on the following issues:

Various interpretations. Bare nouns in Mandarin and Cantonese can have more than one interpretation. In postverbal position, for instance, Mandarin bare nouns can be interpreted as indefinite (1a), definite (1b), or generic (1c). In preverbal position, they can be interpreted as definite (2b) or as generic (2c), but not as indefinite (2a). (SFP = sentence-final particle)

(1) a. Hufei mai shu   qu le.
       Hufei buy book go SFP
       'Hufei went to buy a book/books.'
    b. Hufei he–wan–le   tang.
       Hufei drink–finish–LE soup
       'Hufei finished the soup.'
    c. Wo xihuan gou.
       I   like   dog
       'I like dogs.'
(2) a. Gou yao guo malu.
       dog want cross road
       'The dog wants to cross the road.' *not*: 'A dog wants to cross the road.'
    b. Gou jintian tebie tinghua.
       dog today   very obedient
       'The dog/dogs was/were very obedient today.'
    c. Gou ai   chi rou.
       dog love eat meat
       'Dogs love to eat meat.'

The immediate question that arises is, How do we account for this variety of interpretations and the restriction associated with indefinite bare nominals? (We should note that in this article, except for a brief comment in section 5, we will not discuss demonstratives.)

Differences between Cantonese and Mandarin. When it comes to the distribution and interpretation of bare nouns, the Cantonese pattern is almost the same as the Mandarin pattern illustrated in (1), the crucial exception being that Cantonese bare nouns cannot be interpreted as definite. Instead, Cantonese uses the [Cl + N] combination (see also Leung 1980). In other words, in Cantonese, just as in Mandarin, bare nouns can be indefinite postverbally (3a) but not preverbally (4a), and generic in both positions (3c)/(4c). Definite counterparts of (1b) and (2b), however, are different in Cantonese: instead of bare nouns, Cantonese uses [Cl + N], as illustrated in (3b) and (4b).

(3) a. Wufei heoi maai syu.
         Wufei go   buy book
         'Wufei went to buy a book/books.'
     b. Wufei jam–jyun   *(wun) tong la.
         Wufei drink–finish CL   soup SFP
         'Wufei finished drinking the soup.'
     c. Ngo zungji gau.
         I   like   dog
         'I like dogs.'

(4) a. *Gau soeng gwo maalou.
          dog want cross road
          'A dog wants to cross the road.'
     b. Zek gau gamjat dakbit tengwaa.
         CL dog today special obedient
         'The dog is specially obedient today.'
     c. Gau zungji sek juk.
         dog like   eat meat
         'Dogs love to eat meat.'

It should be noted that [Cl + N] phrases such as the ones in (3b) and (4b) can only be singular, unlike definite bare nouns, which can be plural (see (2) and the discussion in section 3.1). Moreover, in addition to a definite reading, Cantonese [Cl + N] can receive an indefinite, nonspecific reading, as in (5); however, a generic interpretation for [Cl + N] is impossible, as in (6a–b) (see also section 3.3).

(5) Ngo soeng maai bun syu (lei   taai).
     I   want buy CL book come read
     'I want to buy a book (to read).'

(6) a. Zek gau zungji sek juk.
         CL dog like   eat meat
         'The dog likes to eat meat.' *not*: 'Dogs like to eat meat.'
     b. Ngo zungji tong zek gau waan.
         I   like   with CL dog play
         'I like to play with the dog.' *not*: 'I like to play with dogs.'

This brings us to a second difference between Cantonese and Mandarin. As shown in (7), contrary to what is generally assumed (see, e.g., Li and Thompson

1981:104, Tang 1990), it is not the case that in Mandarin a classifier must cooccur with an overt numeral or a demonstrative. The example in (7) shows that Mandarin also has [Cl+N] phrases: in section 3.2 we provide evidence that these are not cases of phonological reduction.

(7) Wo xiang   mai ben shu.
     I   would-like buy CL   book
     'I would like to buy a book.'

However, whereas in Cantonese [Cl+N] phrases can receive either a definite or an indefinite reading, in Mandarin they are restricted to an indefinite interpretation: they cannot be interpreted as definite (or generic).

    In sum, bare nouns and [Cl+N] phrases have the following interpretational possibilities in Cantonese (C) and Mandarin (M):

(8)   Indef   Def   Gen

     Bare nouns   M/C   M   M/C
     [Cl + N]   M/C   C   ——

The difference between the two languages centers on the definite interpretation of bare nouns and [Cl+N] phrases: in Mandarin only bare nouns can be definite, and in Cantonese only [Cl+N] phrases can. In turn, the difference between bare nouns and [Cl+N] phrases centers on the definite interpretation and the generic interpretation. Although both bare nouns and [Cl+N] phrases allow both a definite and an indefinite interpretation, they can't do both in both languages, as just noted. Furthermore, [Cl+N] cannot have a generic interpretation in either language.

    The question of how these differences come about is certainly connected to the earlier question of how the various interpretations arise.

    Chinese NP denotation. The third issue concerns the licensing of bare nouns. The fact that Chinese bare nouns can appear as arguments is interesting in view of the claim made by Stowell (1989), Szabolcsi (1994), Longobardi (1994), and others that only DPs can function as arguments. If this is true, then bare nouns in Chinese must involve more structure than just the bare N (or the bare NP). In contrast, Chierchia (1995, 1998) proposes that Chinese nouns differ from nouns in languages like English in that they are arguments (of type ‹e›) rather than predicates (of type ‹e,t›) and thus can appear bare.

    This article is organized as follows. In section 2 we investigate some of the general issues mentioned in the previous paragraph. In section 3 we address some of the interpretational issues, particularly the indefinite and definite interpretation of bare nouns and [Cl+N] phrases; we look further into the differences between Cantonese and Mandarin; and we discuss generics and specificity. In sections 4 and 5 we briefly touch on more general issues, such as the function of DP and some consequences of our proposals for the analysis of Chinese pronouns and demonstratives, and for languages with articles such as Italian and English.

    [. . .]

3 The Distribution and Interpretation of Bare Nouns and [Cl+N] Phrases

In this section we propose an account of the distribution and interpretation of bare nouns and [Cl+N] phrases in Mandarin and Cantonese. First, we consider the definite interpretation of noun phrases.

3.1 Definite noun phrases

3.1.1 Definite [Cl+N] phrases

In Cantonese, as (4) and (6) illustrate, definite noun phrases are [Cl+N] phrases. Furthermore, definite [Cl+N] phrases are not restricted to postverbal positions. These phrases are thus quite similar to English definite DPs such as *the man*. This is not surprising since, as noted in section 2.2.1 [in the full article], classifiers have a singularizing function, just like D. In other words, if we maintain that Chinese nouns are predicates, classifiers are like Ds in that (a) they are type-shifters, changing predicates into arguments, and (b) they yield the definite interpretation (comparable to an iota operator "ι") (see also Chierchia's (1998) treatment of English).

Definite [Cl+N] phrases in Cantonese nonetheless differ from definite DPs in English in one crucial respect. In English a definite DP is a singular DP if the noun is singular and plural if the noun is plural (e.g., *the student* vs. *the students*). Nouns in Chinese languages do not have number and thus have no singular/plural distinction. However, since classifiers are singularizers, [Cl+N] phrases are generally singular: the classifier picks out a single instance from the count mass domain. To indicate plurality, a special classifier, *di*, is used, as shown in (18), in contrast with (19) (taken from Matthews and Yip 1994:89).

(18) a. Di ce zo-zyu   go ceot-hau.
          CL car block–CONT CL exit
          'The cars are blocking the exit.'
      b. Di leotsi jiu   hou lek   sin dak.
          CL lawyer need very smart only-okay
          'The lawyers had better be very smart.'
(19) a. Gaa ce zo-zyu   go ceot-hau.
          CL   car block–CONT CL exit
          'The car is blocking the exit.'
      b. Go leotsi jiu   hou lek   sin dak.
          CL   lawyer need very smart only-okay
          'The lawyer had better be very smart.'

Di thus picks out multiple instances from the domain. The fact that plurality rests upon classifiers provides further support for Doetjes's proposal regarding the function of classifiers (the connection with number morphology).

3.1.2 Definite bare nouns

Now consider definite noun phrases in Mandarin. As (2) illustrates, definite noun phrases in Mandarin are bare nouns instead of [Cl + N] phrases. This naturally leads to questions such as (a) are definite bare nouns in Mandarin NPs or ClPs? (b) how does the definite interpretation arise? and (c) why are definite [Cl + N] phrases not possible in Mandarin?

As (1) and (2) illustrate, definite bare nouns in Mandarin differ from their indefinite counterparts in that they can appear in both preverbal and postverbal positions. In line with what we have suggested for indefinite bare nouns, we assume that definite bare nouns are not just NPs, but ClPs. The fact that definite bare nouns can appear in preverbal position suggests that the head of ClP is not empty. In the spirit of Longobardi 1994 (and Ritter 1989, among other works), we propose that in cases where a bare noun receives a definite interpretation, N has moved to Cl (i.e., N-to-Cl movement has taken place): once the Cl position has been filled, the phrase is no longer limited to occurring in lexically governed positions.

It should be noted that N-to-Cl movement is covert, as evidenced by noun phrases containing adjectival/possessive modifiers. Overt N-to-D movement of proper names in Italian results in ordering differences, as shown in (20); by contrast, Mandarin bare nouns follow adjectival modifiers, as shown in (21).

(20) a. È venuto il   vecchio Cameresi.
          came   the older Cameresi
          'The older Cameresi came.'
      b. È venuto Cameresi vecchio.
          came   Cameresi older
          (same)
(21) Huangrong de gou jintian tebie tinghua.
      Huangrong DE dog today   very obedient
      'Huangrong's dog was very obedient today.'

Consider now how the definite interpretation arises. We follow Chierchia (1998) in assuming that for languages that do not have a definite article, the nonovert ι operator is available. Chierchia (1998) assumes (following Partee (1987)) that the ι operator is a type-shifter and that it is equivalent to a definite article. Chierchia proposes that if a language (such as English; see also section 3.2.2) has a definite article, ι is not available (i.e., it is blocked). He considers this to represent the last-resort nature of the type-shifting operation: "[I]f there is a determiner D whose meaning is a particular type shifting, then use of that operation as an automatic type-changing functor is blocked" (p. 360). Neither Cantonese nor Mandarin has a definite article. However, both languages have the equivalent of a definite article, namely, classifiers. In Cantonese, classifiers are used for definite noun phrases. That is, Cantonese does not resort to ι for definite interpretation of nouns. However, in Mandarin, definite [Cl + N] phrases are not possible. Leaving aside

for the moment why [Cl + N] is not possible in Mandarin (we will discuss this below), the impossibility of using a classifier in this context allows ι to be used.

We suggest that N-to-Cl movement is a necessary step for use of the ι operator (either because the ι operator changes the NP ‹e,t› into an individual ‹e›, which cannot stay in an NP and thus must undergo movement, or because the N must be in Cl position for the ι operator to function).

N-to-Cl movement in Mandarin, in contrast with classifier insertion in Cantonese, can also explain a clear-cut difference in the definite noun phrases of the two languages: namely, Cantonese definite [Cl + N] phrases are necessarily singular (except when the classifier *di* is used). In Mandarin, however, no overt classifier is present in definite noun phrases; definite bare nouns can thus be interpreted as singular or plural, as shown in (2b).

[. . .]

3.2 Indefinite noun phrases

3.2.1 Indefinite [Cl + N] phrases

Consider first the distribution of indefinite [Cl + N] phrases in Cantonese ((25) is taken from Matthews and Pacioni 1996:(23); (19) is repeated here as (26)).

(25) a. Keoi seung maai gaa ce.
          he   want buy CL car
          'He wants to buy a car.'
      b. Keoi maai-zo gaa ce.
          he   sell-zo CL car
          'He sold the car.'

(26) a. Gaa ce zo-zyu   go ceot-hau.
          CL   car block–CONT CL exit
          'The car is blocking the exit.' *not*: 'A car is blocking the exit.'
      b. Go leotsi jiu   hou lek   sin dak.
          CL lawyer need very smart only–okay
          'The lawyer had better be very smart.' *not*: 'A lawyer had better be very smart.'

When [Cl + N] phrases appear postverbally, as in (25), both definite and indefinite interpretations are possible (the verb *maai* 'to buy' differs from the verb *maai* 'to sell' in tone). Note that in (25b) the predicate *maai-zo* 'sold' could in principle facilitate either a specific indefinite reading or a definite reading [. . .]. The fact that *gaa ce* 'CL car' can be interpreted as definite but not indefinite in (25b) suggests that when [Cl + N] phrases are interpreted as indefinite, they must be nonspecific as well (we will discuss this further shortly below). In contrast, when [Cl + N] phrases appear preverbally, as in (26), only the definite reading is possible. This is reminiscent of the distribution of indefinite bare nouns (i.e., only postverbal position is possible).

Mandarin [Cl + N] phrases display the same distribution, as shown by the contrast between (27) and (28). Note that since definite noun phrases in Mandarin

cannot be [Cl + N] phrases, (28) is ungrammatical (in contrast with Cantonese [Cl + N] phrases, which can be interpreted as definite, as in (26)).

(27) a. Wo xiang   kan ben shu.
         I   would like read CL   book
         'I would like to read a book.'
     b. Men-qian you ge ren.
         door-front have CL people
         'There is someone outside the door.'
(28) *Ben shu bu   hao.
         CL   book not good
         'The/A book is not good.'

As noted earlier, it is not generally believed that [Cl + N] combinations in Mandarin can appear without a numeral or a demonstrative (but see Paris 1981). Sentences like (27a–b) are viewed as cases of phonological reduction of the numeral *yi* 'one'. On this view, *ben shu* 'CL book' in (27a) stands for *yi-ben shu* 'one-CL book' in which the numeral gets suppressed in (fast) speech only.

We have reasons to think that this view is wrong: Mandarin (or Cantonese) [Cl + N] phrases are not simply phonological reductions of [*yi*-Cl + N] 'one Cl + N'. The main reason is that [Cl + N] phrases and [*yi*-Cl + N] phrases have a different distribution. In particular, as (25) shows, indefinite [Cl + N] phrases in Cantonese can be interpreted as indefinite nonspecific only. On the other hand, [*yi*-Cl + N] phrases can be interpreted as specific and nonspecific indefinites. Thus, in contexts where only an indefinite specific interpretation is possible, [Cl + N] phrases should not be able to surface. We present two contexts showing that this prediction is borne out.

The first context involves bounded predicates. Sybesma (1992:176–178) argues that with predicates that are bounded for reasons independent of the object, a strong reading is forced upon the object: a bare noun is interpreted as definite, an indefinite NP as specific. Comparing (29) and (30), we find that a [Cl + N] phrase cannot occur as the object of one of these bounded predicates (however fast the sentence is pronounced), whereas [*yi*-Cl + N] phrases can, with a specific reading. There is no phonological reason why *yi* 'one' could not be suppressed in (30).

(29) a. Wo chi-wan-le   yi-kuai binggan.
         I   eat-finish-LE one-CL cookie
         'I finished a cookie.'
     b. Wo he-wan-le   yi-wan   tang.
         I   drink-finish-LE one-bowl soup
         'I finished a bowl of soup.'
(30) a. *Wo chi-wan-le   kuai binggan.
         I   eat-finish-LE CL   cookie
     b. *Wo he-wan-le   wan tang.
         I   drink-finish-LE bowl soup

The examples in (30) show that [Cl + N] phrases in Mandarin must be nonspecific indefinites. The *ba*-construction (31) provides a similar context of boundedness, and the same pattern emerges: [*yi*-Cl + N] is acceptable, [Cl + N] is not. Once again, there is no reason why *yi* 'one' could not be suppressed here.
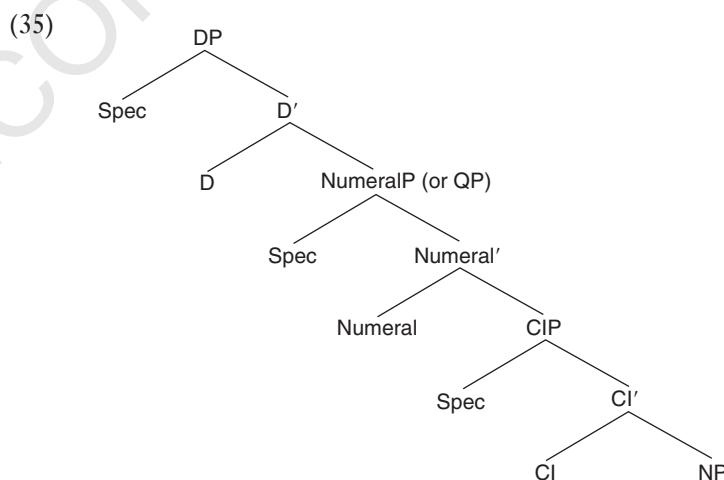
(31) a. Wo ba yi-wan   tang he-wan-le.
          I   BA one-bowl soup drink-finish-LE
          'I finished a (particular) bowl of soup.'
      b. *Wo ba wan   tang he-wan-le.
          I   BA bowl soup drink-finish-LE

[. . .]

3.2.2 The indefinite [CL + N] and why Cantonese and Mandarin differ

From the above data on indefinite noun phrases in Mandarin and Cantonese, it is clear that indefinite bare nouns and indefinite [Cl + N] phrases are similar in interpretation and distribution: both are interpreted as nonspecific indefinites, and both are restricted in distribution (postverbal/governed positions). What accounts for this similarity? In particular, is there a structure that can account for the two indefinite noun phrase types while taking the restriction in interpretation into consideration as well?

Given our earlier hypothesis regarding indefinite bare nouns, the null hypothesis for the indefinite [Cl + N] phrases is that they too contain an empty element that must be lexically governed. The obvious question is what that empty category might be. An obvious choice is Numeral. That is, an indefinite [Cl + N] is in fact a NumeralP, with an empty Numeral head. Tang (1990) suggests that a NumeralP can occur between DP and ClP, as shown in (35). (For Tang (but not for us), demonstratives are Ds. Thus, (35) can be a structural representation of (36).)

(35)

(36)   zhe san   ben shu
        DEM three CL   book
        'these three books'

If it is true that indefinite [Cl + N] involves an empty Numeral head, there are two different structures for surface [Cl + N] strings: (a) [Cl + N] phrases that are ClPs, which yield a definite reading, and (b) [Cl + N] phrases that are NumeralPs with an empty Numeral head, which yield an indefinite reading and have a restricted distribution.

Some questions arise given this account. In particular, (a) why can't Mandarin [Cl + N] phrases yield a definite reading? and (b) do indefinite bare nouns differ from indefinite [Cl + N] phrases in that the former are ClPs and the latter NumeralPs? The following chart will help in answering these questions:

(37)   *The interpretation of the different types of nouns*

	Mandarin		*Cantonese*	
	Indef	Def	Indef	Def
Bare N	+	+	+	−
Cl + N	+	−	+	+
Num + Cl + N	+	−	+	−

(37) shows very clearly that in both languages, noun phrases with overt numerals can only be interpreted as indefinite. That is, overt numerals in a noun phrase consistently lead to an indefinite interpretation. In contrast, both bare nouns and [Cl + N] phrases vary in interpretation (with some language–particular differences).

We suggested above that indefinite [Cl + N] phrases are in fact NumeralPs with an empty Numeral head. Combining this with the claim that noun phrases with overt numerals necessarily yield an indefinite interpretation, owing to the quantificational nature (∃) of numerals, we can make the following generalization:

(38)   The indefinite interpretation of nominals in Chinese is linked to the
        presence of a NumeralP (the head of which may be overt or nonovert).

Superficially, a noun phrase such as *yi–ge xuesheng* 'one-CL student' is interpreted like *one student* in English. However, when classifiers are considered to be singularizers, *yi–ge xuesheng* is more adequately paraphrased as 'one instance of the (kind) student'.
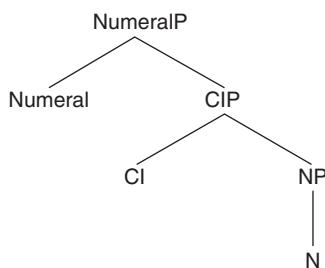
A question arises regarding the indefinite interpretation of bare nouns. Are they also NumeralPs? They must be, in view of (38). In other words, they have not only an empty Cl head, as we have argued, but an empty Numeral head as well.

Consider again the interpretation of definite bare noun phrases. In our proposal, definite bare noun phrases are ClPs, and their definiteness stems from the ι operator, which can be considered to be like a definite article. The numeral apparently
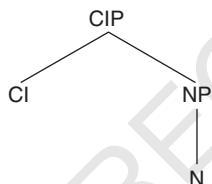
has the effect of undoing the definiteness, just as with [Cl + N] phrases preceded by a numeral. An indefinite bare noun is thus interpreted as '*x* instance of the noun'. This means that both indefinite bare nouns and indefinite [Cl + N] phrases are NumeralPs, the former having an empty Numeral head and an empty Cl head and the latter only an empty Numeral head. This entails that they are essentially interpreted the same way, a result supported by the restriction of the nonspecific reading of bare indefinites and indefinite [Cl + N] phrases (as shown in (33) and (34)).

Under this account, then, surface strings of the form [Cl + N] have two different structural representations: [Cl + N] with an indefinite reading is a NumeralP, as in (39), and [Cl + N] with a definite reading is a ClP, as in (40).

(39)

```
            NumeralP
           /        \
     Numeral         ClP
                    /    \
                  Cl      NP
                           |
                           N
```

(40)

```
         ClP
        /    \
      Cl      NP
               |
               N
```

We propose that the same applies to bare nouns: a surface string in the form of a bare noun can have either the structure in (39) or the structure in (40). Indefinite bare nouns have the structure in (39) (with the Numeral for indefinite interpretation); definite bare nouns have the structure in (40). In other words, the structures in (39) and (40) are in fact the structures for indefinite NPs (surface: Num-Cl–N, Cl–N, N) and definite NPs (surface: Cl–N, N), respectively.

Now consider the fact, shown in (37), that bare nouns in Cantonese, and [Cl + N] phrases in Mandarin, cannot be interpreted as definite. As we have argued, in principle, both languages have the structures in (39) and (40) for indefinite and definite NPs. Thus, the differences must be attributed to language-particular restrictions. What could these restrictions be?

Let us first consider the fact that Mandarin [Cl + N] phrases cannot have a definite interpretation. It is a consequence of our analysis that Mandarin [Cl + N] phrases cannot have the structure in (40). Since the difference between (39) and (40) is that the latter lacks the Numeral projection, the Mandarin restriction suggests that in Mandarin, for reasons we do not understand, overt classifiers simply

cannot appear without a Numeral (even though the Numeral may be nonovert). We state this restriction as follows, leaving open the question of how it can be explained:

(41)   In Mandarin, overt classifiers are always accompanied by a Numeral. The Numeral can be overt or nonovert.

In other words, in Mandarin, whenever there is an overt classifier, it must occur in the structure in (39). Cantonese lacks such a restriction on the occurrence of overt classifiers: in this language the occurrence of a classifier does not automatically imply the presence of a Numeral. Hence, unlike what we see in Mandarin, Cantonese nouns involving a classifier are not necessarily indefinite. This explains why [Cl + N] phrases in Mandarin cannot be interpreted as definite.

Next let us consider why bare nouns in Cantonese cannot be interpreted as definite. Recall that to express definiteness, Cantonese uses [Cl + N] phrases, whereas in Mandarin the bare noun is moved to Cl position (with the ι operator present). In other words, either the Cl position is filled by a classifier (Cantonese), or the ι operator is used, followed by N-to-Cl movement (Mandarin). The question is why Cantonese and Mandarin differ. We would like to suggest that given a choice between insertion of a lexical item and movement, a language always chooses insertion. This is reminiscent of Chierchia's treatment of definite noun phrases in English. Chierchia suggests that "[t]here is a principle that seems to be fundamental for the architecture of grammar which says, roughly, 'Language particular choices win over universal tendencies' . . . or 'Don't do covertly what you can do overtly'" (1998:360). In other words, Cantonese takes the default option. Mandarin, on the other hand, cannot take that option because of the language-particular constraint (41) barring overt classifiers from occurring without a numeral.

[. . .]

3.4   Summary

To conclude this section, let us summarize some of the main claims and hypotheses. We showed that although Chinese count nouns are count mass nouns (semantically count and syntactically mass), there is a count/mass distinction in Mandarin and Cantonese, the two Chinese languages that we examined. Further, bare nouns in Chinese are not really bare: they must at least be embedded in a projection that performs the deictic discourse function of linking the description provided by the noun or noun phrase to a specific entity in the world.

The assumption that N is embedded in another phrase enabled us to straightforwardly apply some of the ideas developed by Longobardi (1994) and account for the distribution of indefinite bare nouns and indefinite [Cl + N] phrases. Assuming that the head of the projection above NP can be empty and that empty heads must be lexically governed, the fact that indefinite bare nouns and [Cl + N] phrases are distributionally restricted to lexically governed positions, and their definite counterparts are not, suggests that the former involve an empty head and the latter do not.

This led to postulating the structures in (39) and (40). The structure in (40) represents definite bare nouns and [Cl + N] phrases: the NP is embedded in a ClP, which (among other things) performs the deictic function assumed to be performed by D in Romance and Germanic languages. We noted that to express definiteness, either the nonovert ι operator or an overt classifier is used; the former option leads to N-to-Cl movement.

The structure in (39) represents indefinite nouns. This structure contains not only the deictic ClP, but also a NumeralP, which is responsible for the indefiniteness (due to the existential quantificational force associated with numerals).

The differences between Mandarin and Cantonese can all be traced back to the restriction in (41), namely, in Mandarin, but not in Cantonese, overt classifiers can occur only with a NumeralP, the head of which may or may not be overt. This, we argued, bars Mandarin from filling the Cl⁰ slot by inserting a classifier, which seems to be the default procedure. Instead, Mandarin must resort to the ι operator (and thus N-to-Cl movement). This explains why [Cl + N] phrases can only be indefinite in Mandarin (the Cl subsumes the presence of a Numeral and is thus indefinite) but indefinite or definite in Cantonese and why bare nouns can be both definite and indefinite in Mandarin but only indefinite in Cantonese (Cantonese only has N-to-Cl movement with generics and proper names because it lacks the restriction formulated for Mandarin in (41)). Another difference connected to the constraint in Mandarin is that definite nominals are neutral for number when they are bare, whereas [Cl + N] phrases are never neutral for number, owing to the presence of the classifier. We furthermore noted that nominals embedded in a NumeralP with an empty head are always nonspecific.

26.3   Questions pertaining to Cheng and Sybesma (1999)

1   Bare nouns in subject position in Mandarin or Cantonese cannot be interpreted as indefinite, whereas in English one can have sentences like *Hurry up, people are waiting for you, Are planes flying overhead right now, do you think?, Books are being published these days that haven't even been proofread*. Discuss the (de)merits of the following proposal: Subject position in such English sentences is filled, not by the bare plural DP itself, but by a larger remnant phrase containing it.

2   Cheng and Sybesma adopt an idea that goes back to Szabolcsi (1987, 1994) and Stowell (1989) according to which arguments must be DPs. Discuss the relevance to this idea of English compounds found in examples such as *an avid magazine-reader, a cat-loving linguist*, bringing in Baker (1988) on noun-incorporation.

3   Cheng and Sybesma take the position that classifiers of the Chinese kind head their own projection. That would imply, if Kayne (2008a) is right, that Chinese classifiers must be verbal. Alternatively, they are nominal and do not head their own projection. Discuss the plausibility of this alternative, bringing in the question of light nouns, as well as Koopman (2003, 2005) and Taraldsen (2012, section 4.3).

4 Cheng and Sybesma take Mandarin (but not Cantonese) phrases of the form "classifier + noun" to necessarily be accompanied by a numeral. If in Mandarin such a phrase is not preceded by an overt numeral, then there must be a silent one present. In this context, discuss the plausibility of the following proposal, bringing in as many other languages as possible: A silent numeral is necessarily present in English *all of the books*.

5 Following Doetjes (1996, 1997), Cheng and Sybesma suggest that "numerals require the presence of a syntactic marker of countability." To what extent is this suggestion compatible with English having *a three-drawer file cabinet, a two-time winner of the marathon*?

6 What is the significance for Cheng and Sybesma's analysis of the fact that in Japanese the classifier can be separated from the following noun by a genitive-like element *no*? Bring in Watanabe (2006).

7 Thinking of Zweig's (2005) proposal that numerals are accompanied by a silent noun NUMBER, discuss the similarities and differences between the overt noun number in English (or a counterpart in another language) and Chinese classifiers.

8 Chinese appears to have no classifier distinction for male vs. female. Nor does it have gender distinctions in its pronouns. To what extent are these two properties related? Bring in as many other languages as possible.

9 Cheng and Sybesma take the position, in agreement with some others, that Mandarin *-men* is a collective suffix, rather than a plural suffix, with one relevant property being that *-men* is incompatible with numerals. Find as many languages as possible with a morpheme similar to Mandarin *-men*. To what extent could English *group* be taken to be a true counterpart of *-men*?

10 As part of their argument that "the count/mass distinction is clearly reflected in the classifier system," Cheng and Sybesma note that "a modification marker *de* can intervene in [massifier + N] sequences but not in [count-classifier + N] sequences." For example, one can have *san bang (de) rou* ('three massifier/pound *de* meat'), but not *ba tou (*de) niu* ('eight count-classifier *de* cow'). How might this contrast within Chinese be related to the English distinction between *a kilo's worth of apples* and **ten's worth of apples*?

11 Cheng and Sybesma emphasize the way in which Cantonese classifiers can apparently play the role of definite articles. How might this (partial) link between numeral classifiers and definite articles be enhanced by the study of Moroccan Arabic?

12 Continuing the previous question, how might one bring to bear that fact that Belgian French allows, with an indefinite interpretation, *beaucoup de la bière* ('great-deal of the beer') where standard French would have *beaucoup de bière*, without the definite article. To what extent might the standard French pair *Elle en a beaucoup, des amis* ('she of-them has great-deal, of-the friends') and *Elle en a beaucoup, d'amis* ('she of-them has great-deal, of friends') be relevant?

13 Cheng and Sybesma have covert N(oun) raising to Cl(assifier) in Mandarin bare definites, recalling in part Longobardi (1994, Ch. 21 of this volume). Such covert noun-raising would not be available in languages that disallow bare NPs being interpreted as definites. The distinction between languages

like Mandarin that allow bare definites (cf. also Déprez 2005) and languages like English that don't would seem, though, to have something in common with the distinction between languages like Italian that allow silent definite articles in contexts discussed by Kayne (2008b) and languages like French that don't, in the same contexts. Discuss the possible implications of the Italian vs. French contrasts on the Mandarin vs. English contrasts.

14 Cheng and Sysbema's "massifiers" might well be taken to include English *amount*. Discuss, then, the implications for the analysis of massifiers of the fact that in English one has *He generally spends only a short amount of time on his work*, but *He generally puts only a small/*short amount of sugar in his coffee*.

15 Discuss the ways in which the facts of the previous question do or do not fit together with Selkirk's (1977) observation that some English allows *those kind of horses*.

References

Baker, M. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Chierchia, G. 1995. Plurality of mass nouns and the notion of "semantic parameter." Ms., University of Milan.

Chierchia, G. 1998. Reference to kinds across languages. *Natural Language Semantics* 6: 339–405.

Déprez, V. 2005. Morphological number, semantic number and bare nouns. *Lingua* 115: 857–883.

Di Sciullo, A. M., and C. Boeckx (eds). 2011. *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford Studies in Biolinguistics. Oxford and New York: Oxford University Press.

Doetjes, J. 1996. Mass and count: Syntax or semantics? In A. Arregui and C. Cremers (eds), *Proceedings of Meaning on the HIL. HIL Occasional Papers in Linguistics* 1, Leiden: Holland Institute of Generative Linguistics, Leiden University, pp. 34–52.

Doetjes, J. 1997. *Quantifiers and selection: On the distribution of quantifying expressions in French, Dutch and English*. Doctoral dissertation, Leiden University.

Kayne, R. S. 2008a. Antisymmetry and the lexicon. In J. Rooryck and J. van Craenenbroeck (eds), *Linguistic Variation Yearbook* 8, Amsterdam and Philadelphia: John Benjamins, pp. 1–31. [Reprinted in Kayne (2010, Ch. 9) and in Di Sciullo and Boeckx (2011, Ch. 14).]

Kayne, R. S. 2008b. Comparative syntax and the lexicon. In J. Durand, B. Habert, and B. Laks (eds), *Proceedings of the congrès mondial de linguistique française – CMLF'08*, Paris: Institut de Linguistique Française, pp. 73–85.

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Koopman, H. 2003. Inside the noun in Maasai. In A. Mahajan (ed.), *Syntax at Sunset 3: Head Movement and Syntactic Theory. UCLA Working Papers in Linguistics* 10, Los Angeles: UCLA Linguistics Department, pp. 77–115.

Koopman, H. 2005. On the parallelism of DPs and clauses: Evidence from Kisongo Maasai. In A. Carnie, H. Harley, and S. A. Dooley (eds), *Verb First: On the Syntax of Verb-initial Languages*, *Linguistik Aktuell/Linguistics Today* 73, Amsterdam and Philadelphia: John Benjamins, pp. 281–302.

Leung, C.-S. 1980. Guo–yue yu bijiao yufa zhaji – dingzhi zhishici (Notes on comparative grammar of Mandarin and Cantonese – markers of definiteness). *Zhong–Ying Yuwen Xue–Yan (Language Learning and Research)* 1: 15–28.

Li, C., and S. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley and Los Angeles: University of California Press.

Longobardi, G. 1994. Reference and proper names: A theory of N-Movement in syntax and logical form. *Linguistic Inquiry* 25: 609–665.

Matthews, S., and P. Pacioni. 1996. Specificity and genericity of NPs in Cantonese and Mandarin. Paper presented at the Symposium on Referential Properties of Chinese Noun Phrases, City University of Hong Kong.

Matthews, S., and V. Yip. 1994. *Cantonese: A Comprehensive Grammar*. London: Routledge.

Paris, M.-C. 1981. *Problèmes de syntaxe et de sémantique en linguistique chinoise*(Problems of syntax and semantics in Chinese linguistics). Paris: Collège de France.

Partee, B. 1987. Noun phrase interpretation and type-shifting principles. In J. Groenendijk and M. Stokhof (eds), *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, Dordrecht: Foris, pp. 115–143.

Selkirk, E. 1977. Some remarks on noun phrase structure. In P. W. Culicover, T. Wasow, and A. Akmajian (eds), *Formal Syntax*, New York: Academic Press, pp. 185–316.

Stowell, T. 1989. Subjects, specifiers, and X-bar theory. In M. R. Baltin and A. S. Kroch (eds), *Alternative Conceptions of Phrase Structure*, Chicago and London: University of Chicago Press, pp. 232–262.

Sybesma, R. 1992. *Causatives and accomplishments: The case of Chinese* ba. Doctoral dissertation, HIL/Leiden University.

Szabolcsi, A. 1987. Functional categories in the Noun Phrase. In I. Kenesei (ed.), *Approaches to Hungarian* 2, Szeged: JATE, pp. 167–189.

Szabolcsi, A. 1994. The noun phrase. In F. Kiefer and K. É Kiss (eds), *The Syntactic Structure of Hungarian. Syntax and Semantics* 27, San Diego: Academic Press, pp. 179–274.

Tang, C.-C. J. 1990. *Chinese phrase structure and the extended X′-theory*. Doctoral dissertation, Cornell University.

Taraldsen, K. T. 2012. The structural object position of verbs and nouns. In L. Brugè, A. Cardinaletti, G. Giusti, N. Munaro, and C. Poletto (eds), *Functional Heads. The Cartography of Syntactic Structures* 7, New York and Oxford: Oxford University Press, pp. 263–274.

Watanabe, A. 2006. Functional projections of nominals in Japanese: Syntax of classifiers. *Natural Language and Linguistic Theory* 24: 241–306.

Zweig, E. 2005. Nouns and adjectives in numeral NPs. In L. Bateman and C. Ussery (eds), *Proceedings of NELS 35*, Amherst: GLSA, University of Massachusetts, pp. 663–678.

27

Remarks on Holmberg's Generalization
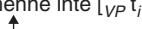
Anders Holmberg

1999

27.1 Introduction

In the Germanic languages, the object sometimes moves out of the VP to a position to the left of the marker of sentential negation. In connection with North Germanic (i.e., Danish, Swedish, Norwegian, Faroese, and Icelandic) this is called Object Shift (Holmberg 1986), whereas in connection with West Germanic (German, Swiss German, Frisian, Dutch, Afrikaans, English, Yiddish, etc.) it is called Scrambling (on the comparability of the two phenomena, see Richards 2004). Two questions that arise are the following: (a) Does Object Shift depend on internal properties of the object? (b) Are there other factors, independent of the properties of the object, that play a role in Object Shift?

Focusing on North Germanic, Holmberg shows that Object Shift is indeed sensitive to internal properties of the object: lexical DPs optionally undergo Object Shift only in some of the Scandinavian languages. In contrast, weak pronouns obligatorily undergo Object Shift in Danish and most varieties of Norwegian, optionally in most varieties of Swedish and some varieties of Norwegian. An example from Swedish, with the weak pronoun *henne*, is given in (1a), with the movement of the object and the verb indicated in (1b):

(1) a. Jag kysste henne inte. (Swedish)
         I   kissed her   not

     b. Jag kysste$_i$ henne inte [$_{VP}$ t$_i$ t]

Much of the discussion in the literature on Object Shift has focused on the second question, namely what are the external conditions on Object Shift.

The answer that has received the most attention was offered in Holmberg (1986) and has come to be known as "Holmberg's Generalization." The core of the generalization can be expressed as follows: Object Shift of the complement of a verb occurs only if the verb has moved out of the VP. For example, Object Shift is possible in (1a) because the finite verb *kysste* has moved out of the VP, as indicated by the fact that it occurs to the left of the negative marker *inte*. But it is not possible in (2a) nor in (2b), because *kysste* is within the VP:

(2) a. *. . . att jag henne$_i$ inte [$_{VP}$ kysste t$_i$]. (Swedish)
        . . . that I her   not [^{VP} kissed t$_i$]
    b. *Jag har henne$_i$ inte [$_{VP}$ kysst t$_i$].
        I   have her   not [$_{VP}$ kissed t$_i$]
    c. Jag har inte [$_{VP}$ kysst henne].
        I   have not   kissed her

Holmberg (1986) had also pointed out that the object had to be "phonetically adjacent" to the element around which it is shifted, that is to say, that Object Shift is not possible if the object is preceded by a preposition, an indirect object (in the case of double object constructions) or a particle, as shown in (3):

(3) a. *Jag talade henne$_i$ inte med t$_i$. (Swedish)
        I   spoke her   not with
    b. *Jag gav den$_i$ inte Elsa t$_i$.
        I   gave it   not Elsa
    c. *Dom kastade mej$_i$ inte ut t$_i$.
        they threw me   not out

In the present article, *Remarks on Holmberg's Generalization*, Holmberg attempts to find a unified explanation for the ungrammatical examples in (2) and (3). He does so by arguing that the dependence of Scandinavian Object Shift on verb movement is a special case of a more general condition that prevents Object Shift from taking place across any phonologically overt category within VP. In other words, Holmberg argues, Object Shift is sensitive to the phonological overtness of potential interveners, rather than to their morphosyntactic features.

In this paper, Holmberg concludes that Object Shift is a PF-phenomenon; more precisely, noting that shifted objects cannot be focused, he proposes that it is induced by a phonological feature [–Focus]. In section 9, he also contemplates, along the lines of Kayne's (1998) proposals concerning leftward VP-fronting within IP, the possibility that the objects which we call "shifted" have actually never moved out of VP, but have rather been carried along with the verb by VP-fronting. While such an approach faces some empirical challenges, it automatically derives the dependence of Object Shift on verb movement (cf. Nilsen 2003).

The sensitivity of Object Shift to the overt position of VP-material has been further explored in Fox and Pesetsky (2005). They suggest that constraints on object shift derive from the process of linearization of syntactic structure: at the

end of each cycle/phase, precedence relations are imposed on the words present in a syntactic structure, and linearization in subsequent cycles/phases must respect previous assertions of precedence. Consequently, if the verb precedes the object in the vP cycle, it must continue to do so in all other cycles. Therefore the object can precede material inserted in later cycles (such as the negative marker) only if the verb does so as well.

27.2   From "REMARKS ON HOLMBERG's GENERALIZATION"

1   Introduction

Scandinavian Object Shift is dependent on verb movement in the sense that an unmoved verb will always block Object Shift, as shown in (1). (All examples in this paper are Swedish, except where indicated otherwise; t_v = verb trace, t_o = object trace, t_s = subject trace, t_{io} = indirect object trace.)

(1)   a.   Jag kysste henne inte $[_{VP} t_v t_o]$ a′.   (*)Jag kysste inte henne.
           I   kissed her   not   I   kissed not her
      b.   *Jag har henne inte $[_{VP}$ kysst $t_o]$.   b′.   Jag har inte kysst henne.
           I   have her   not   kissed   I   have not kissed her
      c.   *. . . att   jag henne inte [VP kysste t_o].   c′. … att   jag inte kysste henne.
               that I   her   not   kissed   that I   not kissed her

(1a) is a licit application of Object Shift derived by V-to-I-to-C and Object Shift of a weak pronoun, while (1b, c) are illicit, by hypothesis because the verb governing the object position has not moved. In (1b) the verb has not moved because the auxiliary verb blocks movement of the main verb, and in (1c) because there is no verb moement [sic] in embedded clauses in Swedish (and generally Mainland Scandinavian). The status of (1a′) varies across Scandinavian languages and dialects; in Danish and most varieties of Norwegian it is unacceptable if the object pronoun is weak, i.e. unstressed and simple. In most varieties of Swedish and some varieties of Norwegian it is acceptable even when the object is a weak pronoun. This interplay of verb movement and Object Shift is well known, and referred to as "Holmberg's Generalization" (henceforth HG) in some of the literature.[1] Less often mentioned, but no less true, is the fact that not just an unmoved verb, but any phonologically visible category inside VP preceding the object position will block Object Shift.[2] Consider the examples in (2):

(2)   a.   *Jag talade henne inte med t_o.   a′.   Jag talade inte med henne.
           I   spoke her   not with   I   talked not with her
      b.   *Jag gav den inte Elsa t_o.   b′.   Jag gav inte Elsa den
           I   gave it   not Elsa   I   gave not Elsa it
      c.   *Dom kastade mej inte ut t_o.   c′.   Dom kastade inte ut mej.
           they threw me not out   they threw not out me

In all of these examples the verb has moved, yet Object Shift is illicit: in (2a) because it has shifted across a preposition, in (2b) because it has shifted across an indirect object, and in (2c) because it has shifted across a verb particle. Note that the object always follows the verb particle in Swedish, even when it is a pronoun. In Danish the object always precedes the verb particle, while the other Scandinavian languages (i.e. Norwegian, Icelandic, and Faroese) are like English in that the object has to precede the particle if it is a pronoun and may do so if it is a lexical DP.

(3)   a.   Jeg skrev (nummeret/det)   op (*nummeret/*det).   (Danish)
       b.   Jeg skrev (nummeret/det)   opp (nummeret/*det).   (Norwegian)
       c.   Jag skrev (*nummeret/*det) upp (numret/det).   (Swedish)
            I   wrote (the-number/it)   up (the-number/it)
            'I wrote the number/it down.'

In all the Scandinavian languages except Swedish Object Shift can apply to a pronominal object of a verb particle construction. Apparently the particle blocks Object Shift in Swedish. Compare (2c) with (4a, b):

(4)   a.   Jeg skrev det måske ikke t_o op.   (Danish)
            I   wrote it   maybe not up
       b.   De   kastet meg ikke t_o ut.   (Norwegian)
            they threw me   not out

The question is: Is there a unified explanation for the cases in (1) and (2), or must we assume different explanations, perhaps one for (1b, c) and up to three distinct explanations for (2a, b, c)?

   Holmberg (1986) proposed a unified explanation of the facts illustrated in (1) and (2) which makes crucial reference to phonological visibility: If the object is governed by a VP-internal phonologically visible category it will be assigned Case in situ, hence Object Shift is ruled out by the Last Resort condition (in modern terms). This is also the position taken in Holmberg & Platzack (1995). In most, or all, other literature on the subject it seems to be implicitly assumed that there are other explanations for the facts in (2) than for the facts in (1). The only case from (2) which has been discussed at any length, though, is the double object construction. Thus Vikner (1989) argues that (2b) is ruled out by Relativized Minimality, preventing movement of the direct object across the indirect object in the double object construction. More recently Collins & Thráinsson (1996) have hypothesized that (2b) is ruled out by a condition which is specific to double object constructions, or at any rate has nothing to do directly with the conditions ruling out (1b, c); cf. (Collins & Thráinsson 1996:420ff). Since we know that movement in the double object construction is constrained in various ways, at least in part for reasons involving Relativized Minimality or other such fundamental syntactic conditions on movement, it seems initially plausible that (2b) would be another effect of (one of) these conditions. This seems less plausible in the case of (2c), though. None of the well-known conditions on movement predict that DP-movement would be blocked by a verb-particle.

I will show in this paper that the interplay between Object Shift and a construction which I will call Verb Topicalization provides strong support for a unified explanation of the facts in (1) and (2), and that the explanation makes crucial reference to phonological visibility. The Verb Topicalization facts show that Scandinavian Object Shift is free to apply whenever all phonologically visible non-adjunct material has been removed from between the launching site and the landing site of the movement, regardless how it is removed (by head movement, A-movement, or A-bar movement). Correspondingly, no trace, whatever its source, will ever block Object Shift. This is expected if Object Shift is a PF-operation (i.e. applying after spell-out, in the phonological component). The possibility that Object Shift is a PF-operation was discussed in Holmberg (1986) and Holmberg & Platzack (1995), but was rejected there based on a single argument, which can now be shown not to hold water. Once we accept this hypothesis, a number of peculiar properties of Object Shift, including HG, fall out in a more natural and more unified manner than under the hypothesis that Object Shift is a pre-spell-out syntactic operation. On the other hand it is unclear what it means for an operation like Object Shift to apply in the phonological component. Although it does make crucial reference to phonological visibility it is not a typical phonological operation, since it does not make reference to phonological primitives such as vowels, feet, nasality, etc. Nor is it a morphological process, since it does not make reference to primitives such as stems or affixes etc. One approach is to ascribe it to a component of 'stylistic rules' (cf. Chomsky & Lasnik (1977), Rochemont (1978), Chomsky (1995:324f)), which is post-spell-out in that it has access to certain phonological features, namely prosodic features, but feeds the phonological component proper. This is the tack taken in this paper. I will argue that Object Shift is an operation on a feature [–Focus], which is inserted together with other phonological features 'at spell-out', that is at the output of the formal syntactic derivation constructing sentences from the items in the Numeration.

I will discuss mainly Object Shift of weak pronouns, which is the most general form of Object Shift found in Mainland Scandinavian. There is a controversy regarding the relation between shift of weak pronouns and shift of full DPs (lexical DPs and proper names), as found mainly in Icelandic: Are they instances of the same operation or are they fundamentally different? Cf. Holmberg & Platzack (1995: ch. 6) for discussion. My position is that they are instances of the same operation. One good reason to think that they are is that they are both subject to HG, in exactly the same way.

Recently Nilsen (1997) has shown that full DP Object Shift is not impossible in Mainland Scandinavian either. One of Nilsen's Norwegian examples, slightly modified, is (5).

(5) Etter dette slo   Guri (Per) heldigvis (?Per) ikke (Per) lenger (Per)
     alltid   (Per) i   sjakk.
     after this beat Guri (Per) luckily   (Per) not (Per) longer (Per)
     always (Per) in chess
     'After this, Guri luckily didn't anymore always beat Per in chess.'

As indicated, the object *Per* may precede any of the adverbs, with the possible exception of the negation. Speakers' reactions vary somewhat when presented with such sentences, but apparently many Norwegian and Swedish speakers accept them quite happily. The conclusions that Nilsen draws from his observations are very much in line with the conclusions of the present paper, namely that Object Shift is, in some sense, a PF-phenomenon. There is still a clear difference between Icelandic and Mainland Scandinavian as regards full DP Object Shift. For example, if a lexical DP is substituted for the pronoun in (1a) the result is clearly ungrammatical in Mainland Scandinavian (although it is improved if the negation is contrastively stressed), but entirely well formed and unmarked in Icelandic. But the explanation of this difference will have to be reconsidered.

 [. . .]

3   Verb topicalization

Consider the construction exemplified in (11) (brought to my attention by Tarald Taraldsen, p.c.):

(11)   a.   Kysst har   jag henne inte (bara hållit henne i   handen).
            kissed have I   her   not only held her   by the hand
       b.   Sett har han mej kanske (men han vet   inte vad   jag   heter).
            seen has he   me perhaps but he   knows not what I am-called

This looks like a case of VP-fronting, except that the object has been left behind, and has undergone Object Shift. The semantic/pragmatic effect is that the verb is contrastive. (9) is in this respect similar to (12), an ordinary case of VP-fronting.

(12)   a.   Kysst henne har   jag inte (bara hållit henne i   handen).
            kissed her   have I   not only held her   by the hand
            'Kissed her I haven't (only held her by the hand).'
       b.   Sett mej har han kanske (men han vet   inte vad   jag   heter).
            seen me has he perhaps but he   know not what I am-called.
            'Seen me he may have done (but he doesn't know my name).'

A possible, initially plausible analysis of (11) is that it is a case of Remnant Topicalization, a term coined by den Besten & Webelhuth (1987) for the German construction (13):

(13)   a.   Ein Buch gegeben hat er dem Jungen nicht.
            a   book given   has he the boy   not
            'He didn't give the boy a BOOK.'

The analysis which den Besten & Webelhuth argued for is that the indirect object has first been scrambled out of VP, after which the VP containing the trace of the indirect object has been topicalized. Analogously, we might analyze (11) as derived

by Object Shift of the pronoun out of VP followed by Remnant Topicalization of the VP to specCP. Note, however, that in order to derive (11) this way, Object Shift must apply across an unmoved verb, violating HG. But then the structure is repaired by topicalizing the VP, thus in a way obliterating the traces of the violation. This derivation is sketched in (14):

(14)   a.   Jag har henne inte [$_{VP}$ t$_s$ [$_{V'}$ kysst t$_o$]] (Object Shift violating HG)
       b.   [$_{VP}$ t$_s$ [$_{V'}$ kysst t$_o$]] har jag henne inte t$_{VP}$ (VP-topicalization)

This amounts to saying that HG is due to a 'surface filter': Violation of HG is all right so long as the structure undergoes other operations which yield the surface order where the object is followed by adjuncts but no visible head preceding the original object position. (15) is a formulation of the putative filter, applying, we assume, to S-structure.

(15)   The Object Shift Filter: *Obj Adv X^o t$_o$, unless X^o is phonologically empty.

(14.a) would be filtered out by (15) at S-structure, but (14.b) would not. So given that other conditions are met, (14.b) is well formed. It can be shown, however, that (15) is not empirically adequate, and that (14) is therefore most likely not the correct derivation of (11). Consider the following examples:

(16)   a.   Jag hörde henne inte [$_{SC}$ t hålla föredrag].
            I   heard her   not   give talk
            'I didn't hear her give a talk'
       b.   *Jag har henne inte [$_{VP}$ hört [$_{SC}$ t hålla föredrag]]
            I   have her   not   heard   give talk
(17)   a.   Tidsskillnaden gör mej alltid [$_{SC}$ t förvirrad].
            the-time-difference makes me always confused
       b.   *Den har mej alltid gjort [$_{SC}$ t förvirrad].
            it has me always made confused
(18)   a.   Han låter den aldrig sjunka till botten.
            he   let   it   never sink to the- bottom
       b.   *Han har den aldrig låtit sjunka till botten.
             he   has it   never le sink to the- bottom

(16), (17), and (18) each contains a verb taking a SC complement. The subject of the SC is a weak pronoun which has undergone Object Shift up into the matrix clause. Object Shift is licit in the (a)-sentences, but illicit in the (b)-sentences where it crosses the unmoved main verb. The question now is, can we repair the (b)-sentences by topicalizing the VP containing the SC? The filter-based theory sketched above predicts that we can. The prediction is false, however: (19), (20), and (21) are ill-formed.

(19)   *Hört hålla föredrag har   jag henne inte.
         heard give talk   have I   her   not

(20)   *Gjort förvirrad har den mej alltid.
         made confused has it   me always

(21)   *Låtit sjunka till   botten har han den aldrig.
         let sink   to the-bottom has he it   never

Topicalizing a VP containing a SC is not a problem in general, as shown by (22).

(22)   a.   Hört henne hålla föredrag har   jag inte.
              heard her   give talk   have I   not
       b.   Gjort mej förvirrad har den alltid.
              made me confused has it   always
       c.   Låtit den sjunka   till   botten har han aldrig (gjort).
              let   it sink   to the-bottom has he   never (done)

Furthermore, constructions corresponding to (19–21) are well formed in German (Gert Webelhuth, p.c.), showing that Remnant Topicalization after scrambling/ object shift of a small clause subject out of the matrix VP is tolerated by UG. The structure of (23a) should be roughly (23b):

(23)   a.   Rauchen gelassen hat er seine Tochter nicht.
              smoke   allowed has he his   daughter not
              He hasn't allowed his daughter to smoke.
       b. $[_{VP} [_{SC} t_s$ rauchen] gelassen] hat er [seine Tochter]$_s$ nicht t$_{vp}$

The conclusion we can draw from this is that (19)–(21) are ill-formed because their derivation includes an illicit application of Object Shift. That is to say, HG is not a matter of S-structure word order but derivation: A violation of HG cannot be repaired by subsequent operations. This implies that (11a) is not derived as shown in (14), i.e. by Object Shift followed by Remnant Topicalization. Instead I claim that it is derived by Verb Topicalization to specCP, followed by Object Shift. The derivation is as shown in (24):

(24)   Infl [inte [har [$_{VP}$ jag [$_{V}'$ kysst   henne]]]]
              not   have I   kissed her
       $[_{CP}$ kysst$_V$ [$_{C}'$ har$_{aux}$ [$_{IP}$ jag$_s$ Infl inte t$_{aux}$ [$_{VP}$ t$_s$ [$_{V}'$ t$_v$ henne]]]]]
       $[_{CP}$ kysst$_V$ [$_{C}'$ har$_{aux}$ [$_{IP}$ jag$_s$ Infl henne$_o$ inte t$_{aux}$ [$_{VP}$ t$_s$ [$_{V}'$ t$_V$ t$_o$]]]]]

(24a) is the underlying structure. In (24b) the subject has moved to specIP, the finite auxiliary verb to C, presumably via Infl, while the main verb has moved to specCP. In (24c), finally, the object has shifted to a position preceding the negation.

    [. . .]

Sten Vikner (p.c.) points out that the theory, as it stands, predicts that (28) should be well formed, which it is not. The same point is made in Kaiser (1997).

(28)   ?*Sett har jag honom inte röka   (men jag har känt   hans andedräkt).
            seen have I   him   not smoke (but I   have smelled his   breath)

Here the verb has been topicalized, permitting Object Shift of the SC subject. We expect it to be roughly equal to (29).

(29)   SETT honom röka   har   jag inte (men . . .)
        seen   him   smoke have I   not   (but . . .)
        'I haven't SEEN him smoke (but . . .)'

Note first that (28) is ill formed independently of Object Shift:

(30)   *Sett har   jag inte Per   röka   (men . . .)
         seen have I   not   Peter smoke (but . . .)

It seems you cannot topicalize a verb stranding a small clause complement. This casts some doubt on the analysis of (11a, b) as derived by V-topicalization. Perhaps it is Remnant VP-topicalization after all? [. . .]

    At this point we need to consider the properties of the construction which I have called V-topicalization in a little more detail, even though a full discussion is beyond the scope of this paper. The construction has not been discussed in the literature before, to the best of my knowledge, in relation to the Scandinavian languages. To begin with it is not universally accepted: Some speakers are uncomfortable with the construction, while other speakers accept it without hesitation. As shown by the examples (11a, b, [. . .]) the movement can strand a nominal object. For some reason this is most acceptable when the verb moved is a participle in the periphrastic perfect or pluperfect construction. (31), where an infinitival main verb is fronted, is dubious, and (32), where a tensed verb is fronted, in construction with the pro-verb *göra* 'do' is clearly not well formed (Mainland Scandinavian allows fronting of tensed VPs; cf. Källgren & Prince, 1989)).

(31)   *?Träffa ska   jag henne inte, men vi   ska   hålla kontakt per e-mail.
          meet   shall I   her   not   but we shall keep in touch by e-mail
(32)   *Säljer gör   han den inte (men han kanske lånar ut den   ibland).
          sells does he   it   not (but he   may   lend it from time to time)

If grammatical, (32) would mean 'He won't sell it, but . . .'. A PP complement may, however, be stranded by an infinitival or tensed verb:

(33)   Bo ska han i Malmö, men han ska jobba i   Köpenhamn.
        live will he in Malmö, but he   will work in Copenhagen
        'He will live in Malmö, but work in Copenhagen.'
(34)   a. Såg   gjorde han på henne (men han sa   ingenting).
            looked did   he at her   (but he said nothing)
            'Look at her he did, but . . .'
        b. *Såg gjorde han henne (men . . .)
             saw did   he her   (but . . .)

Note the contrast between (34a, b), where the complement of the verb *se* in (34a) is a PP (in which case the reading is 'look at'), while in (34b) it is a DP (in which case the reading is 'see'). The acceptability contrast between (34a and b) implies that the problem in (31) and (32) is licensing the stranded nominal object. Let us say that the trace of the topicalized infinitival verb cannot assign (or check) the Case of the stranded nominal object. This problem is not faced by (33) or (34a), where the stranded argument is governed by a preposition.

Finally, as illustrated by (28) and (30), a small clause cannot be stranded in the V-topicalization construction. I suggest that the problem here, too, is licensing of a stranded nominal, namely the small clause subject. If so, we can tentatively formulate the following generalization: Stranding of a nominal complement by V-topicalization is highly restricted: It is possible principally if (a) the main verb is a participle, and (b) the nominal complement is a direct object of the verb (i.e. it is assigned a theta-role by the verb).

I conclude, therefore, that (28) does not provide good enough reason to abandon the V-topicalization analysis of constructions such as (11a, b) in favour of Object Shift plus Remnant VP-fronting.

[. . .]

6   Object Shift and Case

I will begin by reviewing the account of HG in Holmberg (1986), basically taken over by Holmberg & Platzack (1995: ch. 6), which I believe is on the right track but wrong on one crucial point.

[. . .]

The crucial error in Holmberg (1986), Holmberg & Platzack (1995), and Vikner (1994) is the assumption that Case is the feature triggering Object Shift. True, with a few exceptions (mainly some argument-like locative proforms; cf. Haider, Olsen & Vikner 1995) Object Shift moves only nominal categories, which is consistent with the hypothesis that Case is crucial. But as noted, in fact Object Shift affects only a subcategory of nominal categories, namely definite, light, nonfocused nominals, and in the case of pronouns, only weak pronouns, with some crosslinguistic variation regarding the range of nominal types affected. So the triggering feature seems to be a feature distinguishing between nonspecific, heavy, focused, and (for pronouns) strong nominals on the one hand, and specific, light, nonfocused, and (for pronouns) weak nominals on the other hand. I propose that the crucial feature is [± Foc]: Object Shift affects only nominal objects which are [−Foc]. This captures what I take to be common for Scandinavian Object Shift, Scrambling, and Clitic Movement: They move arguments which are not focused out of VP, with crosslinguistic variation regarding the range of argument categories which are so moved as well as regarding the manner of movement.

[. . .]

Before elaborating this hypothesis further, I will add a few more arguments against the hypothesis that Object Shift is triggered by Case. One of the reasons

in Holmberg (1986) for assuming that Case is crucial was the observation that only categories with morphological Case undergo Object Shift: In Icelandic, where lexical nouns have morphological case, full DPs undergo Object Shift, while in Mainland Scandinavian, where only pronouns have morphological case, only pronouns undergo Object Shift. The correlation is extremely weak, though: To begin with, full DPs undergo Object Shift to some extent in Mainland Scandinavian, too (cf. Nilsen (1997) and section 1 above). Second, as noted by Vikner (1994), Faroese has morphological case on nouns, yet follows the Mainland Scandinavian rules for Object Shift. Third, strong pronouns have case morphology just as much as weak pronouns (cf. Cardinaletti & Starke, [1999]), yet Object Shift of strong pronouns is restricted in the same way as Object Shift of full DPs in Mainland Scandinavian.

Furthermore, Object Shift is not movement to check a particular Case, such as accusative. Nominal objects undergo Object Shift in the same way regardless of their Case (cf. Vikner 1994, 1995:173f.). Even nominative objects (triggering verb agreement in the case of Icelandic) in ergative, passive, and psych-verb constructions undergo Object Shift if other conditions are met.

(51) Mér likar hún/tölvan ekki.   (Icelandic)
      me-DAT like-3SG it/the-computer-NOM not
      'I don't like it/the computer.'

Hence if the crucial triggering feature is Case, it is a general [+Case] without a specific value. This is more or less the same as saying that the crucial feature is [+Nominal]. But not all nominals undergo (overt) Object Shift. In particular, focused nominals do not. So the triggering features apparently include something like [−Foc].

[. . .]

9 A remark on Kayne (1998)

In a recent paper R. Kayne outlines a theory of sentence structure with potentially interesting consequences for Object Shift and HG. Very briefly, the derivation of a sentence in a VO language such as English involves first movement of focused material out of VP, followed by movement of the remnant VP to a position preceding the focus position. In his paper Kayne discusses mainly various types of overt focusing particles, such as *only*, *too*, etc., and the negation. These particles are heads situated just outside VP to the spec of which the focused part of VP moves. However, one may consider extending the theory to (virtually) all sentences as follows: Assume that all material in the VP which is focus in the wider sense of 'new information', including all indefinite DPs, must move out of VP, while constituents conveying old information can, or must stay within VP. Subsequently the remnant VP moves to a higher position, preceding all sentence adverbs including the negation. Assume the language is Icelandic. The derivation of (61) will then be as shown in (62):

(61) Ég les aldrei nyjar bækur.
     I read never new books

(62) a. [[nyjar bækur]$_i$ [$_{VP}$ ég les t$_i$]] (Move indefinite DP out of VP.)
     b. [aldrei [[nyjar bækur]$_i$ [$_{VP}$ ég les t$_i$]]] (Merge the adverb.)
     c. [[$_{VP}$ ég les t$_i$]$_j$ [aldrei [[nyjar bækur]$_i$ t$_j$]]] (Move VP.)

This structure is then presumably followed by movement of the verb to a higher head position (I or C) and the subject to a higher spec position, emptying the preposed VP. Now assume a definite DP can but need not move out of VP. Then the analysis of (63) will be as shown in (64), prior to verb movement and subject movement out of the preposed VP:

(63) Ég les þessar bækur aldrei.
     I read these books never

(64) [[$_{VP}$ ég les þessar bækur]$_j$ [aldrei t$_j$]]

That is to say, the definite object ends up in a position preceding the sentence adverb not as a result of DP-movement but as a result of VP-movement. What we need to say to account for Object Shift of pronouns is that in many varieties of Scandinavian (but not all; cf. fn. 25) weak pronouns cannot move out of VP prior to VP-movement, and therefore obligatorily move along with the VP to a position preceding the sentence adverbs.

The most attractive aspect of this analysis, in the present context, is that it explains why Object Shift presupposes V-movement: It is because the object moves only as part of the VP. Furthermore, the 'PF-properties' of Object Shift discussed in section 5 [of the full article] are also predicted: Object Shift cannot license a parasitic gap simply because the object does not move, except as part of a VP. For the same reason it cannot affect binding relations. That it actually does affect binding under Principle C is unexpected, though.

Not surprisingly, there are many problems as well. Consider again the sentences in (1) and (2). A theory along Kaynean lines generates (1a) but not (1b, c), simply because there is no movement of weak pronominal objects in the theory. On the other hand, without further stipulations, it predicts wrongly that (65) is well formed.

(65) *Jag har kysst henne inte.
     I have kissed her not

In this case the VP moved to the pre-adverbal position is more complex, containing two verbs. This derivation can be excluded if the lower VP containing the infinitival verb plus the object is first moved out of the larger VP, after which the remnant VP, now containing only the finite verb, is preposed. The question is, of course, whether there is any independent motivation for this additional VP movement. A related problem is posed by V-Topicalization, discussed in section 3:

(66) Kysst har   jag henne inte.
      kissed have I   her   not

The lower VP movement just described will move the object pronoun out of the
larger VP, and thus block the derivation where the pronoun moves along with the
finite verb to pre-adverb position. That is to say, the problem is deriving (66)
without also deriving (65).

More problems are posed by (2b, c). The analysis of (2c) depends on which
theory of the particle construction is assumed, in particular in the case of a lan-
guage like Swedish. Since this is a controversial issue we may put (2c) aside.
Consider (2b), though, repeated here as (67):

(67) *Jag gav den inte Elsa.
       I   gave it not Elsa

The Kaynean theory here predicts that it should be well formed with the following
derivation: (a) The indirect object moves out of VP, (b) the remnant VP containing
the verb and the direct object move across the indirect object and the negation to
the higher VP-position. A stipulation to the effect that an indirect object can never
move out of VP in this way would of course prevent this derivation. However,
given that all other types of VP-constituents (PPs, CPs, small clauses, etc.) can
move out of VP in this theory, it seems ad hoc to deny only the bare (preposition-
less) indirect object this privilege.

The 'Kaynean' theory of Object Shift merits further study, though. After all,
no other theory around has an explanation of HG which is not based on a set of
more or less questionable assumptions, and furthermore they generally fail to
account for more than the 'core cases' of HG in (1). The theory elaborated in
previous sections has a broader empirical coverage, but this is achieved only with
the help of some assumptions, each of which may be called into question, having
to do with the nature of the postulated feature [±Foc] and the organization of the
grammar. Future research will hopefully tell whether these assumptions are well
founded or not.

Notes

1 For some recent views on Object Shift and Holmberg's Generalization, see
   Bobaljik & Jonas (1996), Collins & Thráinsson (1996), Ferguson (1996), Zwart
   (1994, 1997: 237ff.). The most detailed study of the properties of Scandinavian
   Object Shift is Holmberg & Platzack (1995: ch. 6). Cf. also Vikner (1994),
   Kaiser (1997). For an overview of the syntax of pronouns in Scandinavian, see
   Hellan & Platzack (1995). On the relation between shift of weak pronouns and
   shift of lexical DPs, see the text below.
2 Cf. Holmberg (1986) where this generalization was referred to as 'the phonetic
   adjacency condition' on Object Shift: The object has to be 'phonetically
   adjacent' to the adjuncts around which it is shifted.

27.3 Questions pertaining to Holmberg (1999)

1 To what extent is Holmberg's generalization a property of weak pronouns? Justify your answer.

2 Holmberg states that within Scandinavian the correlation between the richness of morphological Case and the robustness of object shift is weak. Relate this point of Holmberg's to discussions of the null subject/pro-drop parameter. (For recent discussion of null subjects in Germanic, see Rosenkvist 2009.)

3 Relate this point of Holmberg's to Pollock's (1989, Ch. 15 of this volume) discussion of verb-raising.

4 Scandinavian object shift does not move PPs, even when the object of the preposition is a pronoun. Discuss the extent to which this makes Scandinavian object shift similar to Romance clitic placement. (Extra credit: Extend the discussion to Semitic or Berber or Celtic.)

5 Along the same lines of question 4, when the object of the preposition is a pronoun, Scandinavian cannot move just the pronoun, stranding the preposition. To what extent might this be akin to English heavy–NP-shift (cf. Larson 1988, Ch. 13 of this volume; den Dikken 1995) not being able to apply to the object of a preposition? (Extra credit: Bring in other (non-)cases of P-stranding discussed by Kayne 1998.)

6 Scandinavian negative-phrase shift (cf. Christensen 1986) is freer than object shift; for example it is not subject to Holmberg's generalization, insofar as the negative phrase can cross a verb to its left. This recalls to some extent Kayne's (1981) discussion of the difference between French pronominal clitic movement and the movement in French of *tout* ('everything') and *rien* ('nothing'). Why might French and Scandinavian resemble each other in this respect?

7 Holmberg discusses the "possibility that Object Shift is a PF-operation," but goes on to note that "it is unclear what it means for an operation like Object Shift to apply in the phonological component," given that "it does not make reference to phonological primitives such as vowels, feet, nasality, etc." Find at least two of putative PF–operations in the recent syntactic literature to which Holmberg's skepticism would seem to apply equally well.

8 Holmberg notes that V(P)-topicalization involving a participial verb frees up object shift much more readily than does V(P)-topicalization involving an infinitival verb. This recalls the fact that pronominal clitic movement in French can escape a participial VP much more readily than an infinitival VP (cf. Kayne 1991). How might one try to fold this property of French into the derivation of "object shift+VP-topicalization" in Scandinavian? (Hint: Look at the next question.)

9 If the derivation of Scandinavian "object shift+VP-topicalization" involves a step in which the object pronoun moves to the left of the auxiliary, to what extent is that likely to link up to Holmberg's observation that (in embedded contexts) Scandinavian negative morphemes and floated subject quantifiers must precede the auxiliary?

10 Holmberg further notes that infinitival VP-topicalization, while not readily able to strand a prepositionless object pronoun, is however capable of

stranding a PP complement. Discuss whether or not this fact should be taken to be essentially the same as the fact that in English there is a sharp difference between the somewhat marginal *?. . .and allude he will to her* and the completely impossible *. . . and mention he will her*.

11 In his 1986 dissertation, Holmberg had noted that, in Swedish, DP-movement to subject position in passives is blocked by particles (of the *up, down, in, out* . . . sort). In this paper, he notes that object shift is blocked by particles. What considerations might bear on the question whether these two restrictions are or are not just one?

12 Andréasson (2009) presents evidence suggesting that demonstratives fail to undergo object shift in Swedish even when they are segmentally identical to pronouns. Why might that be? Bring to bear Cardinaletti and Starke (1999, Ch. 25 of this volume) and Leu (2007, 2008).

13 Like mainland Scandinavian languages in embedded contexts, English disallows an adverb from appearing between the verb and its (nonheavy) object (cf. Pollock 1989, Ch. 15 of this volume), as in *They should read often the news*. Yet English allows particles to intervene, as in *They should read out the news*. What, then, would Holmberg's analysis lead you to expect concerning the possibility of object shift in English? (Bring in Johnson 1991.)

14 Chomsky (2001: .26) suggests that object shift is available in all languages. How is his suggestion supported by Kayne's (2005, Ch. 11) discussion of the variety of English discovered by Kimball and Aissen (1971), in which sentences like *Which kids do John think should be invited to the party?* are possible?

15 Thráinsson (2007: 34) suggests that "V pronoun/DP negation," as produced by object shift in Scandinavian languages, is akin to "V pronoun/DP particle," as found in English sentences such as *They picked them/the books up*. Discuss the implications of Thráinsson's proposal from the perspective of Holmberg's section 9.

16 Although Icelandic readily allows object shift of lexical DPs, that is not true if the object DP is a determinerless indefinite. On the other hand, German allows sentences like *Ich glaube, dass er Bücher hätte lesen sollen* ('I think that he books had read should'='I think that he should have read books'), in which the determinerless indefinite object DP *Bücher* ('books') appears separated from its verb *lesen* ('to read') by an auxiliary. Why would German differ in this respect from Icelandic? (Hint: Bring in den Besten and Webelhuth 1990 and Koopman and Szabolcsi 2000.)

17 Chomsky (2001:33) suggests that the interpretation of an object-shifted phrase might be akin to that of a phrase moved to subject position (Spec, T). If so, then there is a certain tension between the restriction against object-shifted determinerless indefinites in Icelandic and the fact that English allows sentences such as *Books about linguistics can be found in the last aisle* and *Tabs should not be kept on them*. Discuss how a remnant movement approach to such English sentences might solve the problem, bringing in Chomsky (1995, Ch. 33). (Extra credit: What problem does subject–verb agreement raise? How might it be solved by recourse to Koopman's (2006) idea that agreement can hold between a head and the Spec of its Spec?)

References

Andréasson, M. 2009. Pronominal object shift – not just a matter of shifting or not. *Working Papers in Scandinavian Syntax* 84, Department of Scandinavian Languages, University of Lund, pp. 1–20.

den Besten, H., and G. Webelhuth. 1987. Remnant topicalization and the structure of VP in the Germanic SOV languages. *GLOW Newsletter* 18: 15–16.

den Besten, H., and G. Webelhuth. 1990. Stranding. In G. Grewendorf and W. Sternefeld (eds), *Scrambling and Barriers. Linguistik Aktuell/Linguistics Today* 5, Amsterdam and Philadelphia: John Benjamins, pp. 77–92.

Bobaljik, J., and D. Jonas, 1996. Subject positions and the role of TP. *Linguistic Inquiry* 27: 195–236.

Cardinaletti, A., and M. Starke. 1999. The typology of structural deficiency. In H. van Riemsdijk (ed.), *Clitics in the Languages of Europe*, Berlin: Mouton de Gruyter, pp. 145–233.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, pp. 1–52.

Chomsky, N. and H. Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8: 425–504.

Christensen, K. K.. 1986. Norwegian *ingen*: A case of postsyntactic lexicalization. In Ö. Dahl and A. Holmberg (eds), *Scandinavian Syntax*, Stockholm: Institute of Linguistics, University of Stockholm, pp. 21–35.

Collins, C., and Thráinsson, H. 1996. VP-internal structure and object shift in Icelandic. *Linguistic Inquiry* 27: 391–444.

den Dikken, M. 1995. Extraposition as intraposition, and the syntax of English tag questions. Ms., Vrije Universiteit Amsterdam / HIL.

Ferguson, K. 1996. Shortest move and object case checking. In W. Abraham, S. D. Epstein, H. Thráinsson, and J.-W. Zwart, *Minimal Ideas*, Amsterdam: John Benjamins, pp. 97–111.

Fox, D., and D. Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31: 1–46.

Hajder, H., S. Olsen, and S. Vikner. 1995. Introduction. In H. Haider, S. Olsen, and S. Vikner (eds), *Studies in Comparative Germanic Syntax*, Dordrecht: Kluwer, pp. 1–46.

Hellan, L., and C. Platzack. 1995. Pronouns in Scandinavian languages: an overview. *Working Papers in Scandinavian Syntax* 56, Department of Scandinavian Languages, University of Lund, pp. 47–69.

Holmberg, A. 1986. *Word order and syntactic features in the Scandinavian languages and English*. Doctoral dissertation, University of Stockholm.

Holmberg, A., and C. Platzack. 1995. *The Role of Inflection in Scandinavian Syntax*. Oxford and New York: Oxford University Press.

Johnson, K. 1991. Object positions. *Natural Language and Linguistic Theory* 9: 577–636.

Kaiser, L. 1997. The morphological cliticization of object-shifted weak pronouns in Swedish. Ms., Department of Linguistics, Yale University

Källgren, G., and E. Prince. 1989. Swedish VP-topicalization and Yiddish V-topicalization. *Nordic Journal of Linguistics* 12: 47–58.

Kayne, R. S. 1981. Binding, quantifiers, clitics, and control. In F. Heny (ed.), *Binding and Filtering*, London: Croom Helm, pp.191–211.

Kayne, R. S. 1991. Romance clitics, verb movement and PRO. *Linguistic Inquiry* 22: 647–686. [Reprinted in Kayne (2000, Ch. 5).]

Kayne, R. S. 1998. Overt vs. covert movement. *Syntax* 1: 128–191.

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kimball, J., and J. Aissen. 1971. I think, you think, he think. *Linguistic Inquiry* 2: 242–246.

Koopman, H. 2006. Agreement configurations: In defense of "Spec head". In C. Boeckx (ed.), *Agreement Systems*, Amsterdam: John Benjamins, pp. 159–199.

Koopman, H., and A. Szabolcsi. 2000. *Verbal Complexes*. Cambridge, MA: MIT Press.

Larson, R. 1988. On the double object construction. *Linguistic Inquiry* 19: 335–391.

Leu, T. 2007. These HERE Demonstratives. In T. Scheffler, J. Tauberer, A. Eilam, and L. Mayol (eds), *Penn Working Papers in Linguistics* 13. *Proceedings of the 30th Annual Penn Linguistics Colloquium*, Philadelphia: Penn Linguistics Club, pp. 141–154.

Leu, T. 2008. *The internal syntax of determiners*. Doctoral dissertation, New York University.

Nilsen, Ø. 1997. Adverbs and argument shift. *Working Papers in Scandinavian Syntax* 59, Department of Scandinavian Languages, University of Lund, pp. 1–31.

Nilsen, Ø. 2003. *Eliminating positions: the syntax and semantics of sentence modification*. Doctoral dissertation, Utrecht University.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Richards, M. D. 2004. *Object shift and scrambling in North and West Germanic: A case study in symmetrical syntax*. Doctoral dissertation, University of Cambridge.

Rochemont, M. 1978. *A theory of stylistic rules in English*. Doctoral dissertation, University of Massachusetts, Amherst.

Rosenkvist, H. 2009. Referential null subjects in Germanic languages – an overview. *Working Papers in Scandinavian Syntax* 84, Department of Scandinavian Languages, University of Lund, pp. 151–180.

Thráinsson, H. 2007. *The syntax of Icelandic*. Cambridge: Cambridge University Press.

Vikner, S. 1989. Object shift and double objects in Danish. *Working Papers in Scandinavian Syntax* 44, Department of Scandinavian Languages, University of Lund, pp. 141–155.

Vikner, S. 1994. Scandinavian object shift and West Germanic scrambling. In N. Corver and H. van Riemsdijk (eds), *Studies on Scrambling*, Berlin: Mouton de Gruyter, pp. 487–517.

Vikner, S. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford and New York: Oxford University Press.

Zwart, C. J.-W. 1994. On Holmberg's Generalization. In A. de Boer, H. de Hoop, and H. de Swart (eds), *Language and Cognition* 4. *Yearbook 1994 of the Research Group of Theoretical and Experimental Linguistics of the University of Groningen*, Groningen: Department of Linguistics, University of Groningen, pp. 229–242.

Zwart, C. J.-W. 1997. *Morphosyntax of Verb Movement: A Minimalist Approach to the Syntax of Dutch*, Dordrecht: Kluwer.

28

Movement and Control

Norbert Hornstein

1999

28.1 Introduction

One of the empirical findings brought to light by research in generative syntax is the distinction between two types of predicates that can embed infinitival complements: so-called *raising* and *control* predicates. Both can take an infinitival complement whose subject has no overt phonetic realization and is referentially dependent on the subject of the matrix clause:

(1) John seemed to like it. (*seem*: raising predicate)
(2) John expected to like it. (*expect*: control predicate)

However, these two apparently similar sentences contrast in a number of ways. One difference is that, in (1), *John* is an argument of the embedded predicate *like*, but not of the matrix predicate *seem*. This has led to an analysis that views the subject as raising from the embedded clause, where it is assigned a theta-role, to the matrix clause, where its need for case is satisfied. An analysis along these lines is compatible with the Theta Criterion, the principle that stipulates that every theta-role that a predicate bears must be assigned to an argument, and that an argument can bear exactly one theta-role (Chomsky 1981). In a sentence like (2), in contrast, *John* is interpreted as both an argument of the embedded predicate (*like*) and of the matrix predicate (*expect*). This has traditionally been taken to suggest that *John* does not move from the embedded to the matrix clause, because it would be assigned two theta-roles and violate the Theta Criterion. Among the most influential proposals put forth to account for control predicates, Chomsky (1981) proposed an analysis in which the infinitival subject is a special

unpronounced nominal element (PRO), whose distribution (and interpretation) is simultaneously subject to binding conditions A and B, the conditions on anaphors and pronouns. On this view, all the properties distinguishing control from raising predicates are seen as deriving (directly or indirectly) from the distinction between a movement and a binding relationship. Later, Chomsky and Lasnik (1993) suggested that the special distribution of PRO can be accounted for by stipulating a special kind of case ('null Case') that is only associated with certain kinds of inflectional heads.

In this provocative paper, *Movement and Control*, Hornstein departs from these and other views of control predicates. He argues instead that the referential dependency between the matrix and the embedded subject in obligatory control sentences should be analyzed in terms of movement:

(3)   John$_i$ expects [$_{TP}$ t$_i$ to like it]

Just as in the case of raising predicates, the subject starts out in the embedded clause and moves to the matrix clause, where its need for case is satisfied. However, as mentioned above, in the case of a control predicate the noun phrase plays the role of argument for both the embedded and the matrix predicate. To allow for this, Hornstein dispenses with the restriction imposed by the Theta Criterion that an argument can bear at most one theta-role. He argues instead in favor of a view of theta-roles as morphological features of predicates, which a noun phrase checks when it is merged with that predicate. In this view, a noun phrase is allowed to check more than one such feature, and thus to serve as an argument of more than one predicate. By viewing control predicates as involving movement of the subject from the embedded to the matrix clause, Hornstein can also dispense with the stipulation of the existence of a special unpronounced nominal element, PRO. In his view, the null subject of the embedded clause is simply a trace (or copy) of the overt noun phrase that has undergone movement. Moreover, Hornstein argues in this paper, he can derive the basic properties of obligatory control structures, while simplifying the inventory of grammatical primitives.

Hornstein's proposal (cf. Boeckx et al. 2010; Hornstein and Polinsky 2010) has received important criticism (cf. Landau 2000, 2001, 2003, 2004) and spurred a fruitful debate on the proper characterization of control structures. In addition to being at the center of that debate, it is also part of the literature that addresses the question of whether movement should be seen as the only mechanism by which grammar establishes relations at a distance. The empirical scope of this question includes, in addition to referential dependencies, agreement (Kayne 1989; Koopman 2006) and polarity licensing (Postal 2005). In work on clitic doubling, the literature has suggested the possibility that a noun phrase might start out as a complex element, a piece of which moves out (cf. Kayne 1972; Sportiche 1996; Uriagereka 1995). Kayne (2002) has explored a similar idea for the relation between a pronoun and its antecedent, adopting a movement approach that can extend to cases of control, maintaining some of the advantages of Hornstein's approach while being compatible with the Theta Criterion.

28.2   From "MOVEMENT AND CONTROL"

1   Introduction

This article is an exercise in grammatical downsizing. Since the earliest days of generative grammar (Rosenbaum 1967), control and raising constructions have been treated differently, with different rules and/or formatives involved in the two structures. In the beginning there was Equi-NP Deletion. Equi, a deletion process, contrasted with Subject-to-Subject Raising, a movement process. Subsequently, in most versions of the Extended Standard Theory, control was relegated to binding theory – the binding of an abstract expression PRO – whereas raising remained an instance of movement. This dual-track approach persisted into the Government–Binding (GB) era.

In GB, control sentences like (1a) have structures like (1b). These contrast with raising sentences, (2a), and their phrase markers, (2b). In particular, the relation between *John* and the embedded subject position in (1a) is mediated through the binding of a grammatically distinctive lexical formative in control configurations, namely, PRO. In raising structures like (2a) the relation between the matrix and embedded subjects is a by-product of movement and results in an A-chain in which the head, the antecedent, binds the tail, its trace.

(1)   a.   John expects to win.
      b. $John_i$ expects [PRO_i to win]
(2)   a.   John seemed to win.
      b. $John_i$ seemed [t_i to win]

The differences do not stop here. The distribution of PROs in GB is attributed to binding theory – the PRO Theorem, to be precise. The distribution of NP-traces, in contrast, is the province of the Empty Category Principle (ECP). Traces must be properly governed. PROs, on the other hand, cannot be governed at all. PROs head their own chains; traces, by definition, cannot. PROs are base-generated; traces are produced through movement. Thus, in most every respect, GB funda-mentally distinguishes NP-traces from PROs. Their one commonality within GB is that both are Caseless and phonetically null.

To date, standard work in the Minimalist Program has left matters pretty much in this GB state. There are good reasons for this. Empirically, the distinction reflects the fact that the antecedent of PRO in cases like (1a) bears two θ-roles whereas the subject in (2a) has but one. This semantic difference is theoretically ensconced in the different *kinds* of binding assumed to hold in control versus rais-ing. The theoretical basis for the distinction in GB technically rests on distin-guishing a level of D-Structure. D-Structure is the sole locus of lexical insertion, an operation that precedes all other transformations. Lexical insertion is subject to θ-requirements. In particular, D-Structure is defined as the phrase marker that purely represents GF-θ, the level at which all and only thematic positions of the sentence are occupied by lexical material. Subsequent transformations move the

lexical expressions located in θ-positions to non-θ-positions. These movements are further restricted by the θ-Criterion so that going from one θ-position to another is strictly forbidden.

This GB package of assumptions (the combination of D-Structure and the θ-Criterion) forces a distinction between PRO and trace, and thereby between binding and control. Two suppositions are central, and both are retained in the Minimalist Program: first, the θ-Criterion (the assumption that (A-)chains are constrained to possess but a single θ-position; i.e., movement from one θ-position to another is strictly forbidden); second, the priority of θ-marking over movement (i.e., the requirement that θ-positions coincide with the foot of a chain).

The first requirement prevents movement to θ-positions in the course of a derivation, just as it did in GB theories. The second retains a central feature of D-Structure. Chomsky ([1995a]) operationalizes the thematic restriction on lexical insertion by restricting θ-assignment to the merger of trivial chains. This recapitulates within the Minimalist Program the assumption that D-Structure is the locus of pure GF-θ. Thus, in Chomsky's ([1995a]) version of the Minimalist Program, a D/NP can legitimately enter a derivation only through the thematic door; that is, nominal expressions all enter the derivation via Merge. Given the provision that only trivial chains can be θ-marked, an NP so merged must merge to a θ-position on pain of never receiving a θ-role. Chomsky ([1995a]) further assumes that all subsequent movement is restricted to nonthematic targets. This is technically executed by assuming (3).

(3) a. θ-roles are not features.
    b. Movement must be greedy.

As θ-roles are not "checkable" features, movement to θ-positions cannot be greedy and so is prohibited. In short, the Minimalist Program retains the θ-Criterion.

All of this suggests that the minimalist abandonment of D-Structure as a level (Chomsky 1993) is less radical than often perceived. Chomsky's argument does not lead to a general repudiation of the core characteristics of D-Structure. Rather, D-Structure's earlier properties are packed into restrictions on the computational operations. In fact, the only feature of D-Structure that the Minimalist Program forswears is the principle that *all* lexical insertion precedes the application of *all* other transformations; in other words, the rule Satisfy has been dropped (Chomsky 1993). The other features of D-Structure have been retained.

This article submits these other assumptions to minimalist scrutiny. How well motivated are they? Why assume that chains are biuniquely related to θ-roles? What goes wrong if movement takes place from one θ-position to another? Why distinguish trace from PRO? As is generally the case with minimalist meditations, I assume that the burden of proof is on those who wish to promote these assumptions and invoke these distinctions. What is not at issue is that control and raising sentences manifest different properties. The minimalist question is whether these differences require the technical apparatus standardly invoked to distinguish them.

In the particular case of control, methodological skepticism is fully warranted. The distinction between raising and control multiplies the inventory of empty categories. Furthermore, the distinction massively complicates the grammar. PRO brings with it two big theoretical complications: (a) a control module whose job it is to specify how PRO is interpreted and (b) theoretical modifications to account for PRO's distribution. In GB (b) is handled by the binding theory. PRO is analyzed as a pronominal anaphor. The contradictory requirements that standard versions of the binding theory place on pronouns and anaphors within governing domains force such expressions to be ungoverned. Hence, PROs can appear only in ungoverned positions (see Chomsky 1986).

This GB approach to (b) has several conceptual and empirical problems (see Bouchard 1984, Chomsky and Lasnik 1993). Furthermore, the Minimalist Program cannot adopt the PRO Theorem, since it relies on the notion of government, which is not an acceptable minimalist primitive. Consequently, governing categories and domains cannot be defined or theoretically exploited.

Chomsky and Lasnik (1993) propose that the distribution of PRO is regulated by Case theory. Theirs is the standard minimalist account for the distribution of PRO. In particular, they propose that PRO has "null" Case. This is a Case special to PRO in the sense that only PRO bears it and Is that assign/check it license no other sorts of Case. It is fair to say that null Case accounts for the distribution of PRO largely by stipulation.

The theory of the control module does not fare much better. What principles determine the antecedents of PRO, and whether or not all instances of control are actually the same, is quite controversial. It seems safe to say that control theory has not been one of the bright stars in the GB firmament.

In sum, neither part of the control conglomerate has been uncontroversial even within GB. Given a minimalist sensibility, its technical complexities are ripe for reevaluation.

The article is organized as follows. Section 2 reviews why we need a theory of PRO and control. In particular, sections 2 and 3 review the distribution and interpretive requirements of PRO in obligatory control (OC) and nonobligatory control (NOC) configurations. Sections 4 and 5 argue that the general properties of OC structures can be reduced to movement if we abandon the residues of D-Structure still extant within the Minimalist Program and abandon the θ-Criterion-based prohibition against moving into θ-positions. Section 6 briefly addresses how NOC and OC structures are selected. Section 7 concludes the exercise.

[. . .]

Finally, given minimalist inclinations, the deepest question concerning PRO is whether such a formative even exists. PRO is a theory-internal construct. In GB, PRO is structurally analogous to NP-traces and *wh*-traces. All have the same shape, namely, $[_{NP}\ e]$. The main difference between traces and PRO is the source of their indices: the former derive from movement, the latter are assigned via the control module. In the Minimalist Program, however, this machinery is all suspect. There is little reason to think that traces (qua distinctive grammatical constructs) exist at all. Traces are not grammatical formatives but the residues of the copy-and-deletion operations necessary to yield PF/LF pairs. As such, traces have no common structure

in the Minimalist Program as they do in GB. They are simply copies of lexical material and so have no specific shapes whatsoever. Thus, they cannot be structurally analogous to PRO. This leaves the theoretical status of PRO up in the air. What kind of empty category is it? Why do grammars have it?

Section 1 has provided answers to these questions. PRO exists because of θ-theory. If chains could bear more than one θ-role and if θ-roles could be accreted in the course of a derivation, there would be little reason to distinguish PROs in OC configurations from NP-traces. As these restrictions on θ-assignment are not conceptually necessary, the theoretical basis for distinguishing PROs from NP-traces weakens. Put more bluntly, distinguishing trace from PRO requires *additional* assumptions about θ-assignment and chains. The burden of proof, therefore, resides with those who favor such assumptions. In section 4 I argue that forgoing these stipulations permits a more empirically and theoretically adequate account of OC. I propose that PRO, like NP-trace, is the residue of movement. Strictly speaking, then, there is no grammatical formative like PRO. Rather, PRO is simply a residue of movement – simply the product of copy-and-deletion operations that relate two θ-positions.

4   An alternative

I have argued that the null hypothesis is that OC PRO is identical to NP-trace; that is, it is simply the residue of movement. NOC PRO is to be identified with pro, the null pronominal found in various Romance and East Asian languages. This section is concerned with demonstrating the empirical virtues of these assumptions. The main focus is on OC PRO, since handling the OC data requires the most radical departures from standard GB and minimalist technicalia. For what follows, I adopt the following assumptions:

(18)   a.   θ-roles are features on verbs.
         b.   Greed is Enlightened Self-Interest.
         c.   A D/NP "receives" a θ-role by checking a θ-feature of a verbal/predicative phrase that it merges with.
         d.   There is no upper bound on the number of θ-roles a chain can have.
         e.   Sideward movement is permitted.

(18a) treats θ-roles as morphological features. This is required if movement to a θ-position is to conform to the principle of Greed. If OC is to be reduced to movement, then this assumption is conceptually required given other minimalist assumptions. (18b) interprets Greed as requiring at least one of the relata to check a feature (Lasnik [1995]). Thus, if A moves to merge with B, then at least one feature of either A or (the head of) B is checked. Treating θ-roles as features on the verb or predicate allows a D/NP to move to a θ-position and respect Greed by checking this feature. Analyzing θ-roles thus permits us to "mechanize" θ-role assignment as in (18c): to receive a θ-role is just to check the relevant thematic feature of the predicate. One might think of this as "transferring" the verbal θ-feature to the nominal expression. In effect, checking conforms to Chomsky's

([1995a]:226) vision of syntactic operations as the "rearrangements of properties of the lexical items of which they are ultimately constituted" – that is, the features of the elements in the array. (18d) is logically required to analyze OC in terms of movement given that control involves the relation of at least two θ-positions. It is also the null hypothesis, I believe. The requirement that chains be restricted to a single θ-role needs substantial empirical justification. (18e) comes into play in the analysis of adjunct OC. I discuss it further in that context. What is important here is that c-command is not part of the *definition* of movement. Thus, the computational system does not prohibit the copying of an expression to a position that does not c-command the "movement" site.

The assumptions in (18) suffice to accommodate OC in terms of movement given standard minimalist technology. Their empirical virtue is that they permit a radical simplification of the grammar of control and a derivation of the basic properties of OC structures. Consider the details.

First, consider the basic interpretive properties of OC structures. As noted in section 2, these structures require c-commanding local antecedents (see (4a–c)). This is what one expects if OC PROs are NP-traces. For illustration, let us look at (19).

(19)   a.   John hopes to leave.
       b. $[_{IP}$ John $[_{VP}$ John [hopes $[_{IP}$ John to $[_{VP}$ John leave]]]]]

The derivation begins with *John* merging with *leave*, thereby checking the verb's θ-role. *John* then "raises" to the embedded [Spec, IP] to check the D-feature of the IP. This is *not* a Case-marking position, so the Case of *John* cannot be checked here. *John* raises again to [Spec, VP] of *hope* and checks the external θ-feature of the verb. By (18c), each time *John* checks a θ-feature of a predicate, it assumes that θ-role. Thus, *John* (or the chain it heads) has two θ-roles, the leaver role and the hoper role. *John* raises one last time to [Spec, IP] of the matrix, where it checks the D-feature of the IP and nominative Case. Note that this is the only place where *John* checks Case. On the assumption that it was inserted into the derivation with nominative Case features, the derivation converges. In more conventional notation, the copy *John* in the embedded [Spec, IP] corresponds to PRO, and the copy in the matrix [Spec, IP] is the antecedent. The requirement that OC have a local c-commanding antecedent follows from the fact that PRO is an intermediate link in an A-chain. As such, it must have an antecedent. Furthermore, the antecedent must conform to general A-chain strictures and thus both c-command the traces in the A-chain (i.e., the PRO in [Spec, IP]) and be local to it, given conditions on movement like the MLC. In short, the first three properties of OC PRO follow straightforwardly (see (4a–c)).

Treating OC PRO as the residue of movement also derives the prohibition against split antecedents. Two (nonconjoined) expressions cannot both antecede OC PRO because they cannot have both moved from the same position. In other words, the ban against split antecedents in this case is equivalent to the ban against one and the same trace having two distinct antecedents. In the Minimalist Program this reduces to the fact that two distinct expressions cannot be merged into a single position.

The required sloppy reading of OC PRO follows as well. Note that in raising constructions only a sloppy reading is available.

(20)   Mary seems to be happy and Sally does too.

(20) must be understood to mean that it seems that Sally is happy. For the same reason, OC PRO must carry the sloppy reading since it too is an NP-trace.

The movement analysis also accounts for the required *de se* interpretation of OC PRO. The movement underlying OC PRO ends up assigning two θ-roles to a single expression; for example, in (19a) *John* has two θ-roles. The semantic form of the predication in (19) is equivalent to (21), a predication that ascribes a reflexive property to the subject *John*.

(21)   John λ*x* [*x* hopes *x* leave]

Movement, then, semantically forms a compound monadic predicate by having one and the same expression saturate two argument positions. Salmon (1986) discusses these semantic issues at some length. Of importance here is his observation that relating the semantic value of an expression to two θ-positions via the formation of a reflexive predicate is semantically very different from relating two expressions in different θ-positions to each other via coreference. The former operation results in changing the semantic argument structure of the predicate; the latter leaves it intact. The former operation reflexivizes the predicate and thus forces a *de se* reading; the latter does not. Treating OC as the reflex of movement, then, yields the correct interpretation for the structures – the one exemplified in (21).

Finally, the observed reading in (4g) (repeated here) follows as well.

(22)   Only Churchill remembers giving the BST speech.

The reading on which someone other than Churchill could recall this event requires the paraphrase in (23).

(23)   Only Churchill remembers Churchill giving the BST speech.

This cannot underlie the structure of (22). The PRO here is of the OC variety. This means that *only Churchill* has raised from the embedded position and has the reflexive property noted in (24). This is semantically equivalent to the reading on which Churchill alone has the required memory.

(24)   only Churchill λ*x* [*x* remembers *x* giving the BST speech]

In sum, the six basic properties of OC reviewed in section 2 [of the full article] follow directly from assuming that OC PRO is identical to NP-trace, the residue of movement. In addition, these properties are derived without the problems reviewed in section 3. Once again, consider the details.

The distribution of OC PRO does not require the services of null Case. This Case, specially designed for PRO by Chomsky and Lasnik (1993), is unnecessary if OC PRO is an NP-trace. In fact, the existence of null Case in [Spec, IP] of control infinitives is *incompatible* with the movement analysis, since it would prevent raising out of the embedded [Spec, IP].

Abandoning null Case in this context does not lead to any empirical difficulties. Recall that null Case has been postulated to replace the assumption that PRO must be ungoverned. Its principal empirical effect is to block the derivation of (25) and license PRO only in [Spec, IP] of nonfinite clauses.

(25)   *We never expected [PRO$_i$ to appear to t$_i$ that . . .

The proposed account rules (25) out on the same basis as an account that postulates null Case. On the latter view, PRO cannot move to [Spec, IP] of the embedded clause because it would be moving from one Case-marking position (inside PP) to another ([Spec, IP]). This either violates Greed or causes a feature mismatch. In either case the derivation fails to converge. However, if [Spec, IP] is an *intermediate* NP-trace, as it would be on the proposed account, then the same reasoning prohibits movement through this position. In effect, a PRO in (25) should be no better than an NP-trace in (26).

(26)   *We$_i$ were expected [t$_i$ to appear to t$_i$ that . . .

Furthermore, on the proposed account we expect to find OC PRO in positions from which movement is licit. This should roughly coincide with non–Case-marked positions, such as [Spec, IP] of nonfinite clauses. Note that this is compatible with treating inherently reflexive verbs like *wash*, *dress*, and *shave* as simply not Case-marking their objects – in effect, as allowing derivations like (27) to be licit.

(27)   a.   Mary washed.
        b.   [$_{IP}$ Mary [past [$_{VP}$ Mary [wash Mary]]]]

Case is checked in [Spec, IP]. *Mary* receives two θ-roles since it checks both the internal and external θ-role of *wash*.

In effect, then, by assuming that PRO is identical to an NP-trace, an intermediate NP-trace to be exact, we derive its distribution without having to assume null Case. This account has two further benefits. First, it allows us to treat *wanna* contraction over PRO and NP-trace as one and the same phenomenon (see (10) and (11)). Second, the null phonetic status of PRO is explained in whatever way we explain the null phonetic status of NP-trace. One natural assumption is that Case is required for phonetic "visibility." Both NP-trace and PRO will therefore fail to meet the requirements for having phonetic content.

The movement approach to OC PRO also accommodates the classical data used to distinguish raising from control. It was argued, for example, that idiom chunks and expletives could raise but not control.

(28) a. The shit seems [t to have hit the fan].
    b. There seems [t to be a man in the garden].
(29) a. *The shit expects [PRO to hit the fan].
    b. *There expects [PRO to be a man in the garden].

The distinction between these cases is preserved in the present account even if PRO in (29) is just an NP-trace. The basis for the distinction is that in (28) *the shit* and *there* bear the external θ-role of *expect*. If this θ-role is not checked, then, I assume, the derivation fails to converge since there is an unchecked θ-feature at LF. However, the only nominals that can check the relevant θ-roles, *there* and *the shit*, are not expressions that can bear θ-roles because of their inherent idiomatic or expletive semantics. As a result, we retain a difference between raising and control structures in cases such as these but attribute it not to an inability to control PRO but to an inability to support a θ-role that must be discharged for grammaticality to ensue.

    This section has demonstrated that OC structures can be treated in terms of movement and that there is considerable empirical payoff in doing so. In particular, we can dispense with null Case, and we can derive the six basic properties of OC exemplified in (4). The next section turns to perhaps the biggest advantage. It appears that treating OC PRO as the residue of movement comes very close to allowing us to eliminate the PRO module entirely.

    To sum up: OC PRO is the residue of movement and has all the characteristics of NP-trace. The only real distinction between raising and control structures is that the former involve raising a D/NP to a non-θ position whereas the latter raises expressions to θ-positions. Both raising and control chains (generally) terminate in Case positions.

28.3 Questions pertaining to Hornstein (1999)

1   To judge by auxiliary selection (cf. Perlmutter 1989; Burzio 1986), i.e., by its taking auxiliary *be*, the French or Italian verb corresponding to English *intervene* must be unaccusative. Yet it simultaneously seems to be agentive, as in *John purposely intervened in the dispute*. Might this fact justify movement of the argument (*John*, in the previous example) from an object theta-position within VP into a higher theta-position Spec,vP? If so, to what extent would (or would not) that resemble Hornstein's proposed derivation for control structures?

2   Hornstein argues for allowing movement from one theta-position into another. How is that similar to or different from Kayne's (2002) proposal concerning pronouns and their antecedents?

3   If Hornstein's proposal to reduce control to movement turns out to be incorrect (as Landau 2003, 2006 argues that it is), we will need to ask why exactly the language faculty fails to treat control as movement in Hornstein's way. Is it that movement from one theta-position to another is barred in general? If so, to what extent might that be related to a prohibition against movement from one Case position to another?

4   The Italian counterpart of *seem* allows raising, much as in English, but also control by the dative, as if English were to allow *It seems to me to have bought myself the wrong book*. As discussed by Kayne (1981), there is a difference in Italian between the raising case, in which *sembrare* ('seem') takes an infinitive not preceded by any complementizer (as in *Gianni sembra aver capito* ('G seems to-have understood')), and the control case, in which the infinitive is preceded by complementizer *di* (as in *Mi sembra di aver capito* ('me(dative) seems *di* to-have understood')). How might Hornstein try to cope with these facts?

5   The Italian facts mentioned in the preceding question could be taken to suggest that control in all languages invariably involves a full CP complement (as opposed to raising, where the complement is IP). If so, then Hornstein's movement approach to control will in all cases require movement out of the subject position of a CP. To what extent would such movement violate Chomsky's (1981) ECP, or Chomsky's (2001) PIC, or Rizzi's (2006) criterial freezing?

6   Control can be found within derived nominals, in cases like *John's attempt to solve the problem*. How might Hornstein integrate these into his proposal? Take one restriction on derived nominal control from Pesetsky (1991) and discuss how Hornstein or Landau (2006) might try to account for it.

7   Discuss the similarities and differences between Hornstein's proposal and that made by Cinque (2006) for Italian *volere* ('to want').

8   Hornstein claims that his proposals "permit a radical simplification of the grammar of control." To what extent is it legitimate to focus on the question of simplification of one subpart of the theory of syntax, as opposed to the theory as a whole? To what extent is Landau (2003) correct to deny Hornstein's claim about simplification?

9   How might Hornstein account for the incompatibility of *for* with control, in standard English examples like *John would like very much (*for) to win the race*? How might he try to extend his account to the different facts of Belfast English discussed by Henry (1995)?

10   English disallows control with *believe* (as opposed to French and Italian), e.g., **Mary believes to be right*. Again as opposed (on the whole – see Pollock (1985) for nuances) to French and Italian, English does have ECM with *believe*, as in *Mary believes Susan to be right*. To what extent does this apparent complementarity fit with Hornstein's approach to control? Show how bringing in the question of English gerunds enriches the entire question.

11   Hornstein takes the interpretation of *Only Churchill remembers giving the BST speech* to indicate that the unpronounced subject of *giving* is of the OC type. To what extent is that compatible with the acceptability of *Only John remembers Churchill giving the BST speech*?

12   At the end of his paper, Hornstein suggests that reflexives are also a residue of movement. How does this suggestion of his interact with the morphological complexity of English reflexives and with Chomsky's notion of inclusiveness? Bring into the discussion Helke's (1971, 1973) point about the similarity of English reflexive sentences to others like *John has lost his cool*.

13   The absence of a "strict" reading with control and ellipsis plays a role for Hornstein. How important is it that for at least some English speakers, the following are only somewhat marginal, rather than impossible, with the "strict" reading?: *?John remembers having gotten ridiculously drunk at the party last week, and his wife remembers that perfectly well, too, ?John remembers having gotten ridiculously drunk at the party last week, as does his wife, unfortunately for him, ?John would like to remain calm and his wife would like that, too*. (In the strict reading, this last sentence, for example, is interpreted as ". . . and his wife would like him to remain calm, too.")

14   English is not a null subject language, yet it allows imperatives like *Do that right away!*, with a null subject. An attractive hypothesis, thinking of Ross (1970), is that imperatives are exceptional in English by virtue of their being instances of control, with a silent matrix verb akin to *tell*. Find at least two arguments against taking imperatives to be instances of control and evaluate their weight.

15   Hornstein touches on the fact that control is not possible with expletive *there*, as shown by *There can't possibly be a solution without *(there) being a problem* and **There was a problem long before being a solution*. How might the analysis of expletive *there* in Kayne (2008) provide an account?

16   (Extra credit) English shows a sharp distinction between *whether* and *if* in *They don't know whether/*if to leave right away or not*. Some Romance languages act like English, while others do not. A proposal involving verb movement was made in Kayne (1991, Ch. 18 of this volume). To what extent could Hornstein incorporate some version of that proposal into his?

17   Hornstein notes an intriguing similarity between ordinary control and sentences like *John shaved early this morning*. This latter type of sentence is also sometimes possible in English with a reciprocal, rather than reflexive, reading, e.g., *They met/kissed early this morning*. Ordinary control, though, seems never to allow a reciprocal reading, i.e., *They want to be elected* can't be interpreted as *They want each other to be elected*, despite the existence of what Landau (2000) calls "partial control." Why might that be? To what extent are (or are not) these facts about reciprocals damaging to Hornstein's proposed common analysis of *John shaved* and ordinary control? Bring into the discussion the fact that German apparently reflexive *sich* can yield a recip-rocal interpretation as a direct object, but not as a prepositional object.

18   The facts about Rosenbaum's (1967) Minimal Distance Principle discussed by Hornstein seem to be mimicked in *I asked him if he/*I would please help them* (cf. Postal 1970: 468–476, 488; Jenkins 1972, Ch. 4). How might Hornstein try to account for this?

19   Control, as opposed to raising, is not possible with argumental small clauses. Thus while the raising example *John seems intelligent* is fine, there is a sharp contrast in *John claims *(to be) intelligent*. Compare Hornstein's approach to control with Landau (2006) and with Chomsky (1981) with respect to this fact.

20   Kayne (2006) argues that silent elements must have moved to a special posi-tion. Discuss the similarities and differences between Kayne's proposal and Hornstein's claim that control involves movement.

References

Boeckx, C., N. Hornstein, and J. Nunes. 2010. *Control as Movement*. Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.

Bouchard, D. 1984. *On the Content of Empty Categories*. Dordrecht: Foris.

Burzio, L. 1986. *Italian Syntax: A Government-Binding Approach*. Dordrecht: Reidel.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.

Chomsky, N. 1993. A minimalist program for linguistic theory. In K. Hale and S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press, pp. 1–52. [Reprinted in Chomsky (1995b, Ch. 3).]

Chomsky, N. 1995a. Categories and transformations. In Chomsky (1995b, Ch. 4).

Chomsky, N. 1995b. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, pp. 1–52.

Chomsky, N., and H. Lasnik. 1993. The theory of principles and parameters. In J. Jacobs, A. von Stechow, W. Sternefeld, and T. Vennemann (eds), *Syntax: An International Handbook of Contemporary Research*, Berlin: Mouton de Gruyter, pp. 506–569. [Reprinted in Chomsky (1995b, Ch. 1).]

Cinque, G. 2006. Restructuring and Functional Heads. The Cartography of Syntactic Structures 4. New York and Oxford: Oxford University Press.

Helke, M. 1971. *The grammar of English reflexives*. Doctoral dissertation, MIT.

Helke, M. 1973. On reflexives in English. *Linguistics* 106: 5–23.

Henry, A. 1995. *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford Studies in Comparative Syntax. New York and Oxford: Oxford University Press.

Hornstein, N., and M. Polinsky (eds). 2010. *Movement Theory of Control*. Amsterdam and Philadelphia: John Benjamins.

Jenkins, L. 1972. *Modality in English syntax*. Doctoral dissertation, MIT.

Kayne, R. S. 1972. Subject inversion in French interrogatives. In J. Casagrande and B. Saciuk (eds), *Generative Studies in Romance Languages*, Rowley, MA: Newbury House, pp. 70–126.

Kayne, R. S. 1981. On certain differences between French and English. *Linguistic Inquiry* 12: 349–371.

Kayne, R. S. 1989. Facets of Romance past participle agreement. In P. Benincà (ed.), *Dialect Variation and the Theory of Grammar*, Dordrecht: Foris, pp. 85–103. [Reprinted in Kayne (2000, Ch. 3).]

Kayne, R. S. 1991. Romance clitics, verb movement and PRO. *Linguistic Inquiry* 22: 647–686. [Reprinted in Kayne (2000, Ch. 5).]

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2002. Pronouns and their antecedents. In S. D. Epstein and T. D. Seely (eds), *Derivation and Explanation in the Minimalist Program*, Oxford: Blackwell, pp. 133–166. [Reprinted in Kayne (2005, Ch. 6).]

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2006. On parameters and on principles of pronunciation. In H. Broekhuis, N. Corver, R. Huibregts, U. Kleinhenz, and J. Koster (eds), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, Berlin and New York: Mouton de Gruyter, pp. 289–299. [Reprinted in Kayne (2010, Ch. 4).]

Kayne, R. S. 2008. Expletives, datives, and the tension between morphology and syntax. In T. Biberauer (ed.), *The Limits of Syntactic Variation*, Amsterdam and Philadelphia: John Benjamins, pp. 175–217. [Reprinted in Kayne (2010, Ch. 6).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Koopman, H. 2006. Agreement configurations: In defense of "Spec head". In C. Boeckx (ed.), *Agreement Systems*, Amsterdam: John Benjamins. pp. 159–199.

Landau, I. 2000. *Elements of Control: Structure and Meaning in Infinitival Constructions*. Studies in Natural Language and Linguistic Theory. Dordrecht: Kluwer.

Landau, I. 2001. Control and extraposition: The case of Super-Equi. *Natural Language and Linguistic Theory* 19: 109–152.

Landau, I. 2003. Movement out of control. *Linguistic Inquiry* 34: 471–498.

Landau, I. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22: 811–877.

Landau, I. 2006. Severing the distribution of PRO from case. *Syntax* 9: 153–170.

Lasnik, H. 1995. Last Resort and Attract F. In *Proceedings of FLSM* 6, Bloomington: Indiana University Linguistics Club, pp. 62–81.

Perlmutter, D. 1989. Multiattachment and the unaccusative hypothesis: The perfect auxiliary in Italian. *Probus* 1: 63–119.

Pesetsky, D. 1991. Zero syntax: vol. 2: Infinitives. Ms., MIT. Available at http://web.mit.edu/linguistics/people/faculty/pesetsky/infins.pdf [accessed January 2013].

Pollock, J.-Y. 1985. On case and the syntax of infinitives in French. In J. Guéron, H.-G. Obenauer, and J.-Y. Pollock (eds), *Grammatical Representation*, Dordrecht: Foris, pp. 293–326.

Postal, P. 1970. On coreferential complement subject deletion. *Linguistic Inquiry* 1: 439–500.

Postal, P. 2005. Suppose (if only for an hour) that Negative Polarity Items are negation-containing phrases. Available at http://ling.auf.net/lingBuzz/001603 [accessed January 2013].

Rizzi, L. 2006. On the form of chains: Criterial freezing and ECP effects. In L. Cheng and N. Corver (eds), *Wh-movement: Moving On*, Cambridge, MA: MIT Press, pp. 97–134.

Rosenbaum, P. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: MIT Press.

Ross, J. R. 1970. On declarative sentences. In R. Jacobs and P. Rosenbaum (eds), *Readings in English Transformational Grammar*, Waltham, MA: Ginn, pp. 222–272.

Salmon, N. 1986. Reflexivity. *Notre Dame Journal of Formal Logic* 27: 401–429.
Sportiche, D. 1996. Clitic constructions. In J. Rooryck and L. Zaring (eds), *Phrase Structure and the Lexicon*, Dordrecht: Kluwer, pp. 213–276.
Uriagereka, J. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26: 79–124.

29

VSO and VOS: Aspects of Niuean Word Order

Diane Massam

2000

29.1   Introduction

Like the article by Anderson and Chung included in Ch. 5 of this volume, Massam's contribution also takes VSO languages as its empirical domain. Whereas the former discusses patterns from Samoan, Tongan, and Breton, this article focuses on Niuean, a Malayo-Polynesian language. Massam's paper builds on the seminal idea contained in Anderson and Chung's (1977) work, that, even in VSO languages, the object holds a special relationship with the verb, and hence dismisses the possibility that Niuean might have a flat structure. As shown in Seiter (1980), the subject c-commands the object in this language. Moreover, Massam provides evidence that objects are structurally closer to the verb than subjects, despite the VSO word order: (i) an object can incorporate into the verb, whereas a subject cannot; (ii) the verb and its object can form an idiom, whereas a verb and its subject don't seem to be able to. Hence, this contribution further develops the idea that VSO languages are like SVO and SOV languages in building their clausal structure by first combining the verb and its object to form a VP constituent.

A second aspect of the paper is the account of the derivation of VSO order. Here Massam builds on the notion of *Remnant Movement* proposed in den Besten and Webelhuth (1990), that is to say, movement of an XP that contains a gap/trace. Massam notes that so-called "noun incorporation" in Niuean is allowed with complex phrases as in (1a), where the counterpart of *good fish and chips* appears non-Case-marked and is fronted along with the verb. This contrasts with (1b), where the object is Case-marked, receives a definite interpretation, and remains in post-subject position:

(1) a. Ne [kai sipi mo e   ika mitaki] a   Sione
        PAST eat chip COM ABS fish good   ABS Sione
        'Sione ate good fish and chips.'
    b. Ne [inu] e   Sione e   kofe
        PAST drink ERG Sione ABS coffee
        'Sione drank the coffee.'

Noting that the object in (1a) is clearly not a head and thus its position cannot be derived via noun incorporation, Massam proposes that in Niuean the entire VP (more generally, the entire predicate) raises across the subject, and not simply the verb. In other words, verb-initial order is derived by VP-fronting (cf. also Rackowski and Travis 2000; Lee 2000 on Zapotec; Pearson 2001 on Malagasy; Coon 2010 on Mayan; Nilsen 2003 and Müller 2004 on Germanic V2). The contrast between (1a) and (1b) is due, according to Massam, to a difference between NP and DP objects: bare NP objects remain in their base-position, complement of V, and are thus moved along with the VP (1a). DP objects, in contrast, move out of the VP to check absolutive Case (in the specifier of an AbsP), and thus are not fronted along with the VP (1b).

Concerning the landing site of VP-fronting, the fact that tense (and negation) precede the fronted verb leads Massam to propose that the VP moves to Spec,IP, and that tense is expressed in C. The subject, in contrast, stays in a low position. The idea is that languages can vary parametrically as to whether the EPP requirement of Infl (that is, its need to have something in its specifier) is triggered by a [D] feature or a [Pred] feature. If the former, a DP will need to move to the specifier of Infl, as in an SVO language like English; if the latter, the predicate will have to move there, as in a VSO language like Niuean.

29.2   From "VSO AND VOS: ASPECTS OF NIUEAN WORD ORDER"

This chapter explores the verb initial theme through an examination of predicate fronting in the Polynesian language Niuean. Beginning with the V-fronting analysis advanced by Emonds (1980) and Sproat (1983), among others, two questions are addressed: Where does the Niuean fronted element land? And what sort of constituent undergoes fronting? In answering the first question in section 1, it is shown that the Niuean verb does not front to C or T, since another head, namely NEG, can appear between the complementizer/tense morpheme and the verb. This necessitates positing an additional functional head between C and VP, which, following convention, I label Infl. Section 1 serves largely to describe Niuean fronting so as to place it in a typology of VSO languages. In sections 2 and 3, related issues are raised which have broader cross-linguistic and theoretical implications. In section 2, a brief examination of non–verbal predicate sentences, and of noun incorporation structures, leads to the conclusion that Niuean is not a verb-fronting language, but rather a predicate-fronting language. This is extended so that in all sentence types it is a maximal projection predicate that is fronted to specifier of IP. Thus, VSO order is in fact $[V t_O]SO$ order. In concluding

section 2, some empirical implications are examined regarding Niuean postverbal morphemes. In section 3, it is proposed that predicate fronting takes place, analogously to subject fronting in SVO languages (cf. Massam and Smallwood 1997), to check an Extended Projection Principle-related feature [Pred]. Thus, Niuean predicate fronting to IP, resulting in the so-called VSO order, is an externalization process, and is fundamentally different from the verb fronting to IP witnessed in SVO languages such as French (cf. Pollock 1989). Since it is claimed that predicates are syntactically externalized, it is necessary to define the notion of EPP predicate in a uniform fashion. An EPP predicate is defined as an X° lexically determined predicate and its internal argument. The conclusion, section 4, reviews the findings of the chapter.

1   Niuean fronting is to IP, not CP

This section begins with a basic description of Niuean sentence structure, and an examination of word order. First, arguments are provided against a flat structure; then arguments are provided that Niuean Tense is in C, and that VSO order derives from fronting to Infl and not to C (cf. McCloskey 1991, [1996], for an overview of these two options).

Niuean exhibits a V-S-O-IO-Obl word order, as shown in (1) below.

(1) a. Ne   tala aga   e ia   e   tala   ke he   tagata.
         Past   tell Dir3p ErGhe Abs story to prt   man
         'He told the story to the man.'
     b. Hifo   a   Lemani ki tahi mo   e vaka.
         go-down   Abs Lemani to sea with Abs canoe
         'Lemani went down to sea with a canoe.'

If we adopt the common assumption that the verb and its object originate within V′, then all Niuean clauses with a verbal predicate involve some kind of movement. It could be taken simply as a theoretical assumption that the verb and object originate in V′, but there is in fact some evidence for this, in that verbs form a closer D-structure bond with their objects than with their subjects. For instance, verbs can incorporate internal, but not external arguments, and verbs can form idioms in conjunction with their internal arguments (but not, it seems, with their external arguments). Woolford's (1991) analysis of Niuean as having a flat VSO structure is thus not adopted.[5] The incorporation facts are shown below in (2), while (3) shows a verb-object idiom.

(2) a. Ne   inu   e   Sione e   kofe.
         Past drank Erg Sione Abs   coffee
         'Sione drank the coffee.'
     b. Ne   inu   kofe a   Sione.
         Past drink coffee Abs   Sione
         'Sione drank coffee.'

    c.   Fā   totou   he   tau faiaoga   e   tau tohi.
        HAB read   ERG PL teacher ABS PL book
        '(The) teachers often read books.'
    d. *Fā   totou   faiaoga   e   tau tohi.
         HAB   read   teacher ABS PL   book
        *'Teachers often read books.'

(3)   Loto   a   au   ke   oeli e tau matahui, ti   koli.
      like ABS I   SBJN oil ABS PL knee   then dance
      'I like to get a little drunk, then dance.'

Given a verb/object constituent at the level of thematic representation, I assume that the Niuean verb in sentences such as those above undergoes leftward movement to some c-commanding head position. Following Chomsky (1995), it is assumed that the features of a functional head may contain a strong uninterpretable feature which requires checking prior to spellout, thus attracting some other element, which, as a result of Attract, immediately undergoes a Move operation. A central problem in the study of VSO languages is to determine to which functional head the verb is attracted. The debate usually focuses on C and Infl (or Tense) (cf. Hendrick [2000] and Carnie, Harley, and Pyatt [2000]), so I begin the discussion with these categories.

In Niuean, the two categories of C and Tense are morphologically merged. The sentence begins with a particle which indicates the tense/aspect of the sentence. These particles are given in (4). Note that the particles can be null, as in (1b) and (3).

(4)   *Sentence-initial tense/aspect particles*

PAST	FUTURE	PROGRESSIVE	PERFECT	SUBJUNCTIVE
ne/na	to	hā ne	kua	kia

    a.   *Ne*   tagi   a   ia.
        PAST cry   ABS she
        'She cried.'
    b.   *To* fano   a   ia.
        FUT go   ABS he
        'He will go.'
    c.   *Hā ne* nonofo a   mutolu i hinei.
        PROG stay   ABS you   at this.place
        'Whilst you are staying here.'
    e.   *Kua* fano tuai a   ia.
        PERF go   PERF ABS he
        'He has gone.'

The sentence initial particle expresses the tense or aspect of the sentence. These particles, however, also display complementizer-like properties, as discussed by Seiter (1980), in that the particles vary depending on the complementation status

of the sentence, as follows. Matrix clauses begin with the particles listed above, as do sentential objects of some verbs – verbs of cognition, evaluation, observation, and speaking. These particles are in complementary distribution with other particles which do not have a tense function, such as the causal or factive particle *he*, seen in (5a). They are also in complementary distribution with the subjunctive particle *ke*, which introduces clauses embedded under modal verbs, and verbs of desire and intention as in (5b). Finally, we find a partially distinct series of tense/aspect markers in relative clauses and other operator-extraction contexts, as shown in (6):

(5) a. Gagao foki   nī   a   au he hifo   a   Maka ki tahi.
         sick   also EMPH ABS I   C   go-down ABS Maka to sea
         'I'm also sick of Maka going down to the sea.'
     b. Ne   foa   e   lautolu e   vala vao ke   tā
         PAST clear ERG they   ABS bushland SBJN build
         aki   e   fale pola.
         INST ABS   house
         'They cleared the bushland to build a thatch house.'

(6)   *Sentence-initial tense/aspect particles in operator-extraction clauses*

PAST/PRESENT(NFUT)	FUTURE	PROGRESSIVE	PERFECT
ne	ka	ne fā e	(ne) kua

     a. Ne   inu   e   Sione e   kofe *ne* taute e   au.
         PAST   drink ERG Sione ABS coffee NFUT make   ERG I
         'Sione is drinking the coffee that I made.'
     b. Ko e   tau fale fā hanei *ka*   tā   he maaga.
         'Ko' PL house four these FUT   build in village
         'These are the four houses that are going to be built in the village.'
     c. e   tagata *ne f ā e* onoono hake ke he mahina
         ABS man   PROG   look   up   to PRT moon
         'the man who's looking up at the moon'
     d. Ko   Fao hanā *ne kua* iloa ko ia ne   tā   mai
         'Ko' Fao this PERF   know 'ko' he NFUT bring from
         a   Avatele mo Kavatele mai i   Fonuagalo.
         ABS Avatele and Kavatele from LOC Fonuagalo
         'It was this Fao who is known for bringing Avatele and Kavatele from
         Fonuagalo.'

Sentence initial particles thus express information both as to the tense/aspect of the sentence, and to the grammatical status of the clause (matrix, subject clause, object clause, relativizing clause, etc.). In the case of object clauses, they also express information as to the selectional properties of the governing verb – i.e., whether it selects a subjunctive or a fully tensed clause. Given these facts, the initial particle is posited to be a portmanteau Complementizer/Tense element, CTP (i.e., Comp/Tense Phrase).

We can now question whether the verb-fronting operation in Niuean involves fronting to CTP. In fact, this appears not to be the case. This can be seen by an examination of the Niuean verbal complex, beginning with the preverbal elements, as shown below.

(7)   *Sentence-initial elements*

C/Tense	Negative	Auxiliaries	Verb

To   nākai   liu   feleveia foki   a   taua.
Fᴜᴛ not   again   meet   also   Aʙs   we
'We will never again meet.'

Note that the negative morpheme and auxiliary verb(s) intervene between the verb and the sentence initial CT particle. This means that if the verb is considered always to move to CT position, there must at some point in the derivation be an Xº of the form: [T-Nᴇɢ-Aᴜx-Verb]. It might be the case that the entire complex is a constituent at D-structure, in which case the entire complex moves to CT. Or, one or both of the Aᴜx and Nᴇɢ elements could appear in a head position, with the verb complex moving first to Aᴜx, then to Nᴇɢ, and so on.

These views are problematic, because Nᴇɢ appears to be an independent stem, itself able to host verbal clitics. There is a postverbal perfect marker *tuai* which often co-occurs with the CT perfect marker *kua*. It appears above in (4e). In a negative sentence, instead of appearing after the verb, *tuai* appears after the negative element.

(8)   Kua   nākai   tuai   fano   a   ia.
      Pᴇʀꜰ not   Pᴇʀꜰ   go   Aʙs   he
      'He has not gone.'

As well as the perfect particle, there is another verbal clitic which shifts to the post-negative position in a negative sentence. This is an emphatic marker, *lā*, seen in (9). (Note the negative element *nākai* alternates with a different form *ai* which appears in (9)). *Lā* often co-occurs with another emphatic marker, *ia*. In a negative sentence *lā* appears on the Nᴇɢ, while *ia* occurs on the verb.

(9)   a.   Ai   lā   kitia   e   au   e   pusi.
           not   Eᴍᴘʜ   see   Eʀɢ   I   Aʙs   cat
           'I have not yet seen the cat.'
      b.   . . . ko e   mena   kito   taute   lā ia   he   tau
           . . . 'ko'   thing recent make   just   Lᴏᴄ   Pʟ
           magahala fakamui nai.
           period-of-time-after-this
           'a thing just recently done.'
      c.   Nākai lā   nofo   ia   a   au   he ha   motu tufa a   Niue.
           not   Eᴍᴘʜ live   Eᴍᴘʜ   Aʙs   I   on Nsᴘ island like Aʙs Niue
           'I've never before lived on an island like Niue.'

A constituent question marker can also appear after the negative element:

(10)   Ai   kia kitia e   koe e   laa kua tokoluga?
       not Q   see   ERG you ABS sun PERF high
       'Didn't you see the sun high up?'

It has been argued (Chung 1970) that in other Polynesian languages the negative element is in fact a verb, since, for example, it houses verbal affixes and it takes as its complement a phrase which begins with an embedded CT element (cf. also Bauer 1997, Hohepa 1969, Pearce 1997, and Waite 1987). In Niuean, the negative element does not seem to be a full-fledged verb, since it does not appear with the complete range of verbal clitics (outlined below), and it does not take a CTP complement. However, its behavior is that of a syntactic head, rather than a lexical affix or clitic.

Since none of the material surrounding NEG is phrasal, it might still be possible to maintain the claim that [NEG-*lā*-*tuai*-AUX-Verb-. . .] is a single X° at some level. (See below for discussion of postverbal clitics.) But this necessitates a varying templatic morphology since the order of morphemes within the complex head would differ depending on whether or not it contained a negative morpheme. (In an affirmative sentence we find [AUX-Verb-. . . -*lā*-Y-*tuai*-. . .].) It is hard to explain this variation in morpheme order under a V-to-CTP view, whereas if we assume that NEG intervenes between CT and the fronted verb, we can explain why *tuai* and *lā* appear on the NEG element simply by observing that they are always on the head of the phrase that is a sister to CT (i.e., in the second head position, or "governed" by CTP). It is reasonable that a perfect morpheme need be governed by perfect tense, and that a temporal emphatic such as *lā*, meaning 'just/yet', need be governed by tense. There is no comparable reason why their relative order within a single head should matter.

Finally, as well as being an independent morphological head, NEG has sentential scope rather than scope over the verb alone. It is preferable, then, to consider NEG as an independent item in the syntax. I thus rule out the movement of V to CTP.

The argument presented above for not considering V-movement to be to CTP is based on negative sentences. In sentences where there is no NEG, it might appear that V to CT movement is a possibility (and note that NEG to CT movement remains a possibility). However, this would be true only if the so-called V-fronting is actually head movement. Below I argue that what is fronted is in fact a maximal projection (VP, PP, AP). If this is the case, it is clear that the projection could not move to the Spec of CTP and appear, as it does, to the right of the head of CTP. Note that if the claim of this chapter is granted (i.e., that VSO in Niuean involves maximal or remnant predicate fronting), the ordering of C before the predicate is an argument against movement to CTP in both affirmative and negative sentences.

For the sake of completeness, let us briefly examine auxiliaries. Auxiliaries in Niuean, unlike NEG, are morpho-syntactically inert, in that they never vary in their position with respect to other particles or phrases. They include desideratives,

habituals, and elements meaning 'look like', 'nearly', 'begin', among others (see Seiter 1980). An example (*fia* 'want') appears below:

(11) Ne fia   evaeva a   ia ka e nākai talia
      Pst want   walk   ABS he but not   let
      he   matua ke   taute pihia.
      Erg parent SBJN do so
      'She wanted to go for a walk, but her parent wouldn't let her do so.'

Auxiliaries appear between Neg and V. They do not show evidence of independence from the verb, since their position is fixed and no material may intervene between the auxiliary and the verb. While, as seen above, in a sentence with Neg, the perfect element *tuai* and the emphatic *lā* appear on Neg rather than on the verb, in a sentence with no Neg, but with an auxiliary, *tuai* and *lā* appear on the verb, not on the auxiliary. Thus, auxiliaries, unlike Neg and V, do not display the head-like property of taking particles. As stated above, *tuai* and *lā* appear on the first head after CTP. They never appear on auxiliaries; hence auxiliaries are not heads. Instead, I consider that the auxiliaries are verbal prefixes, or perhaps more correctly, that they form compounds with verbs. Under the entry for *fia* Sperlich (1998) states that ". . . some speakers may want to consider such constructions either as compounds or as verbs with a prefix. . ." This view is supported by the fact that the auxiliaries appear to be completely inert syntactically. It is further (weakly) supported by Seiter's observation that one auxiliary is orthographically represented as a prefix on a verb (*fiakai* 'hungry, lit. 'want-eat', *fiamohe* 'sleepy', lit. 'want-sleep'). Similar support is found in the fact that several auxiliaries are redundant expressions of portions of meaning of the verb with which they appear, e.g., *fā* 'habitual' with *mahani* 'typical', or *fia* 'desiderative' with *loto* 'want' or with *manako* 'desire'.

Given the discussion above, we can adopt the structure in (12) for the Niuean clause, where there is a C/Tense phrase, then a Negative phrase, then the position to which the verb fronts. This position is followed by the VP. Following a convention commonly adopted for other languages, I call the position to which the verb fronts Infl. Since the tense morphology does not actually appear in this slot, and since I claim that what does appear here is a predicate EPP feature, this position might also be termed Pred (cf. Bowers 1993), but here I maintain the more familiar term. I leave aside here detailed discussion of exactly where the subject and object should be positioned at spellout, but see (18), (21), and (25) below, and Massam ([1998]).

(12)   [$_{CTP}$ CT [$_{NEGP}$ Neg [$_{IP}$ Infl [$_{VP}$ Subject V Object]]]]

Niuean VSO order is thus preliminarily considered to arise from movement of V to Infl, with the understanding that Infl does not contain tense features. (This will be revised in section 2 to accommodate predicate fronting of maximal projections.) Note, however, that the attracting feature in Infl is yet to be fully discussed.

2   Maximal Projection Predicate Fronting (VSO is [Vt_o]SO)

2.1   Non-verbal Predicate Fronting (PredSO not VSO)

A central fact of word order in Niuean is that in many instances it is not a verbal element (as it is in 13a) which appears in the V-slot, but that instead it can be an apparent nominal element, as seen in (13b–e), or a prepositional element, as in (14). In these examples, the predicate and its following particles are shown in brackets. This observation has been made for a variety of VSO languages (e.g., for Tahitian by Lazard and Peltzer 1991). The theoretical significance of this fact, however, is not always appreciated (but cf. Carnie 1993, 1995, for Irish, and Lee [2000], for Zapotec). Predicate nominals are marked with *ko* (proper) and *ko e* (common), which I gloss here simply as 'ko', reserving further discussion of its category and meaning until section 3.2 below.

(13) a. [Ne inu] e Sione e kofe.
          PAST drink ERG Sione ABS coffee
          'Sione drank the coffee.'
      b. [Ko Mele] e   faiaoga.
          'Ko' Mele ABS teacher
          'The teacher is Mele.'
      c. Ai [ko e faiaoga] a   Mele.
          not 'ko teacher ABS Mele
          'Mele's not the teacher.'
      d. [Ko e tipolo agaia nī]   ne   inu ai   a   lautolu.
          'Ko' lime still EMPH NFUT drink PRON ABS they
          'It's still only lime juice that they are drinking.'
      e. [Ko e fale   ke   lima aki] e   fale i   kō.
          'Ko' house SBJN five INST ABS house LOC there
          'That house over there is the fifth house.'

Seiter (1980) provides evidence that the fronted *ko (e)* nominal is in the same slot as the fronted verb, since it follows the NEG (13c) and precedes the usual post-verbal adverbs (13d). In (13e), we see a clear example of a fronted maximal projection, namely a nominal with a modifying relative clause (literally translated as 'the house that (one) fives with'.) Other examples with clearly phrasal predicates are given in (16) and (23) below.

There are two immediate theoretical consequences of the observation that the verbal slot may be filled with a maximal projection. First, Niuean VSO order results at least some of the time from non-verbal predicate-fronting rather than from V-fronting. In Minimalist terms, this means that the strong attracting feature in Infl that provokes the predicate initial word order must be [PRED] rather than the purely categorial feature [V] (cf. Bowers 1993). Second, it appears to be the case that the fronted predicate is at times an X°, as in the case of V-fronting as in (13a), and at times an XP, as in the case of predicate nominal (13b–e) or prepositional fronting as in (14).

(14) Hā   he fale   a   ia.
      Pred in house Abs she
      'She is in the house.'

Maintaining a conservative view of X-bar theory (vs., for example, Carnie 1995 or Chung 1990), wherein only non-phrasal X° elements can appear in or adjoined to head positions, it must be the case that phrasal predicates (e.g., 13e) are fronting to the specifier position of IP, rather than to the head position of IP. If, as assumed, in all instances the moved element is fronting to check a [Pred] feature, this raises questions regarding Pied Piping. (15) would appear to hold:

(15) *Bifurcated description* [Pred] *checking* (to be resolved below)
      a.   If the predicate is verbal, it alone moves to check [Pred], none of its arguments may move with it.
      b.   If the predicate is nominal or prepositional, it moves along with a subset of its phrasal arguments/modifiers (obligatorily).

There are two problems with this situation. We must ask what particular separate paths of movement the two predicate types take (head to head vs. specifier to specifier), and whether these might not be expected to correlate with other word order differences (cf. Rackowski and Travis [2000]). I leave this issue aside for now. In addition, there is a philosophical difficulty, namely that there is no apparent reason why predicate types should behave differently in this way. Instead of trying to formulate a justification for this bifurcation in predicate behavior, let us explore the possibility that all [Pred] checking (even purely verbal [Pred] checking) involves the fronting of a maximal projection. In the case of verbal predicates, this will involve so-called remnant VP-fronting (cf. Lee [2000]). To support this position, I first show that there are sentences where verbal [Pred] checking clearly does involve the fronting of a maximal projection.

2.2   VP Predicate-Fronting: Noun Incorporation (VSO vs. VOS)

In this section, I show that "V-fronting" can, at least in some cases, involve VP-fronting. In order to show this, we first corroborate and extend Seiter's (1980) observation that so-called Noun Incorporation (NI) in Niuean can incorporate non-head material, as seen below. In (16a), for example, the nominal *sipi mo e ika mitaki* 'good fish and chips', which would not be considered an X°, is claimed by Seiter to have incorporated onto the verb. In each example below, the "incorporated" nominal is bracketed. (See Massam [1998], for a more detailed discussion of Niuean NI and of the analysis presented briefly here.)
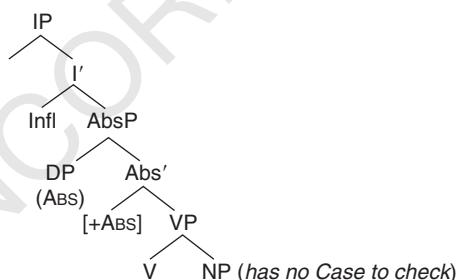
(16) a.   Ne kai [sipi mo e   ika mitaki] a   Sione.
           Past eat chip Com Abs fish good   Abs Sione
           'Sione ate good fish and chips.'
      b.   Ne   holoholo [kapiniu kiva] fakaeneena   a   Sione.
           Past wash   dish   dirty slowly   Abs Sione
           'Sione is washing dirty dishes slowly.'

c. Kua   leva   lahi   e   amaama aki   ke
   PERF longtime   very   ABS awaiting INST SBJN
   fai [pepa pehē   nai] kua tohia ke   fakamau aki   e
   be   book   like   this PERF written SBJN   retain INST ABS
   tau puhala gahua lima   he   motu   ha   toutolu ko Niue. . .
   PL   ways   work hand   of   island of   ours   'ko' Niue
   'There has been a long time of waiting for there to be a book like this,
   which was written to retain the ways of handiwork of our island. . .'

Although Niuean sentences such as these have traditionally been labeled as NI sentences, they cannot be accounted for by analyses of NI, since the putative incorporated nominal is not an X°. We can instead account for sentences such as (16) in Niuean as does Massam ([1998). In that article, I argued that the "incorporated" nominal (bracketed in 16) originates as an NP direct object, rather than a DP direct object, such as found in non-NI transitive clauses. This claim is supported by the fact that incorporated nominals may not be preceded by any of the case number and determiner particles (e.g., *tau* 'PL', as in 2c), which usually appear at the beginning of a DP constituent. An NP object cannot check the absolutive Case, so it does not move out of VP to a checking position, but instead remains *in situ* within the VP, and gets fronted along with the verb by predicate-fronting.
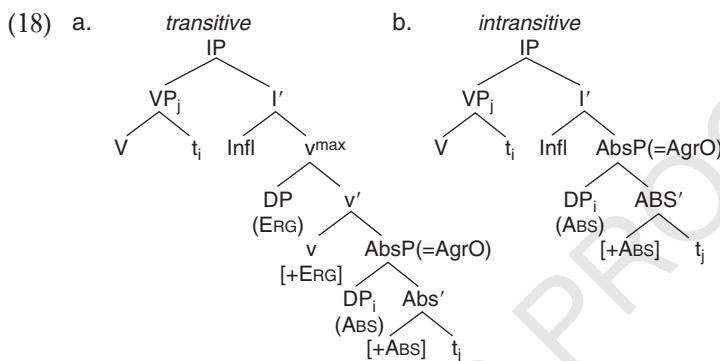
   In (17), we see a structure with so-called noun incorporation. Here, the external DP argument is generated in specifier of AbsP, and an NP is base-generated in object position. This NP cannot check the absolutive Case feature, since an NP (as opposed to a DP) has no Case feature[.] Hence, this NP could not move out of complement position to establish a checking relation with the [+ABS] feature. As a result, when VP fronting occurs, it will involve the movement of the VP in bold to the specifier of IP. In other words, the NP object will front along with the verb.

(17)   "*noun incorporation*"

This analysis, which accounts for all of the basic "NI" facts in Niuean, shows us that non-head material can also be fronted in cases where the predicate is a verb, as well as when the predicate is a DP or a PP. The apparent norm then becomes the exception; that is, the cases where a simple V appears to be fronted are unusual (albeit more common), since most cases of predicate fronting appear to involve XPs rather than X°s. We can gather the one exception into the generalization by considering that in non-"NI" verbal VSO sentences too, the VP is the constituent

that moves (in fact, it is the VP remnant, since all that remains in the VP is the verb). This is shown below. (18a) illustrates a transitive clause, where the DP object moves to AbsP to check the absolutive Case feature. (18b) shows the same situation in an intransitive clause. In these two cases, when the VP fronts to check the [PRED] feature in Infl, the trace of the moved DP fronts also. (Following Legate 1997; I assume obligatory predicate reconstruction at LF, thus allowing the moved DP to c-command its VP-internal trace at LF.)

(18) a. *transitive*   b. *intransitive*

Thus, in non-incorporating verbal predicate sentences, no argument is fronted with the verb, because if the arguments are DPs, PPs or CPs, they will be in Spec of AbsP, rather than included in VP; hence, they are not moved along with VP (see also Lee [2000]).]pi[[. . .]

3 Predicate-fronting as an EPP reflex

3.1 Why does the predicate move?

What is the nature of the Niuean predicate-fronting? I claim that it is an EPP (Extended Projection Principle) reflex, as argued in Massam and Smallwood (1997). Niuean predicate-fronting obviously differs from V-raising to Infl in languages such as French and English, where maximal and non-verbal predicates never substitute for V. Another difference between Niuean and most European languages is that the latter require a grammatical subject, due to the EPP requirement instantiated by the [D] feature in Infl (Chomsky 1995). However, Niuean, a strict VSO language, provides no evidence for [D] in Infl. (This is also argued to be the case for Irish by McCloskey [1996]; see also Doron [2000]). Subjects c-command objects, as evidenced by binding facts (according to Seiter 1980), but are otherwise not much structurally distinguished (by extraction, including raising, quantifier float, etc.). Further, Niuean has no expletives. This is in contrast to English, where a strong nominal [D] in Infl forces a subject–predicate structure for all clauses, resulting in some cases in the presence of an expletive subject as in (22) (cf. Chomsky 1995, and Smallwood 1996, who extends this requirement to small clauses).

(22)   There are students in the classroom.

Niuean (but not English) predicates can check features in Infl, and English (but not Niuean), requires a grammatical subject. Massam and Smallwood (1997) relate these two properties by claiming that the Niuean head of IP has no [D] feature; thus, the specifier need not be filled by an element checking [D], but instead can be filled by the predicate checking the [PRED] feature (which they actually refer to as a Tense feature) on the head of Infl. In this way they account for predicate nominals in Niuean. In this chapter I extend this claim to encompass all clause types in Niuean. Thus, regular verbal predicate clauses and NI clauses, as well as predicate nominals and PPs, involve the fronting of a maximal projection to the specifier position of Infl. It is thus impossible for a language with a true predicate- (as opposed to V-) fronting requirement to instantiate initial subjects, since the specifier of Infl is filled by the predicate. Thus, as in Massam and Smallwood (1997), the Niuean [PRED] feature parallels the English [D] feature, in that the strict EPP nature of English is mirrored by the strict VSO nature of Niuean. [D] and [PRED] are thus in complementary distribution and can be seen as two reflections of a single EPP predication feature. This might explain the noted correlation between VSO and the absence of copular verbs, since maximal non-verbal predicates can undergo predicate-fronting and there is no need for a V in Infl.

3.2 What is a predicate? (X + YP)

It is claimed that the notion of predicate is central to the grammatical structure of Niuean in that each clause must externalize its predicate following the EPP. It is thus important to be clear as to what constitutes a Niuean predicate. It was noted above that when the predicate is nominal, it appears to front with its arguments, whereas when it is verbal, it seems that only the verb itself fronts. It was argued that it is preferable and possible to regard predicate-fronting as uniformly moving a maximal projection. But, in a sense, this has merely shifted the question sideways. While we now consider predicate-fronting always to move a maximal predicate, we must ask why, when the predicate is a nominal, it can front along with an apparent external argument (a possessor) and an internal argument, but when the predicate is a verbal or prepositional phrase, only the internal argument may front (overtly in case of NI, and as a trace in non-NI clauses). That a nominal predicate can include external and internal arguments is seen in (23).

(23) [Ko e   poka-aga   he   tama e   maka]
      'Ko'   push-ING   GEN   child ABS rock
      [ati   matakutaku ai   e   kulī].
      [reason_i fear   PRON_i ABS   dog]
      'The reason the dog was afraid was the child's pushing the rock.'

Of course, the apparent main difference between the nominal and verbal predicate sentences is that in an example like (23) it appears to be the entire DP (including the external argument *he tama* 'the child') which is acting as the predicate. This predicate then has an additional external argument (namely, *ati matakutaku ai e kulī* 'the reason the dog was afraid'). In verbal predicate examples, the verbal

predicate can have no subject independent of one of its own thematic arguments (either the external agent in a transitive, or the moved object in an intransitive).

We can eliminate this last remaining asymmetry, however, by paying attention to the fact that predicate nominals in Niuean always appear with the pre-nominal marker *ko (e)*, as we observed above. In (13b–e) and (23), for example, the nominal phrase is marked by the preposition *ko (e)*, (Likewise, a predicate PP requires an extra preposition, *hā*, to render it a predicate.) Seiter considers this to be a pre-nominal predicate marker. In Massam (1996), however, it is argued that *ko (e)* is best considered a preposition (following Clark 1976). This preposition marks a nominal phrase as being a non-argument. In this view, predicate nominals in fact reduce to predicate PPs.
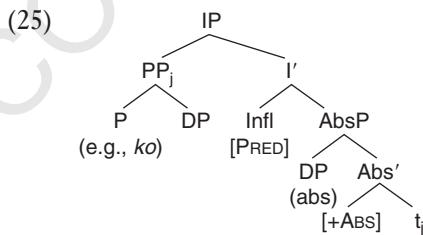
Given the identification of predicate nominal and predicate prepositional phrases, we can now unify the notion "predicate" across Niuean clause types, and explain why nominal arguments other than direct objects can front with a nominal predicate but not with a verbal predicate. Quite simply, the predicate in (13b–e), just as the predicates in (3) and (14) consists of a [–N] element (either *ko (e)*, or another preposition, or a verb) and its complement. The external arguments of all the prepositional predicates, as well as any modifying obliques, are not considered as part of the predicate, just as in the case of verbal predicates.

The notion "predicate" can thus be uniformly defined in Niuean as in (24), which is the revised form of (15). The clitics mentioned in the definition are the first four in (19) – those which appear adjoined to VP, namely the directional and manner particles – and *aki*, *oti*, and *ai*. Remaining unclear is exactly why EPP predicates should be so defined, and the exact role of the clitics in Niuean syntax.

(24)   *Unified description of what checks [Pred]* (Revision of 15)
        [[X YP] + clitics], where X is a lexical predicate, YP its internal argument.

Thus, while a verbal predicate sentence has the structure in (18) (with or without "NI"), or in (21), which shows the clitics, a nominal or prepositional predicate sentence will have the structure in (25).

(25)

The DP argument of the preposition, unlike that of a verb, will always undergo movement as part of the predicate because it receives inherent Case from the preposition, and thus never undergoes movement to AbsP to check Case. All predicate types can thus receive a unified treatment, if we allow that the head predicate in a predicate nominal sentence is not the head of the nominal phrase, but instead is the prepositional element *ko (e)*.

4 Conclusion

This chapter has analyzed the nature of Niuean VSO, arguing that this order is derived, and that it is derived not by movement to C, nor to Tense, but rather to a lower functional head, labeled Infl (or Pred), which houses an uninterpretable EPP feature [Pred]. Next, it was shown that verb-fronting is to be more properly considered as predicate fronting in this language, since maximal phrases which are demonstrably non-verbal also front to the same slot. This raises the question of why there should be a bifurcation in the language between X° verbal predicates and XP non-verbal predicates. Evidence from so-called Noun Incorporation in Niuean allows us to see that at least in some cases in verbal sentences it is VP rather than V° which is fronted. From this, I posited that in all cases predicate-fronting involves XPs, but in case of an object DP, the object will have moved out of VP prior to movement of VP, so that what fronts is a VP remnant consisting of the verb and the trace of the object. Implications of the analysis regarding postverbal elements were discussed. The claim was made (following Massam and Smallwood 1997) that predicate fronting in Niuean is in complementary distribution to subject externalization in SVO languages, and that in both cases the movements are a reflection of an Extended Projection Principle feature: [D] in SVO and [Pred] in VSO. A reanalysis of predicate nominals as predicate prepositional phrases allowed for the notion "Predicate" to be uniformly defined in Niuean as a lexically defined predicative head plus its internal argument and clitics.

This work puts forward the notion that at least in some cases, VSO languages belong to a parametrically different class than SVO languages, in that the former can involve a [Pred] feature rather than a [D] feature as the EPP feature. This chapter can thus be placed within a growing body of work which argues that VSO order is in some cases to be viewed as a variant on VOS order.

Note

5   There are also arguments from binding (Seiter 1980) that the subject c-commands the object, which also argues against a flat VSO structure (though not necessarily for a V′ constituent). In addition, there are strong reasons for assuming that at least absolutive arguments appear at PF in a non-thematic position, since they can be in positions to which they have been raised. This also argues against a flat structure.

29.3   Questions pertaining to Massam (2000)

1   Discuss at least two ways in which the characterization of languages like Niuean as VSO is unsatisfactory.
2   When the verb moves past the subject in Niuean, it accomplishes a movement that recalls what takes place in English in sentences like *Where are the books?* What are the similarities and what are the differences?

3   Is Massam right to reject a "flat" analysis of Niuean VSO? Give your reasons.

4   Massam notes that Niuean tense/aspect morphemes do not co-occur with the subjunctive morpheme *ke*. How might one express the relation between this property of Niuean and the fact that French subjunctives (which do not show a distinctive complementizer) lack a present/past distinction (unlike French indicatives)? (Extra credit: How does all this fit into Cinque's 1999 functional hierarchy?)

5   Facts such as those mentioned in question 4 lead Massam to speak of a "portmanteau" Complementizer/Tense element. What might an alternative approach in terms of silent elements look like?

6   The perfect marker *tuai* is normally postverbal, but when negation is present, Niuean has *tuai* immediately following the negation: '*Neg tuai V. . .*'. How might one relate this to the interaction between negation and agreement in Finnish as discussed by Mitchell (1991)?

7   Massam mentions Chung's (1970) argument that the negative element in some Polynesian languages is a verb, insofar as it is followed by verbal affixes and by a complement beginning with a Complementizer/Tense element. Thinking of Culicover (1971), to what extent do English sentences like *It's essential that he not be back before noon* and *It's essential that he not* suggest an alternative?

8   Massam argues strongly that the verb in Niuean moves to a position below negation rather than to the C/T level. How similar is her analysis of Niuean here to Pollock's (1989) analysis of French and English? (Extra credit: Bring in Cinque 1999 on Italian negation.)

9   Niuean verb movement is shown by Massam to actually be phrasal movement, in part on the basis of the fact that it can carry along an incorporated NP. She takes this incorporated phrase to be NP rather than DP because it is not allowed to contain Case, number or determiner elements. Discuss the extent to which English has comparable NP-incorporation.

10 Massam argues that full DP objects are never carried along by VP-movement in Niuean because full DP objects must raise to a Case-licensing position outside VP prior to VP-movement. How does this point of Massam's tie in with Kayne's (2003) discussion of SOV sentences?

11 Massam mentions a "noted correlation between VSO and the absence of copular verbs." Test this correlation with as many VSO languages (from as many different families) as you can.

12 The EPP is taken by Massam to underlie VP-fronting in Niuean, as well as the more general predicate-fronting that she takes VP-fronting to be a subcase of, the idea being that languages can choose whether to satisfy the EPP by predicate fronting, as in Niuean and other verb-initial languages, or by argument fronting (DP-movement to Spec,IP), as in English and other SVO languages. To what extent does this lead to the expectation that verb-initial and SVO languages should be equally common? (Extra credit: How do SOV languages fit in here?)

13 Fronted predicate nominals in Niuean must be preceded by a preposition *ko (e)*. Why might that be? How plausible is it that there is a link to English *People usually take him *(for) an idiot*. (Extra credit: Bring in Pollock 1983.)
14 How important a role does remnant movement play in Massam's proposal?
15 When the object is a full DP, Niuean has VSO order. There are, on the other hand, languages where VOS order is the normal one even when the object is a full DP. Where should the difference between VOS and VSO be located? Contrast in particular a remnant movement approach with one in which the object scrambles by itself. (Extra credit: Bring in Ordóñez 1998.)

References

Anderson, S. R., and S. Chung. 1977. On grammatical relations and clause structure in verb-initial languages. In P. Cole and J. M. Sadock (eds), *Grammatical Relations. Syntax and Semantics* 8, New York: Academic Press, pp. 1–25.

Bauer, W. (with William Parker, Te Kareongawai Evans and Te Aroha Noti Teepa). 1997. *The Reed Reference Grammar of Māori*. Auckland: Reed Publishing.

den Besten, H., and G. Webelhuth. 1990. Stranding. In G. Grewendorf and W. Sternefeld (eds), *Scrambling and Barriers. Linguistik Aktuell/Linguistics Today* 5, Amsterdam and Philadelphia: John Benjamins, pp. 77–92.

Bowers, J. 1993. The syntax of predication. *Linguistic Inquiry* 24: 591–656.

Carnie, A. 1993. Nominal predicates and absolutive case marking in Irish. In C. Phillips (ed.), *Papers on Case and Agreement II: MIT Working Papers in Linguistics* 19: 89–130.

Carnie, A. 1995. *Non-verbal predication and head movement*. Doctoral dissertation, MIT.

Carnie, A., H. Harley, and E. Pyatt. 2000. VSO order as raising out of IP? Some evidence from Old Irish. In A. Carnie and E. Guilfoyle (eds), *The Syntax of Verb Initial Languages*, New York and Oxford: Oxford University Press, pp. 39–60.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chung, S. 1970. *Negative verbs in Polynesian*. Ms., University of California, Santa Cruz.

Chung, S. 1990. VPs and Verb movement in Chamorro. *Natural Language and Linguistic Theory* 8: 559–620.

Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York and Oxford: Oxford University Press.

Clark, R. 1976. *Aspects of Proto-Polynesian*. Auckland: Linguistic Society of New Zealand.

Coon, J. 2010. VOS as Predicate Fronting in Chol. *Lingua* 120: 354–378.

Culicover, P. 1971. *Syntactic and semantic investigations*. Doctoral dissertation, MIT.

Doron, E. 2000. VSO and left-conjunct agreement: Biblical Hebrew vs. Modern Hebrew. In A. Carnie and E. Guilfoyle (eds), *The Syntax of Verb Initial Languages*, New York and Oxford: Oxford University Press, pp. 75–96.

Emonds, J. 1980. Word order in generative grammar. *Journal of Linguistic Research* 1: 33–54.

Hendrick, R. 2000. Celtic initials. In A. Carnie and E. Guilfoyle (eds), *The Syntax of Verb Initial Languages*, New York and Oxford: Oxford University Press, pp. 13–38.

Hohepa, P. 1969. *Not* in English and *kore* and *eehara* in Māori. *Te Reo* 12: 1–34.

Kayne, R. S. 2003. Antisymmetry and Japanese. *English Linguistics* 20: 1–40. [Reprinted in Kayne (2005, Ch. 9).]

Lazard, G., and L. Peltzer. 1991. Predicates in Tahitian. *Oceanic Linguistics* 30: 1–31.

Lee, F. 2000. VP Remnant Movement and VSO in Quiaviní Zapotec. In A. Carnie and E. Guilfoyle (eds), *The Syntax of Verb Initial Languages*, New York and Oxford: Oxford University Press, pp. 143–162.

Legate, J. 1997. *Irish predication: A Minimalist analysis*. MA thesis, University of Toronto.

Massam, D. 1996. Predicate Fronting in Niuean. In *Proceedings of the 1996 Canadian Linguistics Association Annual Meeting, Calgary Working Papers in Linguistics*, pp. 247–256.

Massam, D. 1998. Niuean noun incorporation: Not nouns, not incorporation. Ms., University of Toronto. Presented at AFLA V, University of Hawaii.

Massam, D., and C. Smallwood. 1997. Essential features of predication in English and Niuean. In *Proceedings of NELS 27*, Amherst: GLSA, University of Massachusetts, pp. 236–272.

McCloskey, J. 1991. Clause structure, ellipsis, and proper government in Irish. In *The Syntax of Verb Initial Languages*. *Lingua Special Edition* 85: 259–302.

McCloskey, J. 1996. On the scope of verb movement in Irish. *Natural Language and Linguistic Theory* 14: 47–104.

Mitchell, E. 1991. Evidence from Finnish for Pollock's theory of IP. *Linguistic Inquiry* 22: 373–379.

Müller, G. 2004. Verb-second as vP-first. *Journal of Comparative Germanic Linguistics* 7: 179–234.

Nilsen, Ø. 2003. *Eliminating positions: The syntax and semantics of sentence modification*. Doctoral dissertation, Utrecht University.

Ordóñez, F. 1998. Postverbal asymmetries in Spanish. *Natural Language and Linguistic Theory* 16: 313–346.

Pearce, E. 1997. Negation and Indefinites in Māori, In D. Forget, P. Hirschbühler, F. Martineau and M.-L. Rivero (eds), *Negation and Polarity: Syntax and Semantics*, Amsterdam: John Benjamins, pp. 271–288.

Pearson, M. 2001. *The clause structure of Malagasy: A minimalist approach*. Doctoral dissertation, UCLA.

Pollock, J.-Y. 1983. Sur quelques propriétés des phrases copulatives en français. *Langue Française* 58: 89–125.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Rackowski, A. and L. Travis. 2000. In A. Carnie and E. Guilfoyle (eds), *The Syntax of Verb Initial Languages*, New York and Oxford: Oxford University Press, pp. 117–141.

Seiter, W. J. 1980. *Studies in Niuean Syntax*. Outstanding Dissertations in Linguistics. New York: Garland.

Smallwood, C. 1996. *A Minimalist account of English predication*. MA thesis, University of Toronto.

Sperlich, W. 1998. *Tohi Vagahau Niue: Niue Language Dictionary*. Honolulu: University of Hawaii Press and Government of Niue.

Sproat, R. 1983. *VSO Languages and Welsh Configurationality*. *MIT Working Papers in Linguistics* 5: 243–276.

Waite, J. 1987. Negatives in Māori: A lexical-functional approach. *Te Reo* 30: 79–100.

Woolford, E. 1991. VP-internal subjects in VSO and nonconfigurational languages. *Linguistic Inquiry* 22: 503–540.

UNCORRECTED PROOFS

30

Derivation by Phase

Noam Chomsky

2001

30.1   Introduction

The Principles and Parameters (P&P) framework takes grammar to be characterized by a set of invariant principles, which account for crosslinguistic commonalities, and a number of parameters, which permit a restricted envelope of variation (cf. Rizzi 1980, 1982, and Travis 1989, excerpted in Ch. 8, 11, and 16 of this volume). In the first important instantiation of the P&P framework, Government and Binding (GB) theory (Chomsky 1981), sentences are generated beginning at the level of D-structure (DS), where properties of lexical items are projected onto syntactic structure. The level of S-Structure (SS) is then derived from DS via a set of movement operations. The SS level serves as input both to a further set of movement operations to give rise to Logical Form (LF), the interface with semantic interpretation, as well as to a set of operations of a different sort to yield Phonological Form (PF), the interface with articulation and perception. Thus a sentence is characterized by four distinct levels of representation. The principles of GB theory have their effect by imposing constraints on these levels. Parameters, in contrast, are choice points, where a language can opt for one or another way to satisfy some constraint. For example, a constraint governing the distribution of *wh*-questions might require that *wh*-elements appear in a certain local configuration, with parameters allowing for variability in whether one or all *wh*-elements must appear in this configuration and in the level of representation at which such a configurational relation needs to hold (cf. Rudin 1988; Huang 1982). The theory thus determines the well-formedness of sentences through restrictions on derivational machinery and constraints on levels of representation.

The Principles and Parameters framework was the source of productive explorations across a range of languages. Yet, as the framework was applied to a wider array of data, the complexity of the set of principles and parameters increased. Chomsky's *Minimalist Program* is an attempt to subject this complexity to careful scrutiny, with the goal of refining the core theoretical underpinnings of syntactic theory while preserving the empirical gains of GB. Part of this scrutiny concerns the levels of representation of the theory. As Chomsky (1993) observes, the PF and LF levels of representation are conceptually necessary in a theory that aims to characterize a mapping between meaning and form. In contrast, DS and SS, however useful, have no such conceptual motivation, and thus in a minimalist theory ought to be eliminated. A second kind of scrutiny concerns the nature of grammatical principles and parameters. In GB, there was little limit on what these could be or where they could apply, and indeed there were some that were surprising in form, even if they did substantial empirical work (e.g., the Empty Category Principle, see Hornstein and Weinberg 1995). With the elimination of DS and SS, well-formedness constraints could no longer apply at these levels. Chomsky suggests that constraints at the interface levels of LF and PF may remain in a minimalist syntax, but should only have the role of "legibility conditions." That is, such constraints should have the effect of ensuring that these levels of representation are appropriately interpretable by the cognitive systems with which they interface. Concerning parameters, they are retained, but as properties of elements of the lexicon, more specifically functional heads, following Borer (1984).

These changes have substantially reduced the amount of explanatory burden that can be borne by constraints on levels of representation. The minimalist program has instead increased the amount of explanatory work done by the properties of the derivational system. Derivations now consist of the interleaved application of structure building and transforming operations. MERGE combines two syntactic objects to form another (like the generalized transformations of Chomsky 1955/1975), MOVE displaces a syntactic object to the edge of a structure containing it, and AGREE establishes a relation of agreement between two syntactic elements. Much work in minimalism has attempted to limit the computational power of these operations, by limiting their application in certain ways. First, unlike the free application of operations in GB theory, where they were only limited by their ability to yield well-formed structures at the various levels of representation, the derivational operations of minimalism must be triggered by the requirements of lexical items, as determined by their "uninterpretable" or "unvalued" features. Further, displacement and agreement are restricted in their scope by a limited search criterion, which enforces a locality condition similar to Relativized Minimality (Rizzi 1990). Finally, when multiple derivational operations are available, economy conditions on derivational operations can force the application of one operation over another (e.g., Merge over Move).

Derivation by Phase develops this derivational picture further, arguing that the syntactic derivation is decomposed into multiple stages, recalling the idea of cyclic domains from Chomsky (1973). In the current paper, a PHASE is a piece of structure whose derivation is encapsulated. The phase plays a significant role in the derivation: it serves as a point at which an intermediate result of the derivation is "spelled out" and given an interpretation at both the PF and LF interfaces. Phases have an

important place in the enforcement of locality conditions on syntactic relations: the PHASE IMPENETRABILITY CONDITION (PIC) demands that movement out of a phase can proceed only from its edge, yielding an effect very similar to the use of the COMP escape hatch familiar from the SUBJACENCY CONDITION. Furthermore, economy conditions can only compare derivational alternatives that are present in the same phase.

The idea of phases has been very influential in subsequent work. The development of phase-based locality conditions has continued, not only for movement (cf. Müller 2011), but also for anaphora and morphological agreement (Polinsky and Potsdam 2001; Bhatt 2005). Phasal spell-out has played an important role in characterizing phrasal stress (Kahnemuyipour 2009; Kratzer and Selkirk 2007), in accounts of object shift (Fox and Pesetsky 2005), in accounting for the distribution of unpronounced elements (Kayne 2006), and in determining semantic, allomorphic, and phonological interpretation at the word level (Arad 2003; Marantz 2007, 2012; Kaye 1995; Newell 2008).

30.2   From "DERIVATION BY PHASE"

What follows extends and revises an earlier paper ("Minimalist Inquiries," MI [Chomsky 2000]), which outlines a framework for pursuit of the so-called Minimalist Program, one of a number of alternatives that are currently being explored. The shared goal is to formulate in a clear and useful way – and to the extent possible to answer – a fundamental question of the study of language, which until recently could hardly be considered seriously and may still be premature: to what extent is the human faculty of language FL an optimal solution to minimal design specifications, conditions that must be satisfied for language to be usable at all? We may think of these specifications as "legibility conditions": for each language L (a state of FL), the expressions generated by L must be "legible" to systems that access these objects at the interface between FL and external systems – external to FL, internal to the person.

The strongest minimalist thesis SMT would hold that language is an optimal solution to such conditions. The SMT, or a weaker version, becomes an empirical thesis insofar as we are able to determine interface conditions and to clarify notions of "good design." While the SMT cannot be seriously entertained, there is by now reason to believe that in non-trivial respects some such thesis holds, a surprising conclusion insofar as it is true, with broad implications for the study of language, and well beyond.

Note the indefinite article: *an* optimal solution. "Good design" conditions are in part a matter of empirical discovery, though within general guidelines of an a prioristic character, a familiar feature of rational inquiry. In the early days of the modern scientific revolution, for example, there was much concern about the interplay of experiment and mathematical reasoning in determining the nature of the world. Even the most extreme proponents of deductive reasoning from first principles, Descartes for example, held that experiment was critically necessary to discover which of the reasonable options was instantiated in the actual world. Similar issues arise in the case at hand (see MI for discussion).

Tenable or not, the SMT sets an appropriate standard for true explanation: anything that falls short is to that extent descriptive, introducing mechanisms that would not be found in a "more perfect" system satisfying only legibility conditions. If empirical evidence requires mechanisms that are "imperfections," they call for some independent account; perhaps path-dependent evolutionary history, properties of the brain, or some other source. It is worthwhile to keep this standard of explanation in mind whether or not some version of a minimalist thesis turns out to be valid.

These considerations bear directly on parametric variation, in this case yielding conclusions that are familiar features of linguistic inquiry. Any such variation is a prima facie imperfection: one seeks to restrict the variety for this reason alone. The same goal is grounded in independent concerns of explanatory adequacy/ learnability, which require further that ineliminable parameters be easily detectable in data available for language acquisition. Both kinds of considerations (related, though distinct) indicate that study of language should be guided by the *Uniformity Principle* (1).

(1)   In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances.

One familiar application is the thesis that basic inflectional properties are universal though phonetically manifested in various ways (or not at all), stimulated by Jean-Roger Vergnaud's influential Case-theoretic proposals 20 years ago [Vergnaud 1977]. Another is the thesis proposed by Hagit Borer [Borer 1984] and others that parametric variation is restricted to the lexicon, and insofar as syntactic computation is concerned, to a narrow category of morphological properties, primarily inflectional. These have been highly productive guidelines for research, extending earlier efforts with similar motivation (e.g., efforts to reduce the variety of phrase structure and transformational rules). What counts as "compelling" is, of course, a matter of judgment: there is no algorithm to determine when apparently disconfirming evidence is real or is the effect of unknown factors, hence to be held in abeyance.

On such grounds, we try to eliminate levels apart from the interface levels, and to maintain a bare phrase structure theory and the *Inclusiveness Condition*, which bars introduction of new elements (features) in the course of computation: indices, traces, syntactic categories or bar levels, and so on. The indispensable operation of a recursive system is Merge (or some variant of it), which takes two syntactic objects α and β and forms the new object $\gamma = \{\alpha, \beta\}$. We assume further that γ is of some determinate type: it has label $LB(\gamma)$. In the best case, $LB(\gamma) = LB(\alpha)$ or $LB(\beta)$, determined by general algorithm. Merge yields the relation Immediately-Contain (IC) (equivalently, Is-a-Member-Of), holding of (γ, α) and (γ, β). Iterated Merge, required in any recursive system, yields Contain (equivalently, Term-Of). Arguably Merge also yields a relation between α and β (sister); transitive closure yields C-Command (and also Contain and Identity, presumably available independently).

While (iterated) Merge "comes free," any other operation requires justification. Similarly, any features of lexical items that are not interpretable at the

interface require justification. That includes most (maybe all) phonological features; these must be deleted or converted to interface-interpretable form by the phonological component. One might ask to what extent the phonological component is an optimal solution to the requirement of relating syntactic input to legible form, a hard question, not yet seriously addressed. We keep here to *narrow syntax*: computation of LF.

The empirical facts make it clear that there are (LF-)uninterpretable inflectional features that enter into agreement relations with interpretable inflectional features. Thus, the φ-features of T (Tense) are uninterpretable and agree with the interpretable φ-features of a nominal that may be local or remote, yielding the surface effect of noun-verb agreement. The obvious conclusion, which we adopt, is that the agreement relation removes the uninterpretable features from the narrow syntax, allowing derivations to converge at LF while remaining intact for the phonological component (with language-variant PF manifestation).

We therefore have a relation Agree holding between α and β, where α has interpretable inflectional features and β has uninterpretable ones, which delete under Agree. The relation Agree and uninterpretable features are prima facie imperfections. In MI and earlier work, it is suggested that both may be part of an optimal solution to minimal design specifications by virtue of their role in establishing the property of "displacement," which has (at least plausible) external motivation in terms of distinct kinds of semantic interpretation and perhaps processing. If so, displacement is only an apparent imperfection of natural language, as are the devices that implement it.

Displacement is implemented by selecting a target P and a related category K to be moved to a position determined by P–P a *probe* that seeks K. The target/probe P determines the kind of category that can be moved to this position (a nominal phrase, a *wh*-phrase, etc.). If uninterpretable inflectional features are the devices that implement displacement, we expect to find uninterpretable features of three kinds:

(2) a. To select a target/probe P and determine what kind of category K it seeks
     b. To determine whether P offers a position for movement
     c. To select the category K that is moved

That seems correct. For movement of a nominal to T, for example, the φ-set and EPP-feature of T serve the functions (2a) and (2b), respectively. The category K that is moved has uninterpretable structural Case, serving the function (2c). Agree is the relation between T and the moved category – more precisely, their relevant subparts. Let us say that the uninterpretable features of P and K render their relevant subparts *active*, so that matching leads to agreement. Locality conditions yield an intervention effect if probe P matches inactive K that is closer to P than matching M, barring Agree(P, M), properties that may be nuanced, as we will see.

The picture seems to generalize over an interesting range. To the extent that this is true, uninterpretable features and the Agree relation are not true "imperfections," despite appearances.

Uninterpretability of features – say, of phonological features, φ-features of T or its EPP-feature, or structural Case – is not "stipulated." The existence of these

features is a question of fact: does L have these properties or not? If it does (as appears to be the case), we have to recognize the fact and seek to explain it: in the best case, by showing that these are only apparent imperfections, part of an optimal solution to design specifications. Though motivated at the interface, interpretability of a feature is an inherent property that is accessible throughout the derivation. The phonological properties [± continuant], for example, are motivated only at the inter-face, but these "abstract" features are accessible throughout the derivation, which ultimately eliminates them in favor of narrow phonetic features interpretable at the interface. Similarly, interpretability of ϕ-features ([+] for N, [−] for T) is accessible throughout the derivation. For convergence, uninterpretable features must be deleted – in narrow syntax, we assume, by the operation Agree, establishing an agreement relation under appropriate conditions.

Suppose that L has generated the syntactic object K with label LB(K). On minimalist assumptions, LB(K) is the only element of K that is immediately accessible to L, so LB(K) must be the element that activates Agree, by virtue of its uninterpret-able features; it is these that constitute the probe that seeks a matching *goal* – another collection of features – within the domain of LB(K). What is the relation Match? The optimal candidate is Identity; we therefore take Match to be Identity.

Interpretability of features is determined in the lexicon, by Universal Grammar (UG) we assume, and the distinction must be indicated not only at that stage but throughout the derivation. The natural principle is that the uninterpretable features, and only these, enter the derivation without values, and are distinguished from interpretable features by virtue of this property. Their values are determined by Agree, at which point the features must be deleted from the narrow syntax (or they will be indistinguishable from interpretable features at LF) but left avail-able for the phonology (since they may have phonetic effects). The conclusion is appropriate in other respects: the values of uninterpretable features are redundant, and there is empirical motivation from intervention effects (see MI). Accordingly, Match is not strictly speaking Identity, but Non-distinctness: same feature, independently of value.

The operation Spell-Out removes LF-uninterpretable material from the syntactic object K and transfers K to the phonological component. It must therefore be able to determine which syntactic features are uninterpretable, hence to be removed. Prior to application of Agree, these are distinguished from interpretable features by lack of specification of value. After application of Agree, the distinction is lost. To operate without reconstructing the derivation, Spell-Out must therefore apply shortly after the uninterpretable features have been assigned values (if they have not been assigned values at this point, the derivation will crash, with uninterpretable features at the interface). Spell-Out must be strongly cyclic (as assumed in MI). The conclusion, which follows naturally from examination of feature interpretability, has other desir-able consequences. In contrast to Extended Standard Theory–based systems [Chomsky 1981], this system has no overt/covert distinction with two independent cycles; rather, it has a single narrow-syntactic cycle. Furthermore, the phonological cycle is not a third independent cycle, but proceeds essentially in parallel.

I will keep here largely to Case/agreement and related systems: ϕ-features, struc-tural Case, EPP, A-movement, and the *core functional categories* T, C, v (T = tense,

C=complementizer, v=a light verb that introduces verbal phrases).[1] Within these systems, probe and goal match if features have values for the goal but not for the probe: if ϕ-features were valued for the probe, it would be inactive and could drive no operation; if they were unvalued for the goal, they would receive no values from the (unvalued) matching features of the probe. If correct, the analysis should generalize to other core syntactic processes. Some extensions to *wh*-movement are suggested in MI, but tentatively, for reasons indicated. This is the easiest case of $\overline{\text{A}}$-movement, since there are grounds to believe that features of probe and goal are involved. In other cases (e.g., topicalization, VP-fronting), postulation of features is much more stipulative; and throughout, questions arise about intermediate stages of successive-cyclic $\overline{\text{A}}$-movement and island conditions.

Matching of probe and goal induces Agree, eliminating uninterpretable features that activate them. A number of questions arise – specifically, with regard to the theses in (3).

(3) a. Goal as well as probe must be active for Agree to apply.
     b. α must have a complete set of ϕ-features (it must be ϕ-*complete*) to delete uninterpretable features of the paired matching element β.

Let us tentatively adopt both theses, returning to the matter.

For the Case/agreement systems, the uninterpretable features are ϕ-features of the probe and structural Case of the goal N. ϕ-features of N are interpretable; hence, N is active only when it has structural Case. Once the Case value is determined, N no longer enters into agreement relations and is "frozen in place" (under (3a)). Structural Case is not a feature of the probes (T, v), but it is assigned a value under agreement, then removed by Spell-Out from the narrow syntax. The value assigned depends on the probe: nominative for T, accusative for v (alternatively ergative-absolutive, with different conditions). Case itself is not matched, but deletes under matching of ϕ-features.

In some cases, an active element E is unable to inactivate a matched element by deleting its unvalued features. E is *defective*, differing in some respect from otherwise identical active elements that induce deletion. The simplest way to express the distinction, requiring no new mechanisms or features, is in terms of (3b): a nondefective probe is ϕ-complete, a defective one is not.

One case, to which we will return, is participle-object constructions, which may manifest partial ϕ-feature agreement but without Case assignment to the object, the participle being defective. Other cases, we will assume, are raising constructions and their exceptional-Case-marking (ECM) counterparts, as shown schematically in (4a), where β is the matrix clause, α is an infinitival with YP a verbal phrase (the case most relevant here), and P is the probe: T with a raising verb (case (4b)), v with an ECM transitive verb (case (4c)).

(4) a. $[_{\beta}\, P\, [_{\alpha}\, [\text{Subj}\, [\text{H YP}]]]]$
     b. i.   there are likely to be awarded several prizes
         ii. several prizes are likely to be awarded
     c. i.   we expect there to be awarded several prizes
         ii. we expect several prizes to be awarded

The Case/agreement properties of Subj in (4a), and its overt location, are determined by properties of the matrix probe P, not internally to α. α is a TP with defective head T_{def}, which is unable to determine Case/agreement but has an EPP-feature, overtly manifested in (4c). Raising-ECM parallels give good reason to believe that the EPP-feature is manifested in (4b) as well, by trace of the matrix subject; preference for Merge over (more complex) Move gives a plausible reason for the surface distinction between [Spec, T_{def}] in (4b) and in (4c) (see MI). In (4bi) and (4ci), the EPP-feature of T_{def} is satisfied by Merge of expletive; in (4bii) and (4cii), by raising of the direct object.

[. . .]

Suppose that the label LB(K) of K has an uninterpretable selectional feature (by definition, an EPP-feature), which requires Merge in [Spec, LB(K)]. That can be satisfied by Merge of an expletive, in which case long-distance agreement may hold between LB(K) and the goal. Alternatively, an active goal G determines a category PP(G) (pied-piping), which is merged in [Spec, LB(K)], yielding the displacement property; long-distance agreement may still appear, if PP(G) fails to satisfy the ϕ-features of the probe (e.g., quirky Case subject). The combination of Agree/pied-piping/Merge is the composite operation Move, preempted where possible by the simpler operations Merge and Agree.

FL specifies the features F that are available to fix each particular language L. The MI framework takes L to be a derivational procedure that maps F to {Exp}, where an expression Exp is a set of interface representations. As a first approximation, take Exp to be {PF, LF}, these being symbolic objects at the sensorimotor and conceptual-intentional interfaces, respectively. We adopt the conventional assumption that L makes a one-time selection [F_L] from F. These are the features that enter into L; others can be disregarded in use of L.

Assume further that L assembles [F_L] to lexical items LI of a lexicon Lex, the LIs then entering into computations as units.

[. . .]

Narrow syntax maps a selection of choices from Lex to LF; the phonological component, in contrast, has further access to [F_L]. Like the extraction of [F_L] from F, these assumptions, largely conventional, reduce the computational burden for the procedure L while adding new conceptual apparatus.

More controversially, MI extends the same reasoning to individual derivations: L makes a one-time selection of a lexical array LA, a collection of LIs (a "numeration" if some are selected more than once), and maps LA to Exp. Again, there is a reduction of computational burden, in this case a vast reduction, since Lex, which virtually exhausts L, need no longer be accessed in the derivation once LA is selected. The new concept LA (numeration) is added, while another concept is eliminated: chains are determined by identity, with no need for indices or some similar device to distinguish chains from repetitions, also violating the Inclusiveness Condition. As in the other cases, the tests are ultimately empirical; on purely conceptual grounds, one could argue either way. As noted, the nature of optimal design that is instantiated in FL (if any) is a matter of discovery, within certain guidelines.

Proceeding further, MI proposes another reduction of computational burden: the derivation of Exp proceeds by *phase*, where each phase is determined by a

subarray LA_i of LA, placed in "active memory." When the computation L exhausts LA_i, forming the syntactic object K, L returns to LA, either extending K to K′ or forming an independent structure M to be assimilated later to K or to some extension of K. Derivation is assumed to be strictly cyclic, but with the phase level of the cycle playing a special role.

A subarray LA_i must be easily identifiable; optimally, it should contain exactly one lexical item that will label the resulting phase. The evidence reviewed in MI suggested that the phases are "propositional": verbal phrases with full argument structure and CP with force indicators, but not TP alone or "weak" verbal configurations lacking external arguments (passive, unaccusative). Assume that substantive categories are selected by functional categories: V by a light verb, T by C. If so, phases are CP and $v*$P, and a subarray contains exactly one C or $v*$.

The choice of phases has independent support: these are reconstruction sites, and they have a degree of phonetic independence (as already noted for CP vs. TP). The same is true of vP constructions generally, not just $v*$P. If these too are phases, then PF and LF integrity correlate more generally.

Suppose, then, we take CP and vP to be phases. Nonetheless, there remains an important distinction between CP/$v*$P phases and others; call the former *strong* phases and the latter *weak*. The strong phases are potential targets for movement; C and $v*$ may have an EPP-feature, which provides a position for XP-movement, and the observation can be generalized to head movement of the kind relevant here.

The special role of strong phases becomes significant in the light of another suggestion of MI that I will adopt and extend here: cyclic Spell-Out, necessary for reasons already discussed, takes place at the strong phase level. The intuitive idea, to be sharpened, is that features deleted within the cyclic computation remain until the strong phase level, at which point the whole phase is "handed over" to the phonological component. The deleted features then disappear from the narrow syntax, allowing convergence at LF, but they may have phonetic effects.

Spell-Out seeks formal features that are uninterpretable but have been assigned values (checked); these are removed from the narrow syntax as the syntactic object is transferred to the phonology. The valued uninterpretable features can be detected with only limited inspection of the derivation if earlier stages of the cycle can be "forgotten" – in phase terms, if earlier phases need not be inspected. The computational burden is further reduced if the phonological component too can "forget" earlier stages of derivation. These results follow from the Phase-Impenetrability Condition (PIC) (MI, (21)), for strong phase HP with head H,

(7)   The domain of H is not accessible to operations outside HP; only H and its
       edge are accessible to such operations.

the *edge* being the residue outside of H′, either specifiers (Specs) or elements adjoined to HP.

H and its edge are accessible only up to the next strong phase, under the PIC: in (8), elements of HP are accessible to operations within the smallest strong ZP phase but not beyond.

(8) $[_{ZP}$ Z . . . $[_{HP}$ α [H YP]]]

Local head movement and successive-cyclic A- and $\overline{A}$-movement are allowed, and both Spell-Out and the phonological component can proceed without checking back to earlier stages. The simplest assumption is that the phonological component spells out elements that undergo no further displacement – the heads of chains – with no need for further specification.

In effect, H and its edge α in (8) belong to ZP for the purposes of Spell-Out, under the PIC. YP is spelled out at the level HP. H and α are spelled out if they remain in situ. Otherwise, their status is determined in the same way at the next strong phase ZP. The question arises only for the edge α, assuming that excorporation is disallowed.

The picture improves further if interpretation/evaluation takes place uniformly at the next higher phase, with Spell-Out just a special case. Assuming so, we adopt the guiding principle (9) for phases Ph_1.

(9) Ph_1 is interpreted/evaluated at the next relevant phase Ph_2.

What are the relevant phases? As noted, because of the availability of EPP, the effects of Spell-Out are determined at the next higher *strong* phase: CP or v^*P. For the same reason, a strong-phase HP allows extraction to its outer edge, so the domain of H can be assumed to be inaccessible to extraction under the PIC: an element to be extracted can be raised to the edge, and the operations of the phonological component can apply to the domain at once, not waiting for the next phase. Keeping to the optimal assumption that all operations are subject to the same conditions, we restate (9) as (10), where Ph_1 is strong and Ph_2 is the next highest strong phase.

(10) Ph_1 is interpreted/evaluated at Ph_2.

On similar grounds, the PIC should fall under (10). We therefore restate the PIC as (11), for (8) with ZP the smallest strong phase.

(11)   The domain of H is not accessible to operations at ZP; only H and its edge are accessible to such operations.

We can henceforth restrict attention to phases that are relevant under (10), that is, the strong phases. For the same reason, we restrict attention to v^* rather than light verb v generally, unless otherwise indicated.

Considerations of semantic-phonetic integrity, and the systematic consequences of phase identification, suggest that the general typology should include among phases nominal categories, perhaps other substantive categories. If categorial features are eliminated from roots, then a plausible typology might be that phases are configurations of the form F-XP, where XP is a substantive root projection, its category determined by the functional element F that selects it. CP falls into place as well if T is taken to be a substantive root, as discussed earlier. Phases are then

(close to) functionally headed XPs. Like TP, NP cannot be extracted, stranding its functional head. The same should be true of other nonphases. Some phases are strong and others weak – with or without the EPP option, respectively, hence relevant or not for Spell-Out and the general principle (10).

Note

For expository purposes, I take the nominal with structural Case to be N and use T and C as cover terms for a richer array of functional categories, as in MI. On light verbs, see among others Hale and Keyser 1993, Harley 1995. Call *v* with full argument structure *v**: transitive *v* or experiencer. Only *v* was considered in MI, and I will largely keep to it below. Where assigned by V, not *v*, Case is inherent. Quirky Case largely falls under general Case assignment principles if understood to be inherent Case with an additional structural Case feature (as in MI). We will return to examples.

30.3   Questions pertaining to Chomsky (2001)

1   French existential sentences such as *Il y a un livre sur la table* ('il there has a book on the table'='there's a book on the table') contain two distinct expletive elements. *Il* is a subject clitic that has something in common with English *it*. *Y* is an object clitic that, apart from its final position, resembles English *there* closely. Discuss the implications for Chomsky's approach to expletives of the fact that some languages have two different expletives co-occurring.

2   How exactly would Chomsky's approach to expletives like *there* have to be modified if it turned out that *there* was not a pure expletive in his sense? Bring in Kayne (2008) and at least some of the references cited in the introduction by Svenonius (2002) to his edited volume.

3   In agreement with Svenonius (2002, Introduction, note 5), Chomsky notes that (colloquial) English really allows expletive *there* only with *be*, as illustrated by *There were several angry men in the room* vs. **There came several angry men into the room*. Chomsky notes in addition that English appears to minimally contrast with Dutch, which does allow such sentences with verbs like *come* and with an expletive *er* that in many respects is very similar to English *there*. Do German and the Scandinavian languages look more like English here or more like Dutch? What might be the parameter(s) at issue?

4   Chomsky's Inclusiveness Condition "bars introduction of new elements (features) in the course of computation." Yet uninterpretable features have their values determined by Agree. Although the uninterpretable feature itself is not introduced in the course of computation, its value is introduced in the course of computation. Should this count as a violation of Inclusiveness? Give your reasons.

5   How central is feature-valuation to Chomsky's probe–goal approach to internal merge? Defend your position.

6 Pairs of English sentences like *John believes he's intelligent* and *John believes himself to be intelligent* are interpretively similar to such an extent that one might readily conclude that the *-self* of the second sentence contributes nothing to its interpretation. Would that conclusion be valid? Why, or why not? (Bring in Helke 1973, Pica and Snyder 1997, and Reuland 2011.)

7 Clitic doubling of the Spanish sort invariably involves agreement in person, and often involves agreement in number, between the clitic and the doubled DP/strong pronoun. Should clitic doubling constitute a subcase of Agree, in Chomsky's sense? Why, or why not? (Extra credit: Could agreement in general, thinking of Kayne and Pollock 2012, constitute a subcase of clitic doubling? Why, or why not?)

8 Discuss the relation between the type of agreement found in *They're my friend*(s)* and the notion of (un)interpretable feature.

9 In his note 3, Chomsky considers the possibility of an independent Agr node, in the spirit of Pollock (1989), noting that "this more complex alternative may be justified, but will be put aside here." Discuss the relation between independent Agr and the analysis of person agreement given in Hale (1973), Rizzi (1982, Ch. 4, section 3.2), and McCloskey and Hale (1983). In addition, discuss the extent to which the lack of gender agreement with finite verbs (even in many languages that have robust gender agreement with adjectives and participles) weakens the naturalness of the class of phi-features. (Extra credit: Does number have more in common with person or with gender, or does number form a natural class with neither? Give your reasons.)

10 Chomsky takes a particular language L to assemble features into lexical items. Discuss the relation between "assemble" and "merge." To what extent do they have similar properties?

11 In his (25), Chomsky adopts the position that "in transitive constructions, something must escape the *v*P." Kayne (2011: 6) suggests that "all arguments must move at least once." Discuss the similarities and the differences between these two formulations. Does either seem correct? Could both be correct? Give your reasons.

12 Chomsky discusses the fact that alongside *There entered the room a large number of men* there is no **How large a number of men did there enter the room?*. Give reasons why this contrast might or might not be related to that seen in *Unfortunately, they arrived late* vs. **How unfortunately did they arrive late?*

13 In what ways is or is not the position of *a large number of men* in the first example of the preceding question similar to the position of the subject in French sentences like *À quelle heure aura lieu la manifestation contre le gouvernement?* ('at what time will-have place the demonstration against the government'='at what time will the demonstration against the government take place')? Bring in Kayne and Pollock (1978, 2001). (Extra credit: Pinpoint the advances in syntactic theory that enabled their 2001 paper to achieve deeper results than their 1978 paper.)

14 Chomsky takes PRO and pro to be accessible to external Merge. Discuss the ways in which this is or is not compatible with Hornstein (1999) and with Kayne (2002, 2006).

15 The PIC (phase impenetrability condition) leads, as Chomsky notes at (42), to the conclusion that in sentences like *Guess what John read*, the object *what* must have initially moved in object shift fashion, before moving up to the CP area. Discuss the consequences for English object shift of the fact that some speakers accept relative clauses like *these people, who John all think should be invited*. Bring in Kimball and Aissen (1971), McCloskey (2000), and Kayne (2003). Be explicit as to how one might account for intra-English variation.

16 Chomsky considers there to be "some reasons to suspect that a substantial core of head- raising processes . . . may fall within the phonological component." How exactly do the following English facts bear on this?: *Why didn't anybody help us?* vs. **We know why anybody didn't help us*? Discuss, in addition, the relation between this positions of Chomsky's, to the effect that head-raising processes can fail to have interpretive impact, and his later (to appear) position that language is "meaning with sound (or some other externalization)."

References

Arad, M. 2003. Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language and Linguistic Theory* 21: 737–778.

Bhatt, R. 2005. Long distance agreement in Hindi-Urdu. *Natural Language and Linguistic Theory* 23: 757–807.

Borer, H. 1984. *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Dordrecht: Foris.

Chomsky, N. 1955/1975. *The Logical Structure of Linguistic Theory*. New York and London: Plenum Press. [Also 1985, University of Chicago Press.]

Chomsky, N. 1973. Conditions on transformations. In S. R. Anderson and P. Kiparsky (eds), *A Festschrift for Morris Halle*, New York: Holt, Rinehart, and Winston, pp. 232–286.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1993. A Minimalist program for linguistic theory. In K. Hale and S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press, pp. 1–52.

Chomsky, N. 2000. Minimalist inquiries: The framework. In R. Martin, D. Michaels, and J. Uriagereka (eds), *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, Cambridge, MA: MIT Press, pp. 89–155.

Chomsky, N. to appear. Problems of projection. *Lingua*.

Fox, D., and D. Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31: 1–46.

Hale, K. L. 1973. Person marking in Walbiri. In S. R. Anderson and P. Kiparsky (eds), *A Festschrift for Morris Halle*, New York: Holt, Rinehart and Winston, pp. 308–344.

Hale, K., and S. J. Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In K. Hale and S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press, pp. 53–109.

Harley, H. 1995. *Subjects, events, and licensing*. Doctoral dissertation, MIT.

Helke, M. 1973. On reflexives in English. *Linguistics* 106: 5–23.

Hornstein, N 1999. Movement and control. *Linguistic Inquiry* 30: 69–96.

Hornstein, N., and A. Weinberg. 1995. The Empty Category Principle. In G. Webelhuth (ed.), *Government and Binding Theory and the Minimalist Program*, Oxford: Blackwell, pp. 241–296.

Huang, C. T. J. 1982. Move WH in a language without WH movement. *Linguistic Review* 1: 369–416.

Kahnemuyipour, A. 2009. *The Syntax of Sentential Stress*. Oxford Studies in Theoretical Linguistics. Oxford and New York: Oxford University Press.

Kaye, J. 1995. Derivations and interfaces. In J. Durand and F. Katamba (eds), *Frontiers of Phonology: Atoms, Structures and Derivations*, London and New York: Longman, pp. 289–332.

Kayne, R. S. 2002. Pronouns and their antecedents. In S. D. Epstein and T. D. Seely (eds), *Derivation and Explanation in the Minimalist Program*, Oxford: Blackwell, pp. 133–166. [Reprinted in Kayne (2005, Ch. 6).]

Kayne, R. S. 2003. Some remarks on agreement and on Heavy-NP-Shift. In M. Ukaji, M. Ike-Uchi, Y. Nishimura, and H. Nakajima (eds), *Current Issues in English Linguistics*, vol. 2 of *Special Publications of the English Linguistic Society of Japan*, Tokyo: Kaitakusha, pp. 67–86. [Reprinted in Kayne (2005, Ch. 11).]

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2006. On parameters and on principles of pronunciation. In H. Broekhuis, N. Corver, R. Huibregts, U. Kleinhenz, and J. Koster (eds), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, Berlin and New York: Mouton de Gruyter, pp. 289–299. [Reprinted in Kayne (2010, Ch. 4).]

Kayne, R. S. 2008. Expletives, datives, and the tension between morphology and syntax. In T. Biberauer (ed.), *The Limits of Syntactic Variation*, Amsterdam and Philadelphia: John Benjamins, pp. 175–217. [Reprinted in Kayne (2010, Ch. 6).]

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2011. Why are there no directionality parameters? In M. Byram Washburn, K. McKinney-Bock, E. Varis, A. Sawyer, and B. Tomaszewicz (eds), *Proceedings of the 28th West Coast Conference on Formal Linguistics*, Somerville, MA: Cascadilla Proceedings Project, pp. 1–23.

Kayne, R. S., and J.-Y. Pollock. 1978. Stylistic inversion, successive cyclicity and move NP in French. *Linguistic Inquiry* 9: 595–621.

Kayne, R. S., and J.-Y. Pollock. 2001. New thoughts on stylistic inversion. In A. C. Hulk and J.-Y. Pollock, *Inversion in Romance*, New York and Oxford: Oxford University Press, pp. 107–162.

Kayne, R. S., and J.-Y. Pollock. 2012. Toward an analysis of French hyper-complex inversion. In L. Brugè, A. Cardinaletti, G. Giusti, N. Munaro, and C. Poletto (eds), *Functional Heads: The Cartography of Syntactic Structures* 7, New York and Oxford: Oxford University Press, pp. 150-167. [Also in Kayne (2010).]

Kimball, J., and J. Aissen. 1971. I think, you think, he think. *Linguistic Inquiry* 2: 242–246.

Kratzer, A., and E. Selkirk. 2007. Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24: 93–135.

Marantz, A. 2007. Phases and words. Ms., New York University.

Marantz, A. 2012. Locality domains for contextual allosemy in words. Ms., New York University.

McCloskey, J. 2000. Quantifier float and Wh-movement in an Irish English. *Linguistic Inquiry* 31: 57–84.

McCloskey, J., and K. Hale. 1983. On the syntax of person-number inflection in Modern Irish. *Natural Language and Linguistic Theory* 1: 487–533.

Müller, G. 2011. *Constraints on Displacement: A Phase-based Approach*. Amsterdam and Philadelphia: John Benjamins.

Newell, H. 2008. *Aspects of the morphology and phonology of phases*. Doctoral dissertation, McGill University.

Pica, P., and W. Snyder. 1997. On the syntax and semantics of local anaphors in French and English. In A. M. Di Sciullo (ed.), *Projections and Interface Conditions. Essays on Modularity*, New York and Oxford: Oxford University Press, pp. 235–250.

Polinsky, M., and E. Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19: 583–646.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Reuland, E. J. 2011. *Anaphora and Language Design*. Cambridge, MA: MIT Press.

Rizzi, L. 1980. Violations of the Wh island constraint and the subjacency condition. *Journal of Italian Linguistics* 5: 157–195. [Reprinted in Rizzi (1982, Ch. 2).]

Rizzi, L. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Rizzi, L. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Rudin, C. 1988. On multiple questions and multiple Wh fronting. *Natural Language and Linguistic Theory* 6: 445–501.

Svenonius, P. (ed.), 2002. *Subjects, Expletives, and the EPP*. New York and Oxford: Oxford University Press.

Travis, L. 1989. Parameters of phrase structure. In M. Baltin and A. Kroch (eds), *Alternative Conceptions of Phrase Stucture*, Chicago and London: University of Chicago Press, pp. 263–279.

Vergnaud, J.-R. 1977. Letter to Noam Chomsky and Howard Lasnik on "Filters and Control," April 17, 1977. [Published 2006 in R. Freidin and H. Lasnik (eds), *Syntax: Critical Concepts in Linguistics*, Vol. 5, New York: Routledge, pp. 21–34.]

31

IP-Internal Topic and Focus Phrases

K. A. Jayaseelan

2001

31.1 Introduction

A detailed comparison of different languages often allows us to reach a deeper understanding of their syntactic properties. In a certain line of research, this has led to postulating a richer articulation of clausal structure. For example, a minute comparison of the placement of the verb in English and French led Pollock (1989) to the proposal that what was taken to be the head of the clause, Infl, consists of (not one, but) two functional projections, T and AGR (cf. also Belletti 1990). Applying the comparative approach to languages that are minimally different from one another, Poletto (2000) examined one hundred varieties of Romance spoken in Northern Italy and proposed a highly articulated structure for the portion of the clause in which subject clitics occur. Extending this line of work, in *IP-internal Topic and Focus Phrases* Jayaseelan compares Malayalam and some Germanic languages, and proposes an articulated functional structure for the portion of the clause lower than the inflectional domain and higher than the verb phrase.

Malayalam's unmarked word order is SOV. In *wh*-questions, the *wh*-phrase must immediately precede the verb; if the subject or object are present, they must in turn precede the *wh*-phrase:

(1) a. awan ewiTe pooyi?
         he   where went
         'Where did he go?'
     b. ninn-e   aarə aTiccu?
         you–ACC who beat.PAST
         'Who beat you?'

Jayaseelan argues that this order is derived from movement, and postulates the existence of functional structure right above *v*P. In his analysis, starting from an underlying Spec–head–complement order (Kayne 1994), the *wh*-phrase moves into the specifier of a Focus Phrase, which is postulated to be immediately above *v*P. The object and the subject move past it, to structurally higher positions. More precisely, the subject moves to the specifier of TP. The object may move either to its canonical position, argued to be a functional projection immediately higher than the Focus Phrase; alternatively, if it is a definite object (old information), it moves to the specifier of a Topic Phrase, postulated to be structurally higher than FocusP and of the canonical object position (cf. the phrase structure in (15) in the article).

Evidence for the existence of two possible positions for the object comes from sentences with ditransitive predicates, like those in (2), which show that each position gives rise to a different reading for the object:

(2) a. ñaan oru maratt-inə weLLam ozhiccu
         I   a   tree–DAT   water   poured
         'I poured water to a tree.'
     b. ñaan weLLam oru maratt-inə ozhiccu
         I   water a   tree–DAT   poured
         'I poured the water to a tree.'

The canonical word order of sentences with ditransitive verbs in Malayalam (as in other SOV languages) exhibits the indirect object preceding the direct object, which in turn precedes the verb (i.e., IO > DO > V). When the direct object is in its canonical position, as in (2a), it has an indefinite reading ('some water'). In contrast, when it is not, as in (2b), it has a definite reading ('the water'). The definite reading of the object in (2b) is taken to be evidence that it has moved to the specifier of a Topic Phrase, which is higher than the projection to which the object moves in its canonical position. Jayaseelan draws parallels with Germanic languages like Dutch, German, and Yiddish, which have a low focus position (between the V-final verb cluster and certain VP-adverbs), and a higher position for nonfocalized, specific objects (Diesing 1992, 1997b).

In arguing for the existence of functional projections for focused and topical-ized constituents in the middle portion of the clause, Jayaseelan proposes that the *v*P has a left periphery that is partly analogous to the articulated left periphery of the clause proposed in Rizzi (1997), echoing ideas also found in Villalba (1999b) and Cecchetto (1999). The left periphery of the *v*P lacks the highest and the lowest projections of the CP (Force and Fin); the Topic projections are not immediately above the Focus phrase, but are separated from it by an adverb projection and, in SOV languages, the canonical positions of arguments. Obligatory movement of VP-internal arguments to their canonical positions is seen as the difference between SOV languages like German and Malayalam from SVO languages like English and Yiddish.

Jayaseelan notes that the idea of a left periphery at the edge of both *v*P and CP should be related to Chomsky's (2001) proposal that (transitive) little *v* and C are both phase heads. In this connection, it is interesting to think about Aboh's (2004b)

work, which suggests that D is the head of a phase in the nominal domain and proposes the existence of topic and focus positions in its left-periphery, providing further support to the idea that there might be a connection between the domain of locality expressed by the notion of phase and the presence of positions for focused and topicalized constituents.

31.2   From "IP-INTERNAL TOPIC AND FOCUS PHRASES"

1   Introduction

In this paper I present a series of arguments for postulating a functional projection of Focus above vP. I also postulate an iterable Topic Phrase above this Focus Phrase. The postulation of IP-internal Topic/Focus projections will be shown to lead the way to a new view of the difference between the clause structures of SOV and SVO languages, and to some interesting results about clause-internal scrambling and object shift in such diverse languages as Malayalam, German, Dutch, Yiddish and Scandinavian.

2   Question words contiguous to V

Many languages have a requirement that a question word should be contiguous to V. In Malayalam, although the natural way to ask a question is by clefting, a non-cleft question is possible under one fairly strict condition: the question word must be placed immediately to the left of V in a position 'normally' occupied by the direct object if one is present, Malayalam being an SOV language:
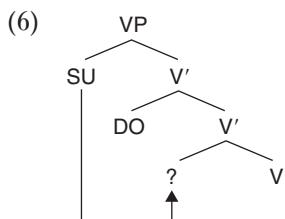
(1)   a.   ni<u>nn</u>-e   aarə aTiccu ?
             you-acc. who beat-past
             'Who beat you?'
      b.   * aarə ni<u>nn</u>-e aTiccu ?
(2)   a.   iwiTe aarə uNTə ?
             here   who is
             'Who is here?'
      b.   * aarə iwiTe uNTə ?
(3)   a.   awa<u>n</u> ewiTe pooyi ?
             he   where went
             'Where did he go?'
      b.   * ewiTe awa<u>n</u> pooyi ?
(4)   a.   nii aa pustakam aar-kkə koDuttu ?
             you that book   who-dat. gave
             'To whom did you give that book?'
      b.   * nii aar-kkə aa pustakam koDuttu ?

Even clefting in questions, one can now see, is possibly a device for positioning the question word next to V:
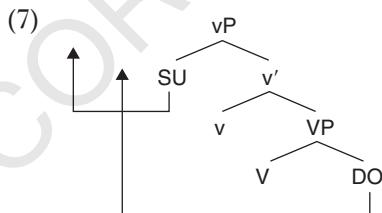
(5)   nii entə aaNə tinn–atə ?
      you what is   ate–nominalizer
      'What did you eat?' (Lit. 'What is it that you ate?')

In a cleft construction, the main verb is the copula; and the question word comes immediately to the left of the copula.[1]
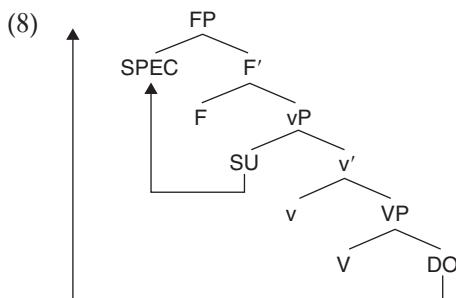
How do we generate this position of the question word? Starting from an underlying SOV word order of the type traditionally assumed in South Asian linguistics, it is difficult to see how one can generate a COMP-like position 'within VP'. Equally impossible are the 'downward' movements we would need to postulate, to move, say, the subject into this position, cf. (1a); this is illustrated in (6):

(6)

However, we can avoid these problems if we assume a universal Spec–Head–Complement order (Kayne 1994); and say that the surface order of the verb's internal arguments in SOV languages is the result of the raising of these arguments into SPECs of higher functional projections. While the subject raises to SPEC,IP, the internal arguments raise to SPECs of functional projections which are intermediate between IP and VP. In the case of a monotransitive verb for instance, the two movements shown in (7) would be the 'normal' movements of an SOV language. ((7) is anticipated, for Dutch, in Zwart (1993).)

(7)

Given this picture, all we need to do, in order to generate the question word's position next to V, is to postulate a Focus Phrase (FP) immediately dominating vP, and to say that the Q-word moves into the SPEC of this FP. All other arguments, and such adjuncts as are generated within vP, e.g. manner, location, time adverbials, would now move 'past' this position into SPECs of higher functional projections by the normal movements which derive the SOV word order. In the case of (1a) for example, the subject is a Q-word and moves into SPEC, FP and the direct object moves 'past' it, as shown in (8):

(8)
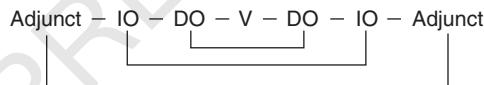
V adjoins to v; there is reason to think that $[_v$ V-v] adjoins to Focus.

[…]

3   Scrambling in Malayalam and the structure of the Malayalam clause

In what is usually taken to be 'the VP' of SOV languages, the canonical order of elements is: Adjunct – IO – DO – V; cf. (10):

(10)   ñaa<u>n</u> innale   Mary -k'k'ə oru kattə ayaccu
       I   yesterday   -dat. a   letter sent
       'I sent a letter to Mary yesterday.'

As a comparison of this sentence with its English gloss shows, the order of elements is the mirror-image of English:

The movements out of the VP that we postulated for SOV languages are apparently 'nested' movements. Interesting questions arise about Relativized Minimality. How do these movements escape minimality effects? There are two sub-questions. One, if SPEC, FP is filled, how do these movements go past it – or (indeed) past SPEC, vP (the 'VP-internal subject')? Two, why are there no inter se minimality effects among them; e.g. why doesn't the landing site of the direct object prevent the indirect object moving to a higher position? The problem of course is that the Malayalam V does not raise any higher than the head of FP, as we just said.

It has been recognized however that we need to postulate two types of movements: one, instantiated by Icelandic object shift, which obeys minimality; the other, instantiated by scrambling in Dutch, which does not obey this constraint (Zwart 1993, Diesing [1997b]). While the reason for this distinction remains puzzling – especially since both types of movement have many things

in common; e.g. they obey a common 'definiteness/specificity' constraint – let us for the time being simply say that the migration of arguments and adjuncts out of the VP in SOV languages is a case of scrambling.
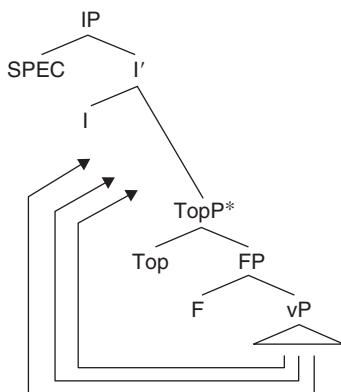
What are the functional heads, higher than FP and lower than IP, which host these moved phrases? It has been claimed for the COMP system (Rizzi 1997) that there are any number of Topic Phrases possible above the FP in COMP; assuming a similar possibility with respect to the FP above vP, it is tempting to say that the 'normal' movements of the internal arguments (and adjuncts) of SOV languages are to SPEC, TopP. This solution has a seeming advantage: repeated applications of Topicalization should be able to produce any order of the elements that undergo the operation, i.e. the base order of these elements can be arbitrarily reordered. This should be able to generate 'scrambling' understood as the free order of a verb's arguments (which is the 'classical' view of scrambling). This advantage however is outweighed by other considerations. Firstly, it cannot account for the canonical order of the verb's internal arguments, the one which we tried to describe in terms of 'nested' movements. But the really serious problem is that the internal arguments in their canonical order – as, for example, in (10) – do not show any topicalization effects. Topics are familiar information in the discourse; they are entities which have already been mentioned, and are therefore definite or specific. In fact, we shall be arguing that the leftward movements showing a definiteness/specificity effect in Scandinavian, Dutch or Yiddish are instances of topicalization – specifically, of movement into TopPs above FP. But there are no definiteness/specificity constraints on the Malayalam verb's internal arguments in their canonical order; cf.

(11)   en̲–ik'k'ə oru aana̲–ye   weeNam
        I-dat.   an elephant–acc. want
        'I want an elephant.'
(12)   nii   puuwə paRik'k'-arutə
        you flower pick–should not
        'You should not pick flowers.'
(13)   awan̲ oru maNDan̲ aaNə
        he an idiot   is
        'He is an idiot.'
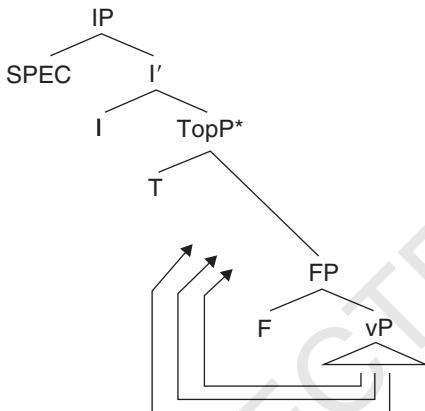
It is difficult to imagine why the indefinite NPs in (11)–(13) should be topicalized.

Our 'nested' movements then are *not* into Topic Phrases. However, suppose we do postulate iterable TopPs above FP in Malayalam also – on the evidence of the European languages, to which we come back presently. Do we say that the 'nested' movements that we are postulating are into positions higher than TopP*, or lower than TopP*? I.e., is (15) the correct picture?

(14)

```
          IP
       /      \
    SPEC       I′
            /
          I
          |
          |
              TopP*
            /     \
          Top      FP
                  /   \
                 F     vP
                       /\
```

(15)

```
          IP
       /      \
    SPEC       I′
            /    \
          I       TopP*
                 /
               T
                    \
                     FP
                    /  \
                   F    vP
                        /\
```

[…]

However, if one of the arguments is indefinite, we get some interesting results, cf.

(19) a. ñaa<u>n</u> awa<u>n</u>-ə oru kattə ayaccu
        I   he-dat. a   letter sent
     b. ?*ñaa<u>n</u> oru kattə awa<u>n</u>-ə ayaccu
        I   a   letter he-dat. sent
        'I sent him a letter.'

(20) a. ñaa<u>n</u> awaL-kkə paNam koDuttu
        I   she-dat.   money gave
     b. ?*ñaa<u>n</u> paNam awaL-kkə koDuttu
        I   money she-dat.   gave
        'I gave her money.'

(21) a. ñaa<u>n</u> ni<u>nn</u>-oo Də oru tamaaəa paRay-aam
        I   you-2nd dat. a joke   say-will
     b. ?*ñaa<u>n</u> oru <u>t</u>amaaəa ni<u>nn</u>-oo Də paRay-aam
        I   a joke   you-2nd dat. say-will
        'I'll tell you a joke.'

If the interchange of positions in the (b) sentences is due to IO moving into a TopP below its canonical position, it is difficult to see why these sentences are unacceptable, since a definite pronoun is always amenable to topicalization. On the other hand, if what is happening in the (b) sentences is the movement of DO into a TopP above its canonical position, the ungrammaticality of these sentences is explained: an indefinite (non-specific) NP has been (illicitly) topicalized. These data then support (15) over (14). As a matter of fact, if the IO is indefinite and the DO definite – the reverse of what is the case in (19)–(21) – the canonical order is somewhat awkward! This is especially so, if the DO is a pronoun, cf.

(22) a. ?? ñaa<u>n</u> oru bhikshakkaara<u>n</u>-ə atə koDuttu
              I   a   beggar-dat.   it   gave
       b. ñaa<u>n</u> atə oru bhikshakkaara<u>n</u>-ə koDuttu
              I   it   a   beggar-dat.   gave
           'I gave it to a beggar.'

If what is happening in (b) is the IO being topicalized in a position lower than its canonical position, the acceptability of this sentence is puzzling – since an indefinite NP is being topicalized. But if the (definite) DO is being topicalized in a position higher than its canonical position, the complete acceptability of the (b) sentence is unsurprising.

A very interesting pair of sentences which helps us to choose between (15) and (14) is the following:

(23) a. ñaa<u>n</u> oru maratt-i<u>n</u>ə weLLam ozhiccu
              I   a   tree-dat.   water   poured
           'I poured water to a tree.'
       b. ñaa<u>n</u> weLLam oru maratt-i<u>n</u>ə ozhiccu
              I   water   a   tree-dat.   poured
           'I poured the water to a tree.'

In Malayalam, the definite article is null. This means that in itself, a form like *weLLam* ('water') is ambiguous between a definite and an indefinite reading. In the (a) sentence, which has the canonical order, the most natural interpretation of *weLLam* is as '(some) water'; i.e. the argument is indefinite. But in the (b) sentence, which has the inverse order, the only permissible interpretation of *weLLam* is as 'the water'; i.e. the argument is obligatorily definite. The (a) sentence could be an answer to the question 'What did you do?' The (b) sentence could only be an answer to the question 'What did you do with the water?' This definiteness constraint on *weLLam* in the (b) sentence is explained if it is a Topic.

[…]

However recall that in Rizzi's (1997) articulation of the COMP system of IP, there are TopPs both above *and below* FP. Is there any evidence of TopPs below FP in the COMP system of vP also? Consider the following sentences:

(24)   aarum   kaND-illa, aana-ye
        nobody saw-neg.   elephant-acc.
        'The elephant, nobody saw.'

(25)   aarə ayaccu, ninn-e ?
        who sent   you-acc.
        'You, who sent?'

If we assume a TopP below FP, we can readily explain the post-verbal elements in these sentences. We can say that these elements are in this TopP; and that furthermore *aarum* 'nobody' (a negative polarity item) in (24), and *aarə* 'who' (a question word) in (25) are in SPEC, FP, and that V has raised and adjoined to F.

   […]

   However, the fact that the below-FP Topic position invariably appears in linear terms post-verbally, suggests also another analysis of these data. Consider (29), where (29b) = (24):

(29)   a.   aana-ye   aarum   kaND-illa
             elephant–acc. nobody saw-neg.
             'The elephant, nobody saw.'
        b.   aarum   kaND-illa, aana-ye
             nobody saw-neg.   elephant-acc.

In (a), *aana-ye* 'elephant-acc.' is plausibly in a pre-IP Topic position. The (b) sentence, we could suggest, is derived from the (a) sentence by preposing IP to the SPEC of a still higher functional head (say, a higher TopP). […] In fact, the preposing of IP and VP must be assumed to take place quite generally in Malayalam, as a result of the familiar property of SOV-language verbs of moving their arguments to the left (the 'nested' movements to canonical positions that we spoke of) – but extended now (at least in Malayalam) to auxiliary verbs and the verbal complementizer. First, note that the auxiliary verbs are stacked on the right-hand side of the lexical verb, in the inverse order_of English; e.g.

(30)   awan atə tinn-iTTuND-aakaam
        he   it eat-perf.-may
        'He may have eaten it.'

The VPs headed by the auxiliary verbs, we may assume, are generated still higher than the Topic and Focus positions – and the 'sandwiched' canonical positions – that we investigated. We can reasonably claim that the auxiliary Vs share with the main V the property that their arguments vacate their base positions which are to the right of V. In the case of each auxiliary V, its complement raises, possibly to its own SPEC position. Repeated applications of this movement give us the inverse order of the stacked auxiliary verbs at the end of the main V. The Malayalam complementizer *ennə*, which occurs at the right edge of the embedded clause, is also a V: historically, it is a non-finite form of a verb meaning 'say'. We may assume that it is generated as the head of the Finiteness Phrase in the Malayalam COMP system; and that it induces its complement IP to move to its left.

4   The 'middle field' of Dutch and German

Our clause structure of Malayalam has interesting explanations for some word-order phenomena of the 'middle field' of the Dutch and German sentence. In the unmarked order, the indirect object precedes the direct object in German. The reverse order can be produced by scrambling, but this is subject to the following definiteness condition (Lenerz 1977:54, cited in Abraham 1986:17):

(31)   * DO [−def] + IO [+/− def]

Consider (32) below. The (a) sentence gives the neutral order; the (b) sentence is acceptable; but the (c) sentence is ungrammatical (Abraham 1986:18):

(32)   a.   ich habe meinem Bruder einen/den Brief geschickt
            I   have my-dat. brother a/the   letter sent
            'I have sent my brother a/the letter.'
       b.   ich habe den Brief meinem Bruder geschickt
            I   have the letter my-dat.   brother sent
       c.   *ich habe einen Brief meinem Bruder geschickt
            I   have a   letter my-dat. brother sent

We have seen the same definiteness condition in operation in the scrambled DO–IO order in Malayalam. The same explanation should carry over. German and Dutch, we can say, are 'SOV languages' in the same sense in which Malayalam is an SOV language: all the arguments and adjuncts of a V-initial VP move out of the VP (by nested movements) to a set of 'canonical positions' – which are higher than a Focus Phrase but below a set of Topic Phrase(s). The scrambled order of (32b) and (32c) is generated by the direct object being topicalized. Lenerz's definiteness condition can now be explained: an indefinite NP which receives an existential interpretation is necessarily 'new information' and therefore cannot be a Topic.

There is apparently an adverb position in German and Dutch immediately below the TopP, and above the canonical positions. In traditional generative analyses of German or Dutch as an SOV language, this adverb position was taken as adjoined to VP. The adverb was used as a diagnostic of scrambling: any phrase which was to the left of the adverb was taken as having moved out of VP. Interestingly it has been noticed that the position to the left of the adverb shows a definiteness effect – a fact which is unexplained if scrambling is a purely optional movement. Thus Diesing ([1997b]) observes that definite NP objects in German are under 'pressure' to scramble (examples from Diesing ([1997b]:378, 380); 'M' means 'marked'):

(33)   a.   ᴹ … weil ich selten die Katze streichle
                  since I   seldom the cat   pet
       b.   … weil ich die Katze selten   streichle
                since I   the cat   seldom pet
            '… since I seldom pet the cat.'

And pronominal objects *must* scramble:

(34) a. *… weil ich selten sie streichle
since I seldom her pet
b. … weil ich sie selten streichle
since I her seldom pet
'… since I seldom pet her.'

In our terms, these facts indicate that there is a preference for topicalizing definite NP objects in German; and in the case of a pronoun topicalization is obligatory. An indefinite NP object which has scrambled cannot have an existential interpretation, but acquires a specific reading (as observed by Diesing [1997b], for German; Zwart 1996, De Hoop 1992, for Dutch); cf. the following Dutch examples (Zwart 1996:91):

(35) a. … dat Jan gisteren een meisje gekust heeft
that John yesterday a girl kissed has
'… that John kissed a girl yesterday.'
b. … dat Jan een meisje gisteren gekust heeft
that John a girl yesterday kissed has
'… that John kissed a (particular) girl yesterday.'

Again, this fact follows from our claim that a scrambled phrase is topicalized.

There is a well-known alternative account of this definiteness/specificity effect of scrambling, offered by Diesing (1992, [1997b]) and De Hoop (1992). This account says that VP is the domain of existential closure; but a definite NP – which introduces a free variable in the semantic representation – must *not* be existentially interpreted and so must move out of VP in order to 'escape' existential closure (see Diesing [1997b]:378–379). However, an indefinite NP may either remain in the VP and get an existential interpretation; or it may scramble out of VP and get other types of interpretation, e.g. a specific interpretation (Diesing [1997b]:377). Observe that this explanation hinges on the assumption that the NP object in its canonical position (in German or Dutch) is within VP. It is obviously incompatible with German or Dutch being an SVO language.

[…]
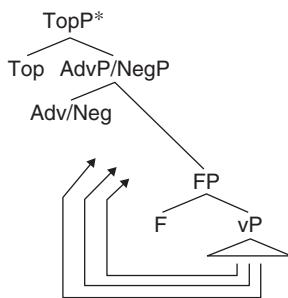
Our explanation of German and Dutch scrambling in terms of a movement to Topic is different from either of these other accounts in the following respects. Unlike the Diesing/De Hoop account, it is consistent with these languages being underlyingly SVO.

[…]
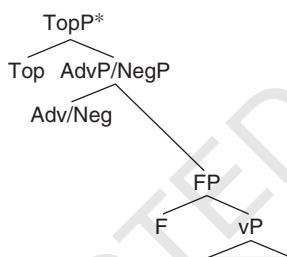
Contrasting two closely related languages – German (an SOV language) and Yiddish (an SVO language) – gives us a very interesting confirmation of the structures we have postulated. Since we assume a universal SPEC-Head-Complement order and therefore a V-initial VP in all languages, for us the difference between SOV and SVO languages is crucially the following: SOV languages generate the 'canonical' positions – above Focus Phrase and below

Topic Phrase(s) and an adverb/neg position – into which they move by 'nested' movements all the (unmarked) arguments and adjuncts in the VP. SVO languages do not make use of this option. The two contrasting structures we have argued for are the following:

(38)   (German)

(39)   (Yiddish)

(38), which may be compared with the lower part of (15) which we adopted for Malayalam,[2] differs from (39) in having the 'canonical positions' indicated by a dotted line, into which the 'nested' movements go

    Now consider a definite DP which occurs to the right of an adverb (and to the left of the verb) in German. We know that in German, a definite DP does not normally stay in its canonical position because there is a strong preference for it to be topicalized – i.e. it normally occurs to the left of the adverb position. Therefore, if it is to the right of an adverb, it could only be because it is contrastively (or otherwise) focused, i.e. it is in Focus Phrase. This prediction appears to be correct, cf. (33a), repeated below, which is perceived as exhibiting a marked order and is acceptable only if *die Katze* 'the cat' is focused:

(33)   a.   ᴹ… weil   ich selten   die Katze streichle
                    since I   seldom the cat   pet
           '… since I seldom pet the CAT.'

An indefinite NP is under no pressure to scramble. Therefore an indefinite NP in the same position, i.e. to the right of an adverb and to the left of the verb, is most probably in its canonical position; here it is existentially interpreted. However it could also be in the Focus Phrase, in which case it will bear contrastive stress.

Now consider Yiddish. The Yiddish indefinite NP is normally placed to the right of the verb, which is its base position. But the definite DP must move to the left of the verb – in fact, to the left of the adverb position, i.e. it must normally be topicalized. Cf. (40) (examples from Diesing [1997b]):

(40) a. Maks hot geleyent a bukh
          Max   has read   a book
          'Max has read a book.'
      b. Maks hot dos bukh (nekhtn) geleyent
          Max   has the book yesterday read
          'Max has read the book (yesterday).'

Now consider the sentences in (41) which have an NP between an adverb and the verb (Diesing's examples):

(41) a. MMaks hot nekhtn   a bukh geleyent
          Max   has yesterday a book read
          'Max read a BOOK yesterday.'
      b. MMaks hot nekhtn   dos bukh geleyent
          Max   has yesterday the book read
          'Max read the BOOK yesterday.'

Given the structure in (39), the NP in question can only be in the Focus Phrase. This prediction is correct, because both the definite *and* the indefinite NP are obligatorily interpreted as contrastively focused, as shown in the gloss. As Diesing points out, the interesting contrast is with the German indefinite NP in the same position, which has an unmarked (non-contrastive), existential interpretation. This contrast is explained by the presence of the 'canonical positions' above vP in German, but not in Yiddish.

Diesing notes that the position between the adverb/neg and the verb in Yiddish can accommodate only one constituent, but that the position to the left of both the adverb/neg and the verb can accommodate more than one constituent. Cf.

(42) *Nekhtn   hot Maks nit dem yingl dos bukh gegebn
       yesterday had Max   not the boy   the book given
       'Max didn't give the boy the book yesterday.'

(43) Nekhtn   hot Maks dem yingl dos bukh nit gegebn
       yesterday had Max   the boy   the book not given
       'Max didn't give the boy the book yesterday.'

These facts again are predicted by our postulated structures, which have only one Focus Phrase but iterable Topic Phrases. Thus we see that our structures (38) and (39), for SOV and SVO languages respectively, make just the right predictions in every case. We take this to be strong confirmation of our analysis.[3]

Notes

1 Other Dravidian languages – and indeed many other languages of the Indian subcontinent – also tend to place their question words to the immediate left of V. But in their case, this positioning is apparently not a strict requirement like it is, in Malayalam. The parametric difference can perhaps be stated in terms of strong/weak features: Malayalam question words have a strong focus feature (see our analysis below), whereas the question words of the other languages have a focus feature which is optionally strong.

2 A small difference is that (15) shows no AdvP/NegP. Neg in Malayalam is a finite auxiliary verb which is clause-final; one must suppose that it heads an auxiliary VP immediately under TP (Tense Phrase). Regarding adverbs, Hany Babu (p.c.) has pointed out to me that Malayalam adverbs seem to behave pretty much like the German adverbs. Cf.

   (i) a. ñaan ennum   oru bhikshakkaaran-ə oru ruupa koDukk-um
            I   every day a   beggar-dat.   a   rupee give-Fut.
            'I give a rupee to a beggar every day.'
        b. ñaan oru bhikshakkaaran-ə ennum   oru ruupa koDukk-um
            I   a   beggar–dat.   every day a   rupee give-Fut.
        c. ñaan oru ruupa ennum   oru bhikshakkaaran-ə koDukk-um
            I   a   rupee every day a   beggar-dat.   give-Fut.

   In (ia), both *a beggar* and *a rupee* have most naturally a non-specific reading. But (ib) means that I give a particular beggar a rupee every day; and (ic) has the odd reading that I give a particular rupee coin (or note) repeatedly to a beggar every day. Thus, we could very well have indicated an AdvP in (15), in the same position relative to TopP* and the canonical positions as in German.

3 Further (and very interesting) confirmation is provided by Modern Persian, a SOV language, in which specific and non-specific DOs occupy different positions in the surface syntax (Karimi 1999; all examples below are from this source). A specific (definite or indefinite) DO precedes IO, and is invariably followed by a special marker *râ*; a non-specific DO follows IO:

   (i) a. Kimea un   ketâb ro   barâ man xarid
            Kimea that book RÂ   for me bought
            'Kimea bought that book for me.'
        b. Kimea barâ man (ye) ketâb xarid
            Kimea for me (a) book bought
            'Kimea bought (a) book for me.'

   A specific DO can license a parasitic gap, a non-specific DO cannot:

   (ii) a. Kimea [$_{NP}$ ye kârgar ro]$_i$ [$_{CP}$ ghablaz inke pro e_i estexdâm
            Kimea   a worker RÂ   before that   hiring
            be-kon-e]   be kâr vâdâsht
            SUBJ-do-3SG   to work forced

'Kimea forced a (specific) worker to work before hiring (her).'

   b.   * Kimea [$_{NP}$ ye kârgar]$_i$ [$_{CP}$ ghablaz inke pro e_i estexdâm be-kon-e] be kâr vâdâsht

We can readily fit these facts into our analysis if we say that *râ* is a Topic marker generated as the head of a Topic Phrase; and that a specific DO in Modern Persian is obligatorily topicalized. Karimi (1999) posits different phrase structure rules to base-generate specific and non-specific DOs in different positions. This move which violates a very widely-accepted 'Uniformity of Theta Assignment Hypothesis' (Baker 1988) can be dispensed with, given our analysis.

31.3   Questions pertaining to Jayaseelan (2001)

1   Given the familiar prohibition against phrasal lowering, what kind of sentences provide direct motivation for Jayaseelan's proposal that the Focus Phrase needed for Malayalam must be located above the merge position of the (agentive) subject?

2   With respect to focus and topic, in what ways is Malayalam different from and in what ways similar to Hungarian as discussed, for example, by Brody (1990) and É. Kiss (2002)?

3   Malayalam requires the interrogative *wh*-word/phrase to occur immediately preceding the verb. How exactly does English differ from Malayalam in this respect?

4   Jayaseelan takes the canonical VP-internal order of arguments and adjuncts to be the mirror-image of that found in English. To what extent is this similar to the phenomena discussed in Cinque (2005)?

5   Give two ways in which Jayaseelan's approach to "scrambling" differs from Diesing's (1992).

6   Compare Romance right-dislocation, as discussed by Cecchetto (1999) or Villalba (1999a), with Jayaseelan's analysis of postverbal arguments in Malayalam. (Extra credit: What points in common are there (or not) between all of these and French stylistic inversion, as discussed by Kayne and Pollock 1978, 2001?)

7   Jayaseelan notes that in Malayalam adverbial elements such as *here* and *yesterday* occur postverbally more freely than do arguments. How might this be related to Rizzi's (1997) discussion of differences between pre-IP adverbials and pre-IP arguments?

8   Jayaseelan notes that the Malayalam complementizer *ennə* is historically related to the verb for *say* (p. 00). How might one try to extend to *ennə* the (synchronic) analysis proposed for Abe by Koopman and Sportiche (1989)? What would such an analysis have in common with Kayne (to appear) on *this* and *that*?

9   In what way is the often-assumed link between universal Spec-Head-Complement order and a possibly universal V-initial VP affected by Larson (1988, 1990)?

10 Jayaseelan takes VO languages such as English or Yiddish to differ from OV languages such as Malayalam or German in that VO languages do not have recourse to the systematic argument movements out of VP that characterize OV languages. Compare this view of English with that of Johnson (1991). Discuss in addition the possible validity of the following conjecture: No argument (in any language) can fail to move at least once.

11 In his note 24, Jayaseelan takes Malayalam negation to be a finite auxiliary verb. What considerations might bear on the choice between this proposal and an alternative one that would take Malayalam negation not to be an auxiliary verb, but to be accompanied by a silent auxiliary verb? Would Malayalam negation then be closer to French *ne* or to French *pas* as discussed by Pollock (1989, Ch. 15 of this volume)? (Extra credit: On the "final" character of Malayalam negation, bring in Biberauer 2008.)

12 What are the similarities and differences between Jayaseelan's analysis of Modern Persian *râ* as a topic marker and Aboh's (2004a) analysis of Gungbe?

13 In his note 27, Jayaseelan notes that some languages, such as Yiddish, allow only one Focus position overtly above vP (see his (42), based on Diesing 1997a). How is Yiddish in this respect similar to, yet different from, languages like Bambara or, in the case of infinitival clauses, from some Irish?

14 Discuss the differences and similarities between Malayalam clefts as used in the most natural way to ask a Malayalam question and the apparently fairly similar property of French, as in Munaro and Pollock (2005).

15 Discuss the tension between Jayaseelan's analysis of pseudo-gapping and Kuno's (1975) observation that pseudo-gapping sentences in English are not appropriate answers to questions.

16 In what respects are Kayne's (2000, Ch. 15) and (2003) proposals concerning Heavy-NP Shift more ambitious than Jayaseelan's? To what extent are they successful?

References

Aboh, E. O. 2004a. *The Morphosyntax of Complement–Head Sequences: Clause Structure and Word Order Patterns in Kwa*. New York and Oxford: Oxford University Press.

Aboh, E. O. 2004b. Topic and Focus within D. In L. Cornips and J. Doetjes (eds), *Linguistics in the Netherlands 21*, Amsterdam and Philadelphia: John Benjamins, pp. 1–12.

Abraham, W. 1986. Word order in the middle field of the German sentence. In W. Abraham and S. de Meij (eds), *Topic, Focus and Configurationality (Papers from the 6th Groningen Grammar Talks, Groningen, 1984)*, Amsterdam: John Benjamins, pp. 5–38.

Belletti, A. 1990. *Generalized Verb Movement*. Torino: Rosenberg & Sellier.

Biberauer, T. 2008. Doubling and omission: Insights from Afrikaans negation. In S. Barbiers, O. Koeneman, M. Lekakou, and M. van der Ham (eds), *Microvariations in Syntactic Doubling. Syntax and Semantics* 36, London: Emerald, pp. 103–140.

Brody, M. 1990. Remarks on the order of elements in the Hungarian focus field. In I. Kenesei (ed.), *Structures and Arguments. Approaches to Hungarian* 3, Szeged: JATE, pp. 95–121.

Cecchetto, C. 1999. A comparative analysis of left and right dislocation in Romance. *Studia Linguistica* 53: 40–67.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, pp. 1–52.

Cinque, G. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36: 315–332.

de Hoop, H. 1992. *Case configuration and noun phrase interpretation*. Doctoral dissertation, University of Groningen.

Diesing, M. 1992. *Indefinites*. Cambridge, MA: MIT Press.

Diesing, M. 1997a. Light verbs and the syntax of aspect in Yiddish. *Journal of Comparative Germanic Linguistics* 1: 119–156.

Diesing, M. 1997b. Yiddish VP order and the typology of object movement in Germanic. *Natural Language and Linguistic Theory* 15: 369–427.

É. Kiss, K. 2002. *The Syntax of Hungarian*. Cambridge: Cambridge University Press.

Johnson, K. 1991. Object positions. *Natural Language and Linguistic Theory* 9: 577–636.

Karimi, S. 1999. A note on parasitic gaps and specificity. *Linguistic Inquiry* 30: 704–713.

Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2003. Some remarks on agreement and on Heavy-NP-Shift. In M. Ukaji, M. Ike-Uchi, Y. Nishimura, and H. Nakajima (eds), *Current Issues in English Linguistics. Special Publications of the English Linguistic Society of Japan* 2, Tokyo: Kaitakusha, pp. 67–86. [Reprinted in Kayne (2005, Ch. 11).]

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2010a. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Kayne, R. S. to appear. Why isn't *this* a complementizer? In P. Svenonius (ed.), *Functional Structure from Top to Toe: A Festschrift for Tarald Taraldsen*, New York and Oxford: Oxford University Press. [Reprinted in Kayne (2010b, Ch. 10).]

Kayne, R. S., and J.-Y. Pollock. 1978. Stylistic inversion, successive cyclicity and move NP in French. *Linguistic Inquiry* 9: 595–621.

Kayne, R. S., and J.-Y. Pollock. 2001. New thoughts on stylistic inversion. In A. C. Hulk and J.-Y. Pollock (eds), *Inversion in Romance*, New York and Oxford: Oxford University Press, pp. 107–162.

Koopman, H., and D. Sportiche. 1989. Pronouns, logical variables, and logophoricity in Abe. *Linguistic Inquiry* 20: 555–588.

Kuno, S. 1975. Conditions for verb phrase deletion. *Foundations of Language* 13: 161–175.

Larson, R. 1988. On the double object construction. *Linguistic Inquiry* 19: 335–391.

Larson, R. 1990. Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry* 21: 589–632.

Lenerz, J. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen. Studien zur deutschen Grammatik 5*. Tübingen: Narr.

Munaro, N., and J.-Y. Pollock. 2005. Qu'est-ce que *(qu)-est-ce-que*? A case study in comparative Romance interrogative syntax. In G. Cinque and R. S. Kayne, *The Oxford Handbook of Comparative Syntax*, New York and Oxford: Oxford University Press, pp. 542–606.

Poletto, C. 2000. *The Higher Functional Field: Evidence from Northern Italian Dialects. Oxford Studies in Comparative Syntax*. New York and Oxford: Oxford University Press.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 281–337.

Villalba, X. 1999a. Right dislocation is not right dislocation. In O. Fullana and F. Roca (eds), *Studies on the Syntax of Central Romance Languages. Proceedings of the III Symposium on the Syntax of Central Romance Languages*, Girona: University of Girona, pp. 227–241.

Villalba, X. 1999b. Symmetry and antisymmetry in syntax. *Syntaxis* 2: 1–25.

Zwart, C. J. W. 1993. *Dutch syntax: A minimalist approach*. Doctoral dissertation, University of Groningen.

Zwart, C. J. W. 1996. *Morphosyntax of Verb Movement. A Minimalist Approach to the Syntax of Dutch*. Dordrecht: Kluwer.

32

The Configurational Structure of a Nonconfigurational Language

Julie Anne Legate

2001

32.1 Introduction

This article investigates Warlpiri, a Pama–Nyungan language that exhibits the characteristics of so-called "nonconfigurational" languages: relative freedom of word order, the existence of discontinuous noun phrases, and the possibility of having null arguments. Since the pioneering work of Ken Hale (cf. Hale [1982b], 1983), these properties have put it at the center of the debate on how to capture the difference between nonconfigurational languages, like Warlpiri and Mohawk (cf. Baker 1996), and configurational languages, like English. Which aspects of the grammatical system account for the presence (or absence) of free word order, discontinuous constituents, and null arguments? Do these three properties cluster together? If so, is there a single parameter that can capture them? Focusing on free word order and discontinuous constituents, do they arise from the fact that non-configurational languages use means of combining words into sentences different from those employed by configurational languages? For example, do they fail to group words into phrasal constituents, or smaller constituents (like a DP) into larger constituents (like a VP)? Thinking more broadly, do notions such as subject and object have a structural correlate across all languages, or are they structurally defined only in configurational languages?

Legate's approach to these questions is similar to that seen in other contributions collected in this volume (like Anderson and Chung's and Massam's, Ch. 5 and Ch. 29) and unlike the approach adopted in Hale (1983), Bresnan and Moshi (1990), Simpson (1991), and Bresnan (2001), in that it finds evidence in favor of postulating hierarchical structure in a language that is quite different from a language like English. One key observation on which the paper builds is that Warlpiri

(like other nonconfigurational languages) does not systematically fail to exhibit properties that can be attributed to the presence of hierarchical structure – like asymmetries between subjects and objects, between direct and indirect objects, or between arguments and adjuncts. Such asymmetries indeed exist in the language, though they are visible only in a subset of the cases in which they appear in configurational languages. Legate argues that these gaps can be explained by examining aspects of the grammar other than phrase structure.

For example, Warlpiri has a ditransitive construction similar to the English double-object construction, in which the subject displays ergative, the indirect object dative, and the direct object absolutive case, and where only the dative argument shows characteristic properties of objects (for example, triggers object agreement and controls an embedded subject). In addition, it has an ethical dative construction, which also exhibits an ergative subject, a dative indirect object, and a direct object, but where both the dative and the direct object exhibit characteristic properties of objects. The parallel behavior of these two arguments, Legate argues, is not due to the lack of a structural asymmetry between them. Rather, it has to do with the nature of a functional head, the Applicative head (cf. Pylkkänen 2008), that is present in the structure: in the ethical dative construction, this head is verbal in nature; thus both the ethical dative and the direct object are in a structural configuration with a verbal head (see (27)). In contrast, in the double object construction, the Applicative head is prepositional in nature, and thus one object enters a relation with a prepositional head, the other with a verbal head (see (26)). If the lack of asymmetry between the two objects in ethical dative constructions is a consequence of the nature of the Applicative head, then it is perfectly compatible with Warlpiri being fully configurational (cf. also Legate 2003).

In the rest of the article, Legate adduces further arguments, from adverb placement and from the structure of the left periphery, showing that Warlpiri exhibits configurational properties in a range of syntactic domains. A similar line of reasoning is also pursued in a different empirical domain in Bruening (2001, 2009).

32.2   From "THE CONFIGURATIONAL STRUCTURE OF A NONCONFIGURATIONAL LANGUAGE"

1   Introduction

Warlpiri is a Pauma-Nyungan language spoken in Northern Territory, Australia, by over 3000 people. A number of properties of this language that made it appear typologically unusual were examined in Hale's (1983) seminal paper, which brought both Warlpiri and "nonconfigurationality" to the forefront of generative linguists. These properties included free word order, possible *pro*-drop of all arguments and adjuncts, and discontinuous noun phrases; these subsequently became the hallmarks of nonconfigurational languages.

Beginning with Hale (1983), Warlpiri has been seen to require adding additional parameters into the typological space of human language. In this paper, I suggest

that such a move is unnecessary, and hence undesirable. Warlpiri syntax may be analysed using a hierarchical structure, consisting of crosslinguistically-motivated projections, in combination with movement operations familiar from other languages.

In Section 2, I outline the "flat-structure" approach to Warlpiri syntax, originally due to Hale (1983), and recently revived by Austin & Bresnan (1996) and Bresnan ([2001]). The remainder of the paper reveals difficulties with such an approach, by presenting evidence for a hierarchical syntactic structure in Warlpiri. Section 3 considers the verb phrase, arguing on the basis of double object and ethical dative constructions for a hierarchically-organized verb phrase in Warlpiri. Section 4 examines the clause structure above the verb phrase. First, I demonstrate that Cinque's (1999) hierarchy of functional projections that introduce adverbs into the syntax applies equally to Warlpiri, and discuss the difficulties this raises for the flat-structure approach. Next, I examine the left periphery (Rizzi 1997) of Warlpiri, demonstrating the existence of distinct and hierarchically ordered projections specialized for two types of topics and two types of foci. Further, I provide evidence that the placement of (*wh*-)phrases in the left periphery is the result of movement rather than base-generation.

The following section begins, with an outline of the flat-structure approach to Warlpiri syntax.

2 Approaches

In this section, I review a number of well-known properties of Warlpiri syntax and outline the flat structure analysis of these data. The analysis of Warlpiri is complex in that certain aspects of the syntax exhibit asymmetries among and between arguments and adjuncts, while others systematically fail to. As mentioned above, word order, the possibility for pro-drop, and the ability for noun phrases to appear discontinuously grant the same freedom to all arguments and adjuncts. Asymmetries between arguments cannot be found in Weak Crossover effects, or Condition C data either, in that WCO effects do not appear in object *wh*-questions, and Condition C behaves as though subjects and objects stand in a relationship of mutual c-command:

(1) *WCO*
   a. *Ngana-ngku kurdu nyanungu-nyangu paka-rnu?*
       who-ERG   child 3-POSS   hit-NPST
       "Who$_i$ hit his$_i$ child?"
   b. *Ngana ka   nyanungu-nyangu maliki-rli wajili-pi-nyi?*
       who   PRES.IMPF he-POSS   dog-ERG   chase-NPST
       "Who$_i$ is his$_i$ dog chasing?" (Hale et al. 1995:1447)
(2) *Condition C*
   a. *Nyanungu-rlu*$_{i/j}$ maliki Jakamarra$_i$-kurlangu paka-rnu*
       3-ERG   dog   Jakamarra-POSS   hit-PST
       "He*$_{i/j}$ hit Jakamarra$_i$'s dog"

b. *Jakamarra*ᵢ-*kurlangu maliki-rli nyanungu**ᵢ/ⱼ *paji-rni*
Jakamarra-POSS   dog-ERG 3   bite-PST
"Jakamarraᵢ's dog bit him*ᵢ/ⱼ" (Laughren 1991:14)

In contrast, Condition A behaves as though the subject asymmetrically c-commands the object, and Condition B distinguishes objects from adjuncts.

(3) *Condition A*
  a. *Purlka-jarra-rlu   ka-pala-nyanu   nya-nyi*
     old.man-DUAL-ERG PRES.IMPF-3DUAL-REFLEX see-NPST
     "The two old men are looking at each other" (Simpson 1991:163)
  b. **Purlka-jarra ka-nyanu-palangu   nya-nyi*
     old.man-DUAL PRES.IMPF-REFLEX-3DUAL.OBJ see-NPST
     Lit: "Each other are looking at the old men."
(4) *Condition B*
  a. **Jakamarra-rlu ka-(nyanu)   nyanungu paka-rni*
     Jakamarra-ERG PRES.IMPF-(REFLEX) 3   hit-NPST
     "Jupurrurlaᵢ is hitting himᵢ" (Simpson 1991:170)}
  b. *Japanangka-rlu-nyanu yirra-rnu mulukunpa nyanungu-wana*
     Japanangka-ERG-REFLEX put-NPST bottle   3-PERL
     "Japanangkaᵢ set the bottle down beside himᵢ." (Simpson 1991:171)

Furthermore, Warlpiri shows suppletion of infinitival complementizers, sensitive to the grammatical function of the controller of the PRO subject. Thus, *-karra* indicates control of the embedded subject by the matrix subject, *-kurra* indicates control by the matrix object, and *-rlarni* is the default, used for control by a matrix adjunct or when the embedded clause has an overt subject.
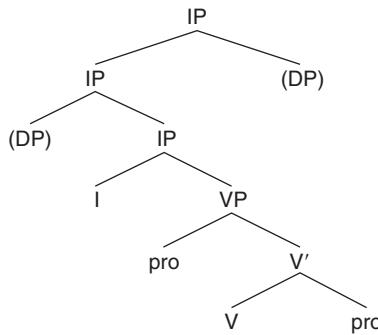
(5) *Embedded complementizers*
  a. *Karnta ka-ju   wangka-mi   [yarla karla-nja-karra]*
     woman PRES.IMPF-1SG speak-NONPST yam dig-INF-SUBJC
     "The woman is speaking to me while digging yams" (Hale 1983:21)
  b. *Purda-nya-nyi   ka-rna-ngku   [wangka-nja-kurra]*
     aural-perceive-NONPST PRES.IMPF-1SG-2SG speak-INF-OBJC
     "I hear you speaking" (Hale 1983:20)
  c. *Wati-rla   jurnta-ya-nu karnta-ku [jarda-nguna-nja-rlarni]*
     man-3SG.DAT away-go-PST woman-DAT sleep-lie-INF-OBVC
     "The man went away from the woman while she was sleeping" (Hale et al. 1995:1442)

Such a bifurcation of behaviours is not unique to Warlpiri, but is attested in a number of "nonconfigurational languages" (see, for example, the papers in Marácz & Muysken [1989a]).

One previous approach to the conflicting data found in nonconfigurational languages like Warlpiri I will term the *pronominal argument approach* (PA); two instantiations of this approach can be found in Jelinek (1984), and Baker (1996).

According to the PA, either all argument positions are filled by clitics, the overt DPs being adjuncts (Jelinek 1984); or the argument positions are filled by *pro*'s, the overt NPs being licensed by agreement morphology on the verb and appearing in a clitic left dislocation–type structure (Baker 1996).

(6)   *Pronominal Argument Approach*

This approach has initial plausibility in allowing a simple, single explanation for the complete range of data in nonconfigurational languages. However, subsequent research has determined that the various data do not seem to have a single source. Austin & Bresnan (1996) (henceforth A&B) examine Australian languages related to Warlpiri and carefully demonstrate that the nonconfigurational properties found in Warlpiri do not consistently co-occur, nor do these properties consistently co-occur with agreement-pronominal clitics, as required by the PA. Thus, a single parametric explanation for the full range of data found in Warlpiri does not seem appropriate, since the same phenomena in related languages cannot be so explained. Furthermore, A&B present a number of difficulties with the hypothesis within Warlpiri itself: several interpretive differences between arguments and adjuncts that would be unexpected on a theory in which all overt DPs are adjuncts; case marking on overt DPs based on lexical idiosyncrasies of particular verbs; the existence of DPs not linked to any agreement/pronominal clitic, and the ability of these DPs to undergo pro-drop. The reader is referred to A&B for details. Given these difficulties with the PA approach for Warlpiri, I will not consider it further.

The alternative approach proposed by A&B has its roots in Hale's (1983) original proposal for the structure of Warlpiri, and Simpson's (1991) related proposal. This approach claims that the syntactic structure of Warlpiri is n-ary branching, overt elements freely base-generated in any order. To account for the hierarchical properties of Warlpiri discussed above, such an approach must posit an additional level of representation which encodes asymmetries between subjects, objects, and adjuncts. A&B thus embed their approach within the framework of Lexical Functional Grammar (LFG), which allows for multiple levels of representation, including: f(unctional)-structure [sic], which encodes grammatical relations, and c(onstituent)-structure, which consists of the surface syntactic tree. Indeed, Bresnan ([2001]) presents Warlpiri as a primary motivation for the multi-level
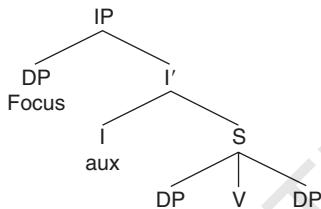
framework of LFG. Under the flat-structure approach, the asymmetric properties of Warlpiri are attributed to asymmetries among grammatical relations in the f-structure, while the symmetric properties of Warlpiri are attributed to a c-structure consisting of an n-ary branching *S*, a constituent which lacks a head and does not project. A&B also posit an IP projection above S, the head of which contains the auxiliary complex and the specifier of which hosts a focused constituent.

(7)　*Flat Structure Approach*
　　f-structure:

$$
\begin{bmatrix}
\text{PRED} & \text{'verb}\langle(f\text{SUBJ})(f\text{OBJ})\rangle\text{'} \\
\text{SUBJ} & [\text{"DP"}] \\
\text{OBJ} & [\text{"DP"}]
\end{bmatrix}
$$

　　c-structure:

The discussion in this paper concentrates on two aspects of A&B's account: the claim that Warlpiri phrase structure is flat, and the claim that it is characterized by free base-generation of elements in any order within the clause. I do not address the symmetric properties of Warlpiri directly, but note that these may be analysed as the result of UG-defined choices familiar from other languages. Thus, DP-splitting from Slavic and Germanic (see van Riemsdijk 1989; Krifka 1998; Ćavar & Fanselow 2000), *pro*-drop, ubiquitous throughout the world's languages; and scrambling that repairs WCO violations, perhaps best studied in German, Hindi (esp. Mahajan 1990), and Japanese (esp. Saito 1989, Miyagawa 1997). The Condition C data, although not unique to Warlpiri (see, for example Marácz & Muysken ([1989b]) for Hungarian), remains mysterious.

　　The next section examines syntax within the verb phrase in Warlpiri, arguing for hierarchy on the basis of double object and ethical dative constructions.

3　Within the Verb Phrase

In this section, I examine double object and ethical dative constructions in Warlpiri, first demonstrating that these represent two types of applicative constructions. Next, I discuss the LFG account of applicatives presented in Bresnan & Moshi (1990), and show that the Warlpiri data raises difficulties for such an account. Finally, I present an analysis of applicative constructions that assumes a hierarchical verb phrase, and show that the Warlpiri data may be accommodated within

such an analysis. To begin, I outline some crosslinguistic generalizations regarding applicative constructions.

Two types of applicatives have been identified crosslinguistically (see esp. Baker 1988, Bresnan & Moshi 1990), which are traditionally called "asymmetric" and "symmetric". As the names suggest, asymmetric applicatives are characterized by asymmetric behaviour between the verbal object (VO) and the applicative object (AO): only the AO shows primary object properties. In contrast, in symmetric applicatives both the AO and VO show primary object properties. Glossing over some interesting complications that arise within particular languages, the cluster of properties of symmetric and asymmetric applicatives are summarized in the following table.

(8)   *Types of Applicatives Crosslinguistically*

Asymmetric	Symmetric
AO shows object properties (agreement, passives, scope, . . .)	AO, VO show object properties (agreement, passives, scope, . . .)
transitivity restriction on verb	no transitivity restriction on verb
animacy restriction on AO	no animacy restriction on AO
AO semantically related to VO	AO semantically related to event

In Legate (2001), I demonstrated that Warlpiri has both types of applicative constructions. Thus, a class of ditransitive verbs are asymmetric applicatives and the ethical dative construction is a symmetric applicative. In the next section I begin with the ditransitives.

3.1   Ditransitives

Warlpiri has a class of verbs with an ERG-DAT-ABS case frame, that is the subject displays ergative case, the indirect object displays dative case, and the direct object shows absolute case. An example of such a verb is *yi-nyi* "give":

(9)   *Warnapari-rli ka-rla   kurdu-ku ngapurlu yi-nyi.*
       dingo–ERG   PRES.IMPF–3DAT child–DAT milk   give–NPST
       "The dingo gives milk to the little one."

I argue that this is not a PP-dative construction, as the translation suggests, but rather an asymmetric applicative construction, akin to the English double object construction: *The dingo gives the little one milk*.

First, the dative AO shows primary object properties for agreement and control (Simpson 1991). Thus, the dative AO triggers object agreement rather rather than the absolutive VO:

(10)   *Ngajulu-rlu kapi-rna-ngku   karli-patu   yi-nyi   nyuntu-ku*
        I–ERG   FUT.C–1SG.S–2SG.O boomerang-pauc give–NPST you–DAT
        "I will give you (the) (several) boomerangs" (Hale et al. 1995:1432)

Furthermore, recall that Warlpiri embedded infinitival complementizers supplete according to the grammatical function of the controller of their PRO subject, see (5) above. When the dative AO controls a PRO infinitival subject, the embedded complementizer -*kurra* is used, registering control by a matrix object. This complementizer cannot be used when the absolutive VO controls the embedded subject.

(11) a. *Karnta-ngku ka-ju   kurdu miliki-yirra-rni*
          woman-ERG PRES.IMPF–1SG.O child show-put-NPST
          nguna-nja-kurra-(ku)
          lie-INFIN-OBJ.C–(DAT)
          "The woman is showing the child to me while I am lying down"
          (Simpson 1991:342)
      b. ⁇*Yu-ngu-rna-rla   kurdu parraja-rla   ngunga-nja-kurra*
          give-PST-1SG.S–3DAT child coolamon-LOC sleep-INFIN-OBJ.C
          yali-ki
          that-DAT
          "I gave the child which was sleeping in the coolamon to that one"
          (Simpson 1991:341)}

Furthermore, ERG-DAT-ABS verbs fall into the familiar crosslinguistic classes of double object verbs (see Levin 1993, Pesetsky 1995).

(12) *Double Object Verb Classes*:
      a. inherently signify act of giving: *yi-nyi* "give"
      b. inherently signify act of taking: *punta-rni* "take away from", *jurnta-ma-ni* "take away from", *jurnta-marda-rni* "take away from", *punta-punta-yirra-rni* "take away from", …
      c. instantaneous causation of ballistic motion: *kiji-rni* "throw" (cf. not *rarra-ma-ni* "drag")
      d. sending: *yilya-mi* "send/throw to"
      e. communicated message: *ngarri-rni* "tell", *payi-rni* "ask", *japi-rni* "ask", *milki-yirra-rni* "show" (cf. not *wangka-mi* "speak/say", *jaalyp(a)-wangkami* "whisper")
      f. continuous causation of accompanied motion in some manner: *kanyi* "carry, bring, take"

Also, there exists an alternation in Warlpiri between the ERG-DAT-ABS and an ERG-ABS-ALL(ative) ditransitive, an alternation comparable to the double object versus PP-dative alternation. In the ERG-ABS-ALL variant, it is the ABS that controls object agreement:

(13) *The Allative Variant*
      Yu-ngu-ju-lu   Jakamarra-kurra
      give-PST-1SG.O–3PL.S Jakamarra-ALL
      "They gave me to Jakamarra" (Laughren 1985)

In addition, asymmetric applicatives crosslinguistically display a characteristic semantics, in which the AO is interpreted as a (potential) possessor of the VO. The

dative AO of ERG–DAT–ABS verbs receives this interpretation, whereas the allative of the ERG–ABS–ALL variant does not. Thus, of the pair in (14),

(14) a. *Ngarrka-ngku ka-rla   kurdu-ku japujapu kiji-rni*
         man–ERG   PRES.IMPF–3DAT child–DAT ball   throw–NPST
         "The man is throwing the child the ball"

     b. *Ngarrka-ngku ka   japujapu kurdu-kurra kiji-rni*
         man–ERG   PRES.IMPF ball   child–ALL   throw–NPST
         "The man is throwing the ball to the child" (Hale [1982a]:253)

Hale ([1982a]) remarks that "[the] dative in [(14a)] implies that the child is the recipient of the ball, not merely the endpoint of motion. The allative in [(14b)], on the other hand, implies that the child – or the child's location – is merely the end-point of the trajectory traversed by the ball." (Hale [1982a]:253)

Finally, related to the possessive semantics, crosslinguistically we find an animacy restriction on the goal (AO) of asymmetric applicatives. This animacy restriction is also found on the dative AO of ERG–DAT–ABS verbs; if the AO is inanimate, the absolutive-allative variant must be used instead.

(15) a. *Purturlu kala-rla   yilya-ja*.
         backbone PST.C–3DAT send–PST
         "He sent her the backbone"

     b. *Marnkurrpa-rna yilya-ja   Yalijipiringi-kirra*
         three-1SG.   S send–PST Alice.Springs–ALL
         "I sent three to Alice Springs"

Thus, I conclude that ditransitive verbs which display the ERG–DAT–ABS case frame should be identified as asymmetric applicatives.

In the next section we consider a second applicative construction in Warlpiri, the ethical dative construction.

3.2 Ethical datives

The Warlpiri ethical dative construction involves the addition of a dative DP, without an overt morpheme to indicate how the additional DP is to be interpreted. An example of this is given in (16):

(16) *Karli   yinga-rla   paka-rni jinta-kari-rli nyanungu-ku*
      boomerang REAS.C–3DAT chop–NPST one-other–ERG he–DAT
      "Because the other one will chop a boomerang for him" (Simpson 1991:381)

This construction proved problematic for previous analyses of Warlpiri lexical structure, notably the detailed LFG account of Simpson (1991). Simpson is forced to posit a new grammatical function for ethical datives, which she calls "EXTERNAL OBJECT", in addition to an optional process promoting ethical datives to the "OBJECT" function.

Examining the construction, we discover that it exhibits distinct behaviour from the double objects considered above. First, both the ethical dative (AO) and the object of the verb (VO) trigger object agreement. Due to a morphophonological restriction against both dative and absolutive agreement in the auxiliary (Simpson 1991), this agreement pattern is visible only when the object of the verb is also dative. In such a case, both datives are obligatorily registered in the auxiliary. Thus, in (17), *warri-rni* "seek" selects a dative object, and the auxiliary agrees with both this VO object and the dative AO.

(17) *Ngarrka-ngku ka-ju-rla   ngaju-ku karli-ki*
     man-ERG   PRES.IMPF-1SG.O-3DAT me-DAT   boomerang-DAT
     warri-rni
     seek-NPST
     "The man is looking for a boomerang for me" (Hale [1982a]:255)

In addition, when either the VO or the AO control an embedded PRO subject, the *-kurra* complementizer appears, indicating control by a matrix object.

(18) *Control by dat*
    a. *Kamina-rlu ka-rla   mangarri purra   ngati-nyanu-ku*
       girl-ERG   PRES.IMPF-3DAT food   cook.NPST mother-self-DAT
       nguna-nja-kurra-ku
       lie-INFIN-OBJ.C-DAT
       "The girl is cooking food for her mother who is lying down."
       (Simpson 1991:385)
    b. *Control by abs*
       Maliki-rna ramparl-luwa-rnu Jakamarra-ku parnka-nja-kurra
       dog-1SG   accident-hit-PST Jakamarra-DAT run-INFIN-OBJC
       "I accidently hit Jakamarra's dog while it was running." (ElD)

Furthermore, unlike asymmetric applicatives (see the table in (8)), there is no transitivity restriction on the ethical dative construction.

(19) a. *Karnta ka-rla   kurdu-ku parnka-mi*
       woman PRES.IMPF child-DAT run-NPST
       "The woman is running for the sake of the child" (Simpson 1991:381)
    b. *Nantuwu ka-rla   Japanangka-ku mata-jarri-mi*
       horse   PRES.IMPF-3DAT Japanangka-DAT tired-INCH-NPST
       "The horse is tiring on Japanangka" (Hale [1982a]:254)

Finally, we do not find the possessive semantics characteristic of asymmetric applicatives in the ethical dative construction. Instead, interpretation of the dative AO "embrac[es] a considerable range of possible semantic connections which may hold between an entity and an event or process" (Hale [1982a]:254), including at least benefactive, malefactive, and possessive:

(20) a. *Nantuwu ka-rla   Japanangka-ku mata-jarri-mi*
       horse   PRES.IMPF-3DAT Japanangka-DAT tired-INCH-NPST

"The horse is tiring on Japanangka"
"Japanangka's horse is tiring"

b. *Ngarrka-ngku ka-rla   kurdu-ku karli   jarnti-rni*
   man-ERG   PRES.IMPF–3DAT child-DAT boomerang trim-NPST
   "The man is trimming the boomerang for the child"
   "The man is trimming the child's boomerang" (Hale [1982a]:254)

In sum, the properties displayed by the Warlpiri ethical datives are those of a symmetric applicative construction. I conclude that Warlpiri has both an asymmetric and a symmetric applicative. In the next section [3.3 of the original article], we discover that this conclusion poses difficulties for an LFG account of applicatives.

[…]
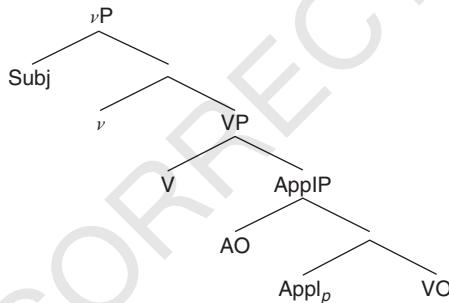
3.4 A structural account
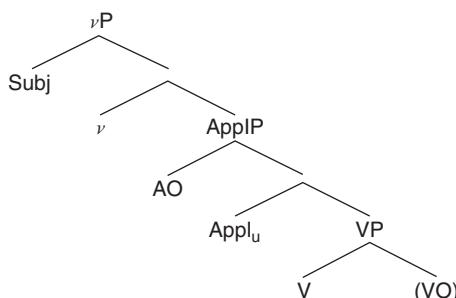
The analysis of applicative constructions I present here is a modification of McGinnis ([2001]). I adopt the insights of her proposal, while eliminating some of the technology by exploiting categorial differences between the applicative head that appears in symmetric applicatives and the applicative head that appears in asymmetric applicatives.

Under this approach, symmetric and asymmetric applicatives differ structurally:

(24) *Asymmetric Applicative (cf. Pesetsky 1995)*

(25) *Symmetric Applicative (cf. Marantz 1993):*
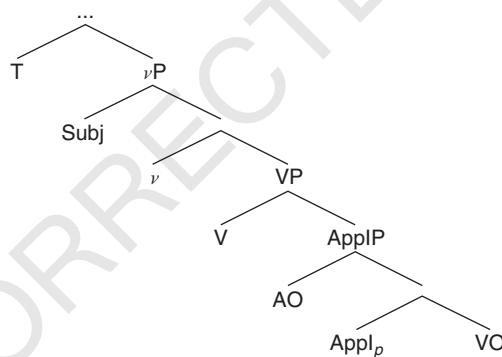
In the asymmetric applicative, the phrase headed by the applicative morpheme appears as the complement to the verb. I assume it is therefore prepositional in nature. This applicative preposition relates the AO, in its specifier, to the VO in its complement, establishing the semantic relationship of (potential) possession between them. The structure therefore captures the inability of asymmetric applicatives to appear with intransitive verbs, as well as the characteristic semantic interpretation of the AO as a potential possessor.
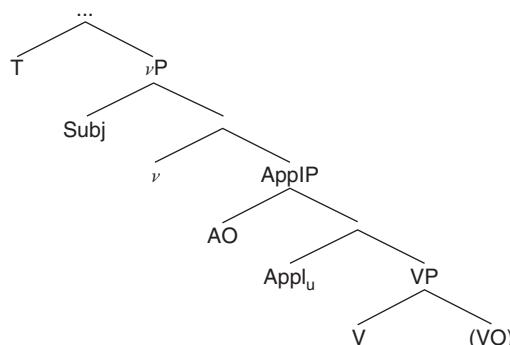
In the symmetric applicative, on the other hand, the phrase headed by the applicative morpheme dominates the verb phrase. I assume that it is therefore a type of light verb, or v. Since the AO is related directly to the VP, this structure captures the lack of transitivity restriction on symmetric applicatives, as well as the interpretation of the AO as being related to the event.

I argue that the distinction between the nature of the applicative morphemes, prepositional for asymmetric applicatives and verbal for symmetric applicatives, has significant repercussions throughout the syntax of the constructions. In the asymmetric applicative, the applicative preposition assigns case to the VO in its complement, and the AO raises to check case and agreement with the v that introduces the subject. In the symmetric applicative, the VO raises to check case with the applicative v, and the AO raises to check case with the v that introduces the subject. This results in the following configurations (before subject raising and verb movement):

(26)   *Asymmetric Applicative*

(27)   *Symmetric Applicative:*

These structures allow us to understand the differing behaviour of VOs between symmetric and asymmetric applicatives. In symmetric applicatives, both the AO and the VO enter an agreement relationship with a *v* head, and thus both exhibit behaviour as objects. In asymmetric applicatives, on the other hand, only the AO agrees with a *v* head, the VO being the object of a preposition, and therefore, only AO behaves as a direct object. One direct consequence of this agreement relationship is that in symmetric applicative constructions, both the AO and the VO may trigger object agreement morphology, since both agree with a *v*, an extended projection of the verb. This is illustrated in (28) with data from Kichaga. In asymmetric applicative constructions, only the AO triggers object agreement morphology, since only the AO agrees with a *v*; the VO is assigned case by a preposition. This is shown in (29) for Chicheŵa.

(28) a. *N–ä–ï–ḿ–lyì–í–à*   *k–èlyâ*.
          FOC–1 S-PRES-1O-eat-APPL-FV 7-food
          "He/she is eating food for/on him/her."
      b. *N–ä–ï–kì–lyí–í–à*   *ḿ–kà*.
          FOC–1 S-PRES-7O-eat-APPL-FV 1-wife
          "He/she is eating it for/on the wife."
      c. *N–ä–ï–kï–ḿ–lyì–ï–à*
          FOC–1 S-PRES-7O0-1O-eat-APPL-FV
          "He/she is eating it for/on him/her."
          (Bresnan & Moshi 1990:150–151)

(29) a. *Amayi a–ku–mu–umb–ir–a*   *mtsuko*.
          woman SP-PRES-OP-mold-APPL-ASP waterpot
          "The woman moulded the waterpot for him."
      b. **Amayi a–na–u–umb–ir–a*   *mwana*.
          woman SP-PST-OP-mold-APPL-ASP child
          "The woman is moulding it for the child." (Baker 1988:247)}

[…]

Returning to Warlpiri, recall that object agreement is triggered by both the AO and the VO in symmetric applicatives, but only the VO in asymmetric applicatives. Although overt nominals in Warlpiri inflect on an ergative-absolutive pattern, agreement morphology shows a nominative-accusative paradigm, requiring a dissociation between case and agreement in the language. Thus, agreement relations in Warlpiri may pattern identically to the case-agreement relationships discussed with respect to the applicative structures above. In the symmetric applicatives, both AO and VO agree with a *v* head, and thus both trigger object agreement morphology. In the asymmetric applicatives, however, only the AO agrees with a *v* head (the VO agreeing with the applicative preposition), and so only the AO controls object agreement.

In addition, embedded infinitival complementizers in Warlpiri register object control when either the AO or VO of a symmetric applicative control the PRO subject of the embedded clause, but only when the AO of an asymmetric applicative controls the subject of the embedded clause. The examples are repeated below:

(32)   a.   *Kamina-rlu ka-rla   mangarri purra   ngati-nyanu-ku*
            girl-ERG   PRES.IMPF-3DAT food   cook.NPST mother-self-DAT
            nguna-nja-kurra-ku
            lie-INFIN-OBJ.C-DAT
            "The girl is cooking food for her mother who is lying down."
            (Simpson 1991:385)

       b.   *Maliki-rna ramparl-luwa-rnu Jakamarra-ku parnka-nja-kurra*
            dog-1SG   accident-hit-PST Jakamarra-DAT run-INFIN-OBJC
            "I accidently hit Jakamarra's dog while it was running." (ElD)

(33)   a.   *Karnta-ngku ka-ju   kurdu miliki-yirra-rni*
            woman-ERG   PRES.IMPF-1SG.O child show-put-NPST
            nguna-nja-kurra-(ku)
            lie-INFIN-OBJ.C-(DAT)
            "The woman is showing the child to me while I am lying down."
            (Simpson 1991:342)

       b. ??*Yu-ngu-rna-rla   kurdu parraja-rla   ngunga-nja-kurra*
            give-PST-1SG.S-3DAT child coolamon-LOC sleep-INFIN-OBJ.C
            yali-ki
            that-DAT
            "I gave the child which was sleeping in the coolamon to that one."
            (Simpson 1991:341)}

The verb phrase structures proposed above allow a simple characterization of this data. Control by a nominal within the *v*P domain registers as object control, whereas control by a nominal (in an A-position) above the *v*P registers as subject agreement, and control by a nominal within the VP triggers the default complementizer. This generalization may be technically implemented in a number of ways, the choice among which seems immaterial here.

To conclude this section, I have demonstrated that Warlpiri exhibits both a symmetric and an asymmetric applicative construction. I showed that the Warlpiri applicative data is problematic for an LFG analysis of applicatives (Bresnan & Moshi 1990), which uses a-structure and f-structure to account for the differing behaviour of noun phrases in applicatives, rather than using syntactic structure. Since a flat-structure analysis of Warlpiri requires differences in the behaviour of noun phrases to be encoded at a-structure/f-structure (by hypothesis no asymmetries between noun phrases are present in the syntactic structure), the applicative data is problematic for a flat-structure analysis of Warlpiri. Finally, I outlined a crosslinguistic analysis of applicative constructions which attributes the differing behaviour of noun phrases to a hierarchical syntactic structure, and showed that the Warlpiri data are compatible with such an analysis.

32.3   Questions pertaining to Legate (2003)

1   What part(s) of Holmberg and Platzack (1995) offer support to this paper of Legate's?

2   How do Romance pronominal clitics bear on the pronominal argument approach to nonconfigurational languages favored by Baker (1996)?

3   Discuss the implications of Ngonyani's (1996) argument that (Bantu) symmetric applicatives are not fully symmetric.

4   Asymmetric applicatives are often said to display an animacy restriction on the goal, as shown by *They sent John/*New York City the money*. In what direction(s) do sentences such as *They gave the ball a kick, They're giving the house a new coat of paint, Its elegance is what gives your idea its appeal* lead?

5   Legate argues against Bresnan and Moshi's (1990) distinction between asymmetric and symmetric (applicative) languages, on the grounds that Warlpiri has both types of applicatives. How exactly is Legate's argument similar to Chomsky's (1995, section 2.6.4) argument against Koopman's (1992) parametric proposal for Bambara?

6   Legate adopts an analysis of asymmetric applicatives in which the applicative head is prepositional. In what respects is her approach to these double object sentences different from earlier approaches formulated in terms of preposition deletion? How could one simultaneously integrate into such an applicative analysis both sentences like *They gave us a book*, which has a (near-) paraphrase with *to* and sentences like *They bought us a book*, which has a (near-)paraphrase with *for*?

7   The LFG flat structure analysis of Warlpiri that Legate critically discusses is incompatible with Kayne's (1981) binary branching hypothesis. Would replacing "binary branching" by "(only) ternary branching" make a difference for the LFG approach? How would changing "binary" to "ternary" affect the power of the theory in general? What would the implications be for intransitive verb phrases? for the role of precedence in syntactic relations?

8   How does Legate's discussion of Warlpiri adverbs bear on the general notion of "second-position clitics"? (Extra credit: Bring in Legate 2008)

9   Legate argues that phrases in the specifier of TopP must be moved there (via internal merge), rather than being merged there directly (via external merge). How might one relate her discussion to Chomsky's (1982, 1995) proposal to ban "vacuous quantification"?

10 Legate subsequently distinguishes TopP from a projection hosting "hanging topics" in the sense of Cinque (1977) and takes hanging topics to be merged directly where they are observed to be. What expectations concerning reconstruction effects are generated by this claim (thinking of Chomsky 1995, Ch. 3 on reconstruction)?

11 How might one link such hanging topics to English sentences like *As for John/ As far as John is concerned, we don't know much about him*, which seem to display surprising reconstruction possibilities, as in *As for her own kittens/As far as her own kittens are concerned, every female cat will defend them ferociously*?

12 With her example (48), Legate suggests that Warlpiri allows focus to co-occur with and to precede interrogative *wh*-phrases. Cinque (1999, p. 225) makes a comparable point for Italian (in embedded contexts only). How close is their use of the term FOCUS to Chomsky's (1972, pp. 89ff.)? To what extent might English sentences like *To John, what did they say?* be relevant?

13   Legate notes, in favor of her hierarchical approach, that Warlpiri allows long distance *wh*-extraction out of infinitivals but not out of finite clauses. Find at least three examples of such an extraction distinction (whether with *wh*-movement or other movements) in other languages.

14   How might the finite vs. nonfinite distinction of the previous question be related to the fact, noted by Legate, that infinitival clauses in Warlpiri show less freedom of word order than finite clauses? (Extra credit: Bring in English derived nominals.)

15   In her note 8, Legate suggests that adverbs that are free morphemes might be in specifier position, whereas the less free Warlpiri preverbs might be heads. To what extent might this idea draw support from Pollock's (1989) discussion of French negation?

References

Austin, P., and J. Bresnan. 1996. Non-configurationality in Australian Aboriginal languages. *Natural Language and Linguistic Theory* 14: 215–268.

Baker, M. 1988. Theta theory and the syntax of applicatives in Chicheŵa. *Natural Language and Linguistic Theory* 6: 353–389.

Baker, M. 1996. *The Polysynthesis Parameter*. New York and Oxford: Oxford University Press.

Bresnan, J. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Bresnan, J., and L. Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21: 147–185.

Bruening, B. 2001. *Syntax at the edge: Cross-clausal phenomena and the syntax of Passamaquoddy*. Doctoral dissertation, MIT.

Bruening, B. 2009. Algonquian languages have A-movement and A-agreement. *Linguistic Inquiry* 40: 427–445.

Ćavar, D., and G. Fanselow. 2000. Discontinuous constituents in Slavic and Germanic Languages. Ms., University of Potsdam.

Chomsky, N. 1972. *Studies on Semantics in Generative Grammar*. The Hague: Mouton de Gruyter.

Chomsky, N. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Cinque, G. 1977. The movement nature of left dislocation. *Linguistic Inquiry* 8: 397–412.

Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York and Oxford: Oxford University Press.

Hale, K. L. 1982a. Some essential features of Warlpiri main clauses. In S. M. Swartz (ed.), *Papers in Warlpiri Grammar: In memory of Lothar Jagst (= Work Papers of SIL-AAB, Series A, Vol. 6)*, pp. 217–315, Berrimah, NT: SIL-AAB.

Hale, K. L. 1982b. Preliminary remarks on configurationality. In J. Pustejovsky and P. Sells (eds), *Proceedings of NELS 12*, Amherst: GLSA, University of Massachusetts, pp. 86–96.

Hale, K. L. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory* 1: 5–47.

Hale, K. L., M. Laughren, and J. Simpson. 1995. Warlpiri Syntax. In J. Jacobs, A. von Stechow, W. Sternefeld, and T. Vennemann (eds), *Syntax. An International Handbook of Contemporary Research*, Berlin and New York: Walter de Gruyter, pp. 1430–1451.

Holmberg, A., and C. Platzack. 1995. *The Role of Inflection in Scandinavian Syntax*. New York and Oxford: Oxford University Press.

Jelinek, E. 1984. Empty categories, case, and configurationality. *Natural Language and Linguistic Theory* 2: 39–76.

Kayne, R. S. 1981. Unambiguous paths. In R. May and J. Koster (eds), *Levels of Syntactic Representation*, Dordrecht: Foris, pp. 143–183. [Reprinted in Kayne (1984, Ch. 7).]

Kayne, R. S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

Koopman, H. 1992. On the absence of case chains in Bambara. *Natural Language and Linguistic Theory* 10: 555–594.

Krifka, M. 1998. Scope inversion under the rise-fall pattern in German. *Linguistic Inquiry* 29: 75–112.

Laughren, M. 1985. Warlpiri reflexives, 'inherent' reflexives and grammatical relations. Ms., MIT.

Laughren, M. 1991. Some data on pronominal disjoint reference in Warlpiri. Ms., MIT.

Legate, J. A. 2001. Warlpiri datives: A high applicative, a low applicative, and an "applicative". Paper presented at the *Appl Fest*, MIT.

Legate, J. A. 2003. Reconstructing nonconfigurationality. In A. M. Di Sciullo (ed.), *Asymmetry in Grammar: Syntax and Semantics. Linguistik Aktuell / Linguistics Today* 57, Amsterdam and Philadelphia: John Benjamins, pp. 99–116.

Legate, J. A. 2008. Warlpiri and the theory of second position clitics. *Natural Language and Linguistic Theory* 26: 3–60.

Levin, B. 1993. *English Verb Classes and Alternations*. Chicago: University of Chicago Press.

Mahajan, A. 1990. *The A/A-bar distinction and movement theory*. Doctoral dissertation, MIT.

Marácz, L., and P. Muysken (eds), 1989a. *Configurationality: The Typology of Asymmetries*. Dordrecht: Foris.

Marácz, L., and P. Muysken. 1989b. Introduction. In L. Marácz and P. Muysken (1989a), pp. 1–38.

Marantz, A. 1993. Implications of asymmetries in double object constructions. In S. Mchombo (ed.), *Theoretical Aspects of Bantu Grammar*, Stanford: CSLI Publications, pp. 113–150.

McGinnis, M. 2001. Phases and the syntax of applicatives. In M.-J. Kim and U. Strauss (eds), *Proceedings of NELS 31*, Amherst: GLSA, University of Massachusetts, pp. 333–349.

Miyagawa, S. 1997. Against optional scrambling. *Linguistic Inquiry* 28: 1–26.

Ngonyani, D. 1996. *The morphosyntax of applicatives*. Doctoral dissertation, UCLA.

Pesetsky, D. 1995. *Zero Syntax: Experiencers and Cascades*. Cambridge, MA: MIT Press.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.

Pylkkänen, L. 2008. *Introducing Arguments*. Cambridge, MA: MIT Press.

van Riemsdijk, H. 1989. Movement and Regeneration. In P. Benincà (ed.), *Dialect Variation and the Theory of Grammar. Proceedings of the GLOW Workshop on Linguistic Theory and Dialect Variation*, Dordrecht: Foris, pp. 105–136.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 281–337.

Saito, M. 1989. Scrambling as semantically vacuous A'-movement. In M. R. Baltin and A. S. Kroch (eds), *Alternative Conceptions of Phrase Structure*, Chicago: University of Chicago Press, pp. 182–200.

Simpson, J. 1991. *Warlpiri Morpho-Syntax. A Lexicalist Approach*. Dordrecht: Kluwer.

33

Antisymmetry and Japanese

Richard S. Kayne

2003

33.1 Introduction

As generative syntax developed, certain notions and certain patterns became so familiar as to be taken for granted – for example, the notion that all phrases have one and only one head; or the so-called V2 pattern, by which some languages exhibit the finite verb in second position in matrix clauses, i.e., as the second constituent from the beginning of the sentence. Kayne's (1994) book *The Antisymmetry of Syntax* shook the field by raising important questions that were being left unanswered as the result of taking such notions and patterns for granted. For example, why should phrases have exactly one head? Why should certain patterns exist across languages, while other equally logical possibilities are unattested (for example, a language where the finite verb occurs as the second constituent from the end, rather the beginning, of the sentence)?

The core idea put forth in Kayne (1994) is that syntax is antisymmetric in the sense that if some subtree (with hierarchical structure and precedence relations specified), say (1a), is well-formed in some human language, its mirror-image, (1b), is not well-formed in any human language:

(1)   a.   A⟨B C⟩   b. *   ⟨C B⟩A

The precise implementation of the core idea of antisymmetry offered in Kayne's book is in terms of the Linear Correspondence Axiom (LCA), which determines the linearization of hierarchical structures by mapping asymmetric c-command among terminal nodes onto precedence.

The antisymmetric proposal makes nontrivial and precise predictions. For example, it predicts the impossibility of "inverse German," that is to say, of a language with V2 (of the German sort) but where the verb is the second constituent from the end of the sentence. It derives the restriction that a phrase must have exactly one head. It also denies the existence of headedness parameters in the traditional directionality parameter sense (cf. Travis 1989, Ch. 16 of this volume; Kayne 2011) because, in an antisymmetric approach, hierarchical structure fully determines linear order. Another interesting consequence of antisymmetry is that it assigns a central role to movement (any discrepancy in linear order must be related to movement), while at the same time restricting it to a single direction (assuming that movement must be upward), namely, leftward. Seeming rightward movement of a constituent must be analyzed as involving leftward movement of one or more constituents; and seeming right-adjunction of a phrase must be analyzed as left-adjunction followed by a set of movements. This, in turn, lends more prominence to derivations involving *remnant movement*, that is to say, movement of a constituent out of which something had previously moved (cf. den Besten and Webelhuth 1990; Müller 1998), as illustrated in (2):

(2) $[\ldots t_k \ldots]_j \ldots XP_k \ldots t_j$

The Antisymmetry of Syntax thus put forth a view of syntax that was more restrictive, and was able to derive some widely-assumed properties of human language syntax that were empirically motivated but had, until then, remained stipulated as primitives.

In the present paper, *Antisymmetry and Japanese*, Kayne offers further support for the antisymmetry hypothesis by examining the existence of crosslinguistic gaps, that is, language types that are easily imaginable but that do not seem to exist. He discusses a number of such gaps, arguing that a theory that incorporates the antisymmetry hypothesis is more readily amenable to accounting for them than one that does not. Kayne's work on some of these issues converges with work by Dominique Sportiche on selection and reconstruction of nominal complements, to the effect that prepositions, complementizers, and determiners may be introduced into the structure independently of the constituent that looks like their complement.

Once it is assumed that syntax is antisymmetric, the question arises – in the spirit of Chomsky (2004) – of why it is antisymmetric. This question implies that antisymmetry may not be a primitive, but a reflection of a more general property of human language (and perhaps other cognitive systems). This more general property might be a ban on optionality, as suggested in Kayne (2008).

33.2 From "ANTISYMMETRY AND JAPANESE"

1 Introduction

In this paper, I would like to focus on certain aspects of the antisymmetry hypothesis of Kayne (1994) and to a certain extent on their implications for Japanese. I will take as a starting point my hypothesis that syntactic structure is universally

and without exception of the form S-H-C. The complement of a head invariably follows that head. The associated specifier invariably precedes both head and complement.

This S-H-C hypothesis is to be taken to hold at all stages of a derivation, both prior to movement and subsequent to movement. (There is no sense in which it is a hypothesis about 'base' structure alone, contrary to the occasional misunderstanding.)

Given this, the relation between S-H-C and the question of OV/VO order cannot be a simple one. If an object can occupy the Spec of its own V (whether as the result of movement or, thinking of Larson (1988) and Barbiers ([2000]), because it is generated/merged there), then it will precede that V. If an object can remain in the complement position of its own V, then it will follow that V. If an object of V ends up in the Spec position of a head that is itself higher than V, then that object will precede V. If what we think of as an object of V ends up in the Spec position of a head lower than the position in which V itself ends up, then that object will follow V.

It follows that the OV order of a language like Japanese directly excludes the possibility that the object in Japanese has remained in the complement position of V. It must rather be in (or within) a Spec position, perhaps that of V itself, but much more likely that of some head higher than V. I return to this question shortly.

In Kayne (1994) I argued that specifiers are instances of phrasal adjunction. I continue to adopt here the position that syntax does not require a notion of phrasal adjunction distinct from the notion of specifier. Related to this was the claim that there cannot be multiple specifiers for a single head. I will maintain this claim, too. (For corroborating argument, see Rizzi (1997) and Cinque (1999).)

The S-H-C hypothesis (combined with the unavailability of adjunction as distinct from specifier and with the ban on lowering operations) leads to the conclusion that there can be no rightward movement operations in any language, as I will continue to hold.

In what follows, I will address some specific questions of Japanese syntax (as well as some more general considerations of syntactic theory). Whatever the subarea of Japanese syntax that one might be interested in, the question will arise as to whether or not one's analysis needs to be compatible with antisymmetry. If antisymmetry is not a valid characterization of UG, then the answer is no. If antisymmetry does constitute a valid characterization of UG, then the answer is yes.

Evaluation of the antisymmetry hypothesis must ultimately rest on evidence from as many languages as possible, in as many areas of syntax as possible. One must test to whatever extent possible the antisymmetry predictions concerning language 'gaps,' i.e. the claim that certain types of languages, though easily imaginable (such as 'reverse German'), will never be found. One must in some cases compare earlier analyses based on rightward movement with alternatives favored or imposed by antisymmetry, often making in exchange more extensive use of leftward movements (as discussed in part below). One must ask how antisymmetry (as opposed to a theory of syntax lacking it) interacts with other general properties of UG.

Since this evaluation will be potentially sensitive to evidence from any language, it is clear (but not surprising) that the question whether analyses of Japanese need to respect antisymmetry cannot be answered solely by looking at aspects of Japanese syntax. Evidence bearing on Japanese need not, to put it another way, come only from Japanese.

2   Japanese

2.1   The position of objects

It is uncontroversial to say that objects in Japanese can be found outside VP, in a higher Spec position, as in sentences whose derivation involves what we informally call scrambling. The question is whether Japanese objects ever surface within VP, in complement position of V (or in some Spec position lower than V). Antisymmetry says not, given OV order.

We can set aside potential cases of incorporation as not directly relevant, since incorporation by definition involves movement out of complement position. The standard view is that incorporated objects adjoin to V, as in Baker (1988). Since antisymmetry leads to the expectation that such head-to-head adjunction will be left-adjunction, the resulting OV order would be perfectly straightforward.

The plausibility of the claim that objects in Japanese are invariably found in (or within) a position higher than V is enhanced by the observation that in some OV languages objects move higher than V in a very visible way. One such case is Malayalam, in which objects must surface in a position preceding that of VP-external focus (SOFocV), as emphasized by Jayaseelan (2001). Another is that of the continental West Germanic languages (if we abstract away from V-2 contexts). In West Flemish, for example, objects precede one of the negation markers (see Haegeman (2001; [2002])). In Dutch and German, that type of (clitic) negation is not present, but the infinitive marker must still intervene between object and verb (O *te/zu* V_{infin}) and so must what are called separable particles, in the order 'O Prt *te/zu* V_{infin}', for example, in German:

(1)   . . . das Buch mitzubringen. ('. . .the book with/along-to-bring' = 'to bring along the book')

The above point about West Flemish negation is made in a more general way by Whitman ([2005]). Using earlier typological work by Dahl (1979) and Dryer (1988; 1992), Whitman notes the importance of the existence of a substantial set of 'SONegV' languages, i.e. languages in which the normal position of negation is between object and following verb. As he shows, they strongly suggest obligatory leftward movement of the object past negation.

On the basis of these considerations, the antisymmetry claim that Japanese objects must end up in or within a position asymmetrically c-commanding V is not very radical.

[. . .]

2.3 Head-Finality

It has often been noted that prenominal relative clauses of the Japanese sort are generally absent from VO languages (although for this to be strongly true (even apart from Chinese) one needs to set aside reduced subject-based relatives of the type found even in English *a recently arrived letter*). And it is sometimes thought that this supports a notion of 'head-finality' that languages like Japanese would display in a particularly consistent fashion.

In fact, I think that prenominal relatives pose a problem for any notion of consistent head-finality, since the supposed generalization rests on a double use of the term 'head'. If by 'head of a relative clause' we mean the material outside of the relative clause proper (but not counting higher determiner elements), then in *the apples that John bought, apples* is the head. The problem is that we also have *the pound of apples that John bought, the allusion to his wife that John got upset at*, etc., in which it looks like the head of the relative must then be *pound of apples* or *allusion to his wife*. But these are not heads (rather, they are phrases) in the sense in which V is a 'final head' in OV languages.

What this suggests, I suspect, is that although the (near-)exclusion of prenominal relatives in VO languages is something we want to find an explanation for, that explanation cannot reside in any notion of 'consistent head-finality.' This in turn is related, it seems to me, to a broader weakness in that notion, which has been based to a significant extent on the supposition that languages by and large pattern either as 'head-final' or as 'head-initial.' But that supposition looks highly questionable, as Kroch (2001: 706) has pointed out, observing that most languages are actually inconsistent in head-directionality. I think that Kroch is correct, especially as one takes into account a wider range of heads than was taken into account at earlier stages of the theory. When one broadens one's view away from just the lexical heads N, V and A (and perhaps P) to encompass complementizers of different sorts, question particles, topic and focus particles, and tense and aspect morphemes of various kinds, the inconsistency that Kroch points to becomes clear.[15]

In fact, it is very plausibly the case that Japanese itself is 'inconsistent,' in that its particles *wa* and *ga* are actually 'initial' heads. This proposal was made in Kayne (1994: 143) (based on an earlier similar proposal for Hungarian made by Brody (1990)) and has since been expanded on by Whitman (2001). The idea is that *wa* and *ga* are high functional heads in the sentential skeleton. *Wa*, for example, may be a Top^0 in Rizzi's (1997) sense, much as that discussed for Gungbe by Aboh (1999).

From an antisymmetric perspective, there is a natural reinterpretation of the notion 'final head,' which would be a head the entirety of whose complement has moved past it to a higher position or positions. A head that is not a 'final head' in this sense would be an 'initial head.' But there is no reason to think that 'final'/'initial' head is a primitive of syntactic theory, and no reason to think that languages must be consistent in having only one type. (In Kayne (1994: xv) I argued that no language could be uniformly head-final at all.)

3   Additional cross-linguistic 'gaps'

[. . .]

3.3   Adverbs and ('heavy') objects

Another cross-linguistic gap (that would be unexpected under a symmetric view of syntax) is found in the positional interaction of adverbs and definite objects. As is well-known, there are VO languages (such as English) in which V and O cannot be separated by adverbs. (In English, when there is only one non-prepositional object, V and O can be separated by a particle – even that is not possible in Danish (also VO).) What seems to be unattested is an OV language that systematically forbids its adverbs from intervening between O and V (at least when O is definite).

This point can be (informally) reformulated in terms of scrambling: An OV language will always allow (some) scrambling, at least with definites. That is not true of all VO languages. (Note, however, that to judge by Slavic some VO languages do allow (some) scrambling.) Part of this almost follows immediately. Since in an OV language the object must, from the present perspective, have moved to some Spec position higher than V, it is natural to think that it can therefore also reach a Spec position higher than at least some adverbs (yielding 'O Adv V'). Although a more precise account will have to be more specific about the character of the relevant landing sites, it is clear that in the absence of any possibility for rightward object movement (and right-adjoined positions for adverbs) there is no expectation that 'V Adv O' can be derived in parallel fashion. (For ways in which 'V Adv O' can be derived via leftward movement(s), see Pollock (1989) and Nilsen (2003).)

[. . .]

In analyzing heavy-NP shift as leftward movement of the object, den Dikken [1995] was led to propose, for (languages like) English, that such sentences also involve leftward VP-movement (to a position higher than the landing site of the object). Consider now English sentences like:

(3)   All of a sudden, there hit the building a shell *? (fired by our own troops).

The fact that this kind of sentence is much more acceptable if the subject argument contains a (reduced) relative, plus the post-object position of that subject argument, suggests a close link with heavy-NP shift sentences. At the same time, as Chomsky (1995: 343) has noted, such sentences appear to be close English counterparts of the well-known Icelandic transitive expletive construction. Jonas and Bobaljik (1993: 75n) in fact take (3) to be an instance of that construction, combined with heavy-NP shift.

Put another way, (3) should be thought of as having a derivation involving a prior stage of the form '. . . (there) a shell fired by our own troops hit the building.' Subsequently, the VP 'hit the building' moves leftward past the subject (but to a position lower than the (final) position of expletive *there*).

The fact that, in order to produce a grammatical output here, English, but not Icelandic, must have VP-movement apply can now be related to another difference between English and Icelandic concerning negation:

(4)   John has seen noone.

(5)   *John has noone seen.

In Icelandic the judgments are reversed, despite the fact that Icelandic is VO. My proposal in Kayne (1998) was that the negative *noone* moved leftward past the verb in both languages. Only in English is that negative phrase movement followed by VP-movement, arguably the same, or nearly the same, leftward VP-movement, as in (3). Leftward VP-movement thus allows us to tie together (3) and (4) in English vs. Icelandic in a natural way.

     [. . .]

3.6   Negation and auxiliaries

Basque is a VAux language that allows Aux . . . V if Aux is accompanied by negation. A conjecture worth testing would be:

(6)   No AuxV language has V . . . Aux licensed by negation.

3.7   DP

Cinque (1996: §4) has shown that Greenberg's (1966) Universal 20 can be explained from the perspective of antisymmetry [cf. Cinque 2005]. The fact seems to be that 'Dem Num Adj' order is cross-linguistically found both prenominally and post-nominally, whereas 'Adj Num Dem' order is found only postnominally. In effect, a bit as with Carstens's point about serial verbs from section 3.1 [in the full article], the stability of 'Dem Num Adj' order relative to the position of the noun gives us a window into UG that allows us to see that that order can be achieved independently of N/NP movement, whereas the order 'Adj Num Dem' cannot be. (A symmetric view of syntax would lead to the incorrect expectation that the two orders should be equally available.)

Some modifications

[. . .]

Consider the following Greenbergian universal, which is exceptionless, according to Hawkins (1990: 225) and Dryer (1992: 102):

(10)   If a language is complementizer-final, then the language is OV.

[. . .]

In effect, (10) says that if C follows IP, then the normal order of the language must have IP-C preceding V and never '*V-IP-C.' Why, however, should the internal order within CP be in any way universally keyed to the order between that CP and the matrix V (an 'external' property of CP). [. . .] I think that the answer lies in the realization that we have to give up the idea that CP is a constituent of the familiar type.

The same holds for PP, I think, once we distinguish the 'functional' adpositions from the 'lexical' (nominal) ones. As a first approximation, I take the functional adpositions to correspond to English *to, at, from, by, with, for, of*. Now Dryer (1992: 83) notes that the following is largely (though not completely) true:

(11)   If a language is postpositional, then it is OV.

Let me strengthen this to:

(12)   If the functional adpositions of a language are postpositions, then that language is OV.

which I conjecture to be exceptionless.

A strong interpretation of (12) has it that there can be no language whose normal word order is:

(13)   *V DP P

where DP is what we call the object of P and P is a functional adposition.

As in the case of complementizers, what we have here is a correlation between what looks like an internal property of PP (whether P follows DP or not) and an external property of PP (whether it precedes or follows the matrix V). Again, I think that we can achieve a satisfactory account of this correlation only if we give up the idea that PP is a constituent of the familiar type.

The proposal is, first, that (functional) P is not merged directly with its 'object,' but is rather merged outside VP. And second, that P is typically paired with a head K (for Kase) that is visible in some languages and is also merged outside VP (but below P). A (simplified) derivation for *John was looking at us* would be as in:

(14)   . . . looking us → merger of K
        . . . K looking us → movement of DP to Spec,K
        . . . us$_i$ K looking t_i → merger of P
        . . . at [us$_i$ K looking t_i] → movement of VP to Spec,P
        . . . [looking t_i]$_j$ at [us$_i$ K t_j]

This derivation produces the correct word order for an English-like language, without having 'at us' be a constituent.

4.3   (Remnant) VP-Movement

The derivation in (14) involves remnant VP-movement in the last step. Although the particular use I am putting it to here is perhaps unfamiliar, VP-movement per

se is a long-standing part of our understanding of syntax. In English, we have sentences like:

(15)   . . . and do it he will.

These are well-known. Less well-known about English is that it allows remnant movement in this construction. As background, note:

(16)   I predicted that John would marry Susan, and marry Susan/her/*Ann he will.

The argument(s) in the preposed VP must not be 'new information.' With this in mind, I find the following acceptable (with a 'coreferential' interpretation):

(17)   I predicted that John would marry Susan, and marry he will the woman he loves.

Stranding a non-heavy object is not possible:

(18)   *. . . and marry he will Susan/her.

Thus, (17) is an instance of remnant VP-movement. The 'heavy' NP/DP *the woman he loves* has been moved out of the VP prior to the VP being fronted. (Recall from section 3.3 that 'heavy'-NP shift is itself a fronting (leftward movement) operation.) The derivation of (17) (simplified) will look like:

(19)   marry the woman he loves $\rightarrow$ heavy-NP shift
the woman he loves$_i$ [marry t_i] $\rightarrow$ merger(s)
he will the woman he loves$_i$ [marry t_i] $\rightarrow$ VP-preposing
[marry t_i]$_j$ he will the woman he loves$_i$ t_j

That 'heavy'-NP shift is involved in the derivation of (17) is supported by the fact that (17) is subject to two well-known restrictions holding of heavy-NP shift in general:

(20)   *?I predicted that John would look at Susan and look at he will the woman he loves.

(21)   *I predicted that John would send Susan some flowers and send some (flowers) he will the woman he loves.

In addition to the preposition stranding restriction seen in (20), we see in (21) that a prepositionless indirect object is not subject to heavy-NP shift.

The problem with (21) is not that the preposed VP contains a visible argument, since we do have:

(22)   I predicted that John would send Susan those flowers, and send her he will the flowers he loves.

(23)   I predicted that John would speak French to Susan, and speak French he
        will to the woman he loves.

In my English, the preposition restriction in (20) is somewhat weak, but gets much
stronger in the presence of a direct object:

(24)   *I predicted that John would introduce you to Susan, and introduce you to
        he will the woman he loves.

As expected, this distinction matches my judgments on 'simple' heavy-NP shift
sentences:

(25)   *?John will look at tomorrow the article you just sent him.

(26)   *John will introduce you to tomorrow the woman he loves.

Remnant VP-movement is thus found in all of (14), (17), (22) and (23), though the
landing sites for the moved VP are not uniform.
        [. . .]

4.6   Non-prepositional complementizers

The stranding seen in (33) has no counterpart with non-prepositional comple-
mentizers like English *that* or *if*:

(40)   *They predicted that he'd be happy that he could help us, and he can help
        us he'll be happy that.

(41)   *They predicted that he wouldn't be sure if he could help us, and he can/
        could help us he won't be sure if.

These contrast sharply with:

(42)   (?)They predicted that he'd be happy to help us, and help us he'll be happy
        to.

[. . .]

   Relative clause complementizers of the English sort, which precede their asso-
ciated IP, like English *that*, were given in Kayne ([2000]: chap. 15) a derivation
similar to that given above in (14) for prepositions but without K. *That* is merged
outside VP. If we extend this to non-relative sentential *that*, and maximize uni-
formity by associating a K with *that*, too, we would have derivations like:

(43)   . . . think they're smart $\rightarrow$ merger of K_{fin}
        . . . K_{fin} think they're smart $\rightarrow$ movement of IP to Spec,K_{fin}

... [they're smart]$_i$ K$_{fin}$ think t_i → merger of *that*
... that [they're smart]$_i$ K$_{fin}$ think t_i → movement of VP to Spec, *that*
... [think t_i]$_j$ that [they're smart]$_i$ K$_{fin}$ t_j

[...]

This, however, leaves open the contrast (with respect to stranding of *to* vs. *that*) between [...] (42) on the one hand and (40) on the other. The solution, I think, is to go back to Rosenbaum's (1967) idea that *that*-clauses are introduced by *it*, and to say that contrary to the impression given by (43), the finite IP *they're smart* is not, and could not by itself be, an argument of the matrix verb *think*. More generally put:

(44)   A finite IP cannot be the argument of a higher predicate.

Let me take this to lead to:

(45)   For an IP to function as the argument of a higher predicate, it must be nominalized.

The idea is that this can happen in one of two ways, broadly speaking – either through nominalizing morphology or through merger with a noun. (These are very likely just two variants of what is essentially one strategy, insofar as nominalizing morphemes are 'bound' variants of 'free' nouns.)

In English, this nominalizing morphology can be of the derived nominal *(-ion)* type, or of the gerundial *(-ing)* type or of the (null) infinitival type. It is probably the case, as a first approximation, that the -*ion*-type is always further embedded under a D, that the -*ing*-type usually is, and that infinitives often are not.

In the absence of nominalizing morphology, i.e. when the verb is finite, the IP must be embedded under a non-affixal noun, which can be *fact*, in which case further embedding under (definite) D is required:

(46)   John mentioned *(the)/*a fact that Mary was away.

This might hold in general, i.e. even when the noun that the finite IP is embedded under is not pronounced. In other words, in:

(47)   John thinks that Mary is away.

there might be both an unpronounced N (imposed by (45)) and an unpronounced D. Yet (46) and (47) diverge when it comes to adjectives:

(48)   John is aware *(of) the fact that they're away.

(49)   John is aware (*of) that they're away.

Let me tentatively take this to suggest than an unpronounced N need not be further embedded under D, that it is not in (49), and that the Case requirement that imposes *of* on (48) is keyed to D rather than to N.

The derivation given in (43) needs to be revised to take into account the presence of this N (and in some cases D, which I will set aside here). The new derivation containing unpronounced N will begin:

(50)   . . . think N they're smart $\rightarrow$ merger of K_{fin}
      . . . K_{fin} think N they're smart

The next step is movement to Spec,K_{fin}. But now a question arises that did not arise earlier (in (43)), namely whether to move IP or NP. Let us entertain the hypothesis that what moves at this point is NP:

(51)   . . . K_{fin} think N they're smart $\rightarrow$ movement of NP to Spec,K_{fin}
      . . . [N they're smart]$_i$ K_{fin} think t_i $\rightarrow$ merger of *that*
      . . . that [N they're smart]$_i$ K_{fin} think t_i $\rightarrow$ movement of VP to Spec, *that*
      . . . [think t_i]$_j$ that [N they're smart]$_i$ K_{fin} t_j

This derivation may help in understanding why complementizer *that* looks like demonstrative *that* (especially if Sportiche (2002) is on the right track), though there is no immediate generalization to the complementizer *that* of English relative clauses. On the other hand, perhaps there should not be one, given languages like German, which have a sentential complementizer (*dass*) that looks like a demonstrative/definite article (*das*) yet does not occur in relatives in the way that *that* does in English.

Consider now the result of replacing *think* in (50)/(51) by a verb-particle combination. The relevant sentence is:

(52)   John pointed out that they're smart.

Assume that the particle *out* is within the matrix VP and that it (therefore) enters the derivation prior to K_{fin}:

(53)   . . . point [N they're smart] out $\rightarrow$ merger of K_{fin}
      . . . K_{fin} point [N they're smart] out $\rightarrow$ movement of NP to Spec,K_{fin}
      . . . [N they're smart]$_i$ K_{fin} point t_i out $\rightarrow$ merger of *that*
      . . . that [N they're smart]$_i$ K_{fin} point t_i out $\rightarrow$ movement of VP to Spec, *that*
      . . . [point t_i out]$_j$ that [N they're smart]$_i$ K_{fin} t_j

In the last step, *out* is carried along by VP-movement, yielding (52). (This is independent of the decision to have NP rather than IP move to Spec,K_{fin}.)

Since the movement of VP carries along the particle *out*, there is in fact no way, given our general approach plus the assumption that *out* is within VP, to derive:

(54)   *John pointed that they're smart out.

(More precisely put, there is no way to derive (54) without introducing the sort of scrambling operation that Kayne (1998) argued to be unavailable with English par-

ticles.) In other words, we have a possible grammatical account of (54) (independent of any processing considerations).

[. . .]

Returning now to (40) (and by extension (41)), we are now in a position to attribute to (40) the presence of an unpronounced non-affixal N just as in (49), in which case we can correlate the deviance of (A-bar-type) IP-preposing in (40) with the presence of that N (thinking perhaps of the complex-NP constraint). If, furthermore, we agree that infinitives, by virtue of being an instance of nominalizing morphology, do not systematically require embedding under non-affixal N, and more specifically that (42) does not contain one, we might then be able to account for the relative well-formedness of (42), since (42) will not run afoul of any complex-NP-like constraint (assuming the nominalizing morphology itself not to act like the non-affixal N of (40)).

[. . .]

Conclusion

Antisymmetry impinges on the analysis of Japanese in many ways, some of which I have touched on here. I have in addition argued that the solidity of the antisymmetry hypothesis is enhanced by a wide range of cross-linguistic gaps, i.e. imaginable language types that appear never to occur. I have argued further that adpositions are paired with a K-head and that DP-movement to Spec,K and VP-movement to Spec, P/P′ cut across prepositional and postpositional languages, in a way quite different from Kayne (1994). Complementizers are interestingly similar to adpositions (and deserve more space than I have given them here.)

Note

15 Travis (1989) was perhaps the first to call attention to the importance of 'inconsistent' languages.

To mention just one little cited example, Western Shoshoni (Crum and Dayley (1993)) is an OV language with postpositions that has an initial *if* (p. 186), as well as various 2nd-pos. elements (that almost certainly reflect the presence of an 'initial' head) and postnominal relatives (that are non-finite).

See also Julien (2002; 2003).

33.3   Questions pertaining to Kayne (2003)

1 To what extent does Kayne's S-H-C proposal translate into an S-V-O proposal?
2 In what sense does evidence bearing on questions of Japanese syntax have to (or not have to) come from within Japanese?

3   To what extent does "S O Aux V" make the same point, relative to the question where objects end up, as "S O Neg V"?

4   To what extent does it matter if a language has "S O Aux V" (or "S O Neg V") some of the time, as opposed to all of the time? (Extra credit: Find three languages that have "S O Aux V" at least some of the time. Partial hint: look at Tamrazian 1994.)

5   What exactly does it mean to say, for example, that there is no "mirror-image English"?

6   In what way is the term "head of a relative" misleading? How may the use of that term have affected discussions of whether languages are or are not consistent in their "directionality"? Bring in Aldridge (2003) and Kornfilt (2005).

7   Using Brody (1990) and Whitman (2001), discuss the viability of the proposal that Japanese *wa* and *ga* are high functional heads in the sentential skeleton, and that *wa*, for example, may be a Top^0 in Rizzi's (1997) sense.

8   What must the proper understanding of the term "OV" be if antisymmetry is correct? Bring in Aboh (2004).

9   What kind of parameter might underlie the fact that some languages lack "heavy- NP-shift"? To what extent is this the same question as for right-dislocation?

10   Diesing (1992) argues that indefinites remain in situ within VP, while definites move out. To what extent is her argument compatible with the alternative claim that both indefinites and definites move but indefinites move less high than definites? (Extra credit: How is Jayaseelan 2010 relevant?)

11   Find one example of a crosslinguistic asymmetry not mentioned in this paper.

12   Dryer (1992, p.83), suggests that it is by and large true that "If a language is postpositional, then it is OV." Kayne suggests recasting this as "If the functional adpositions of a language are postpositions, then that language is OV," which he conjectures to be exceptionless. How would you propose characterizing the notion "functional adposition"? How might this tie in to Kayne's (2008) proposal that there's really only one core categorial distinction, in effect matching that between nouns and verbs?

13   Kayne accepts sentences like *I predicted that John would marry Susan, and marry he will the woman he loves*, with an anaphoric link between *Susan* and *the woman*. . . Some English speakers appear not to accept such sentences. What might the parameter(s) in question be?

14   There is a distinction between the (very) marginal ??*John will look at tomorrow the article you just sent him today* and the impossible *John will introduce you to tomorrow the woman he just met today*. Against the background of Larson (1988, 1990) and den Dikken (1995), how might one relate this contrast to a property of English pseudo-passives?

15   Kayne suggests that "For an IP to function as the argument of a higher predicate, it must be nominalized." How does this compare to Kayne's (to appear) later ideas on sentential complementation?

16   Kayne discusses the impossibility of *John pointed that they're smart out*. Thinking of Rosenbaum's (1967) idea concerning a deleted *it* and Kayne's

(2006) idea that silent (deleted) elements must move to a special position, how might one establish a link with *the problem that they pointed the cause of out*?

17 To what extent can Pearson (2000) on "VO" be transposed to "OV"?

18 What might be the consequences of antisymmetry constituting an "imperfection" in Chomsky's (2001) sense? To what extent is the notion of "imperfection" in this sense testable?

19 In a footnote, Kayne conjectures that "NV incorporation (with that order) is never found in V-initial languages." Think of one further conjecture concerning incorporation and discuss how it (and Kayne's) might bear on whether incorporation is head movement or phrasal movement.

20 Find one language that has been called "consistent in directionality" and show that it is not.

21 Some language families such as Indo-European, Austronesian, and Semitic contain some or many verb-initial languages. Other families such as Niger-Congo appear not to contain any verb-initial languages. Find at least one other family like Niger-Congo. What kind of parameter(s) might be involved? (Hint: Look at Coon 2010; Emonds 1980; Lee 2000; and Massam 2000.)

References

Aboh, E. O. 1999. *From the Syntax of Gungbe to the Grammar of Gbe*. Sierre, Switzerland: Editions à la Carte SA.

Aboh, E. O. 2004. *The Morphosyntax of Complement–Head Sequences: Clause Structure and Word Order Patterns in Kwa*. New York and Oxford: Oxford University Press.

Aldridge, E. 2003. Remnant movement in Tagalog relative clause formation. *Linguistic Inquiry* 34: 631–640.

Baker, M. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Barbiers, S. 2000. The right periphery in SOV languages: English and Dutch. In P. Svenonius (ed.), *The Derivation of VO and OV*, Amsterdam: John Benjamins, pp. 181–218.

den Besten, H., and G. Webelhuth. 1990. Stranding. In G. Grewendorf and W. Sternefeld (eds), *Scrambling and Barriers. Linguistik Aktuell/Linguistics Today 5*, Amsterdam and Philadelphia: John Benjamins, pp. 77–92.

Brody, M. 1990. Remarks on the order of elements in the Hungarian focus field. In I. Kenesei (ed.), *Structures and Arguments. Approaches to Hungarian 3*, Szeged: JATE, pp. 95–121.

Chomsky, N. 1995. *The Minimalist Program*. MIT Press, Cambridge, MA.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, Cambridge, MA: MIT Press, pp. 1–52.

Chomsky, N. 2004. Beyond explanatory adequacy. In A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures 3*, New York and Oxford: Oxford University Press, pp. 104–131.

Cinque, G. 1996. The antisymmetric programme: Theoretical and typological implications. *Journal of Linguistics* 32: 447–464.

Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*, Oxford University Press, New York.

Cinque, G. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36: 315–332.

Coon, J. 2010. VOS as Predicate Fronting in Chol. *Lingua* 120: 354–378.

Crum, B., and J. Dayley. 1993. *Western Shoshoni Grammar. Occasional Papers and Monographs in Cultural Anthropology and Linguistics 1*, Boise, ID: Boise State University.

Dahl, Ö. 1979. Typology of Sentence Negation. *Linguistics* 17: 79–106.

Di Sciullo, A. M., and C. Boeckx (eds). 2011. *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford Studies in Biolinguistics. Oxford and New York: Oxford University Press.

Diesing, M. 1992. *Indefinites*. Cambridge, MA: MIT Press.

den Dikken, M. 1995. Extraposition as intraposition, and the syntax of English tag questions. Ms., Vrije Universiteit Amsterdam / HIL.

Dryer, M. S. 1988. Universals of negative position. In M. Hammond, E. Moravcsik, and J. Wirth (eds), *Studies in Syntactic Typology*, John Benjamins, Amsterdam, pp. 93–124,

Dryer, M. 1992. The Greenbergian word order correlations. *Language* 68: 81–138.

Emonds, J. E. 1980. Word order in generative grammar. *Journal of Linguistic Research* 1: 33–54.

Greenberg, J. H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg, *Universals of Language* (2nd edn), MIT Press, Cambridge, MA, pp. 73–113.

Haegeman, L. 2001. Antisymmetry and verb-final order in West Flemish. *The Journal of Comparative Germanic linguistics* 3: 207–232.

Haegeman, L. 2002. West Flemish negation and the derivation of SOV order in West Germanic. *Nordic Journal of Linguistics* 25: 154–189

Hawkins, J. A. 1990. A parsing theory of word order universals. *Linguistic Inquiry* 21: 223–261.

Jayaseelan, K. A. 2010. Stacking, stranding, and pied-piping: A proposal about word order. *Syntax* 13: 298–330.

Jayaseelan, K. A. 2001. IP-internal topic and focus phrases. *Studia Linguistica* 55: 39–75.

Jonas, D., and J. D. Bobaljik. 1993. Specs for subjects: The role of TP in Icelandic. In J. D. Bobaljik and C. Phillips (eds), *MIT Working Papers in Linguistics* 18: *Papers on Case & Agreement* 1, pp. 59–98.

Julien, M. 2002. *Syntactic Heads and Word Formation*. Oxford University Press, New York.

Julien, M. 2003. Word order type and syntactic structure. In J. Rooryck and P. Pica (eds), *Linguistic Variation Yearbook 1*, John Benjamins, Amsterdam, pp. 19–61,

Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. S. 1998. Overt vs. Covert Movement. *Syntax* 1: 128–191 [Reprinted in Kayne (2000).]

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2006. On parameters and on principles of pronunciation. In H. Broekhuis, N. Corver, R. Huibregts, U. Kleinhenz, and J. Koster (eds), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, Berlin and New York: Mouton de Gruyter, pp. 289–299. [Reprinted in Kayne (2010, Ch. 4).]

Kayne, R. S. 2008. Antisymmetry and the lexicon. In J. Rooryck and J. van Craenenbroeck (eds), *Linguistic Variation Yearbook 8*, Amsterdam and Philadelphia: John Benjamins, pp. 1–31. [Reprinted in Kayne (2010, Ch. 9) and in Di Sciullo and Boeckx (2011, Ch. 14).]

Kayne, R. S. 2010a. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2011. Why are there no directionality parameters? In M. Byram Washburn, K. McKinney-Bock, E. Varis, A. Sawyer, and B. Tomaszewicz (eds), *Proceedings of the 28th West Coast Conference on Formal Linguistics*, Somerville, MA: Cascadilla Proceedings Project, pp. 1–23.

Kayne, R. S. to appear. Why isn't *this* a complementizer? In P. Svenonius (ed.), *Functional Structure from Top to Toe: A Festschrift for Tarald Taraldsen*. New York and Oxford: Oxford University Press. [Reprinted in Kayne (2010b, Ch. 10).].

Kornfilt, J. 2005. Agreement and its placement in Turkic nonsubject relative clauses. In G. Cinque and R. S. Kayne (eds), *The Oxford Handbook of Comparative Syntax*, New York: Oxford University Press, pp. 513–541.

Kroch, A. S. 2001. Syntactic Change. In M. Baltin and C. Collins (eds), *The Handbook of Contemporary Syntactic Theory*, Oxford: Blackwell, pp. 699–729.

Larson, R. 1988. On the double object construction. *Linguistic Inquiry* 19: 335–391.

Larson, R. 1990. Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry* 21: 589–632.

Lee, F. 2000. VP remnant movement and VSO in Quiavini Zapotec. In A. Carnie and E. Guilfoyle (eds), *The Syntax of Verb Initial Languages*, New York and Oxford: Oxford University Press, pp. 143–162.

Massam, D. 2000. VSO and VOS. Aspects of Niuean word order. In A. Carnie and E. Guilfoyle (eds), *The Syntax of Verb Initial Languages*, New York and Oxford: Oxford University Press, pp. 97–117.

Müller, G. 1998. *Incomplete Category Fronting*. Dordrecht: Kluwer.

Nilsen, Ø. 2003. *Eliminating positions: The syntax and semantics of sentence modification*. Doctoral dissertation, University of Utrecht.

Pearson, M. 2000. Two types of VO languages. In P. Svenonius (ed.), *The Derivation of VO and OV*, Amsterdam and Philadelphia: John Benjamins, pp. 327–363.

Pollock, J.-Y. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20: 365–424.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, pp. 281–337.

Rosenbaum, P. S. 1967. *The Grammar of English Predicate Complement Constructions*, Cambridge, MA: MIT Press.

Sportiche, D. 2002. Movement Types and Triggers. *GLOW Newsletter* 48: 116–117.

Tamrazian, A. 1994. *Chains and the auxiliary*. Doctoral dissertation, University College London.

Travis, L. 1989. Parameters of phrase structure. In M. R. Baltin and A. S. Kroch (eds), *Alternative Conceptions of Phrase Structure*, Chicago and London: University of Chicago Press, pp. 263–279.

Whitman, J. 2001. Kayne 1994: p.143, fn. 3. In G. Alexandrova and O. Arnaudova (eds), *The Minimalist Parameter*, Amsterdam: John Benjamins., pp. 77–100.

Whitman, J. 2005. Preverbal elements in Korean and Japanese. In G. Cinque and R. S. Kayne (eds), *The Oxford Handbook of Comparative Syntax*, New York: Oxford University Press, pp. 880–902.

34

"Restructuring" and Functional Structure

Guglielmo Cinque

2004

34.1 Introduction

It is common for infinitives to serve as the complement of a verb. Interestingly, as first observed in Bech (1955) for German, when they do, they do not always display the same set of properties: some infinitival complements behave like distinct clauses, while others do not. This behavior was first explored, within the generative framework, by Evers (1975) in Dutch and German, and by Aissen and Perlmutter (1976) and Rizzi (1976[a], 1978) in Romance. These works noted that, while certain infinitival clauses create boundaries for processes that apply at the clausal level, others do not, as if the boundary between the matrix verb and its infinitival complement were transparent. For instance, while clitic placement is usually clause-bounded in Romance, infinitival clauses that are the complement of certain matrix verbs allow clitic pronouns from the embedded clause to occur to the left of the matrix verb. Rizzi (1978) proposed that the matrix verbs in question (certain modal, aspectual, and motion verbs) undergo a process of RESTRUCTURING, whereby the matrix and the embedded verb are reanalyzed as a single verbal complex, turning a biclausal structure (1a) into a monoclausal structure (1b):

(1) a. Gianni deve [presentare la a Francesco].
         Gianni must [show   it to Francesco]
         'Gianni must show it to Francesco.'
     b. Gianni la [deve presentare] a Francesco.
         Gianni it [must show]   to Francesco
         'Gianni must show it to Francesco.'

The nature of what have come to be known as "restructuring" or "transparency" effects was later tackled in a number of other works, some advocating the presence of a biclausal structure, e.g. Burzio (1986) and Kayne (1989), others the existence of more than one type of categorization on the part of the matrix verb, e.g. Wurmbrand (2001). We refer the reader to footnote 1 of Cinque's article for an extensive bibliographical overview of the different types of approaches that have been adopted over the course of the last forty years.

Cinque's "*Restructuring" and Functional Structure* revisits the issue of what gives rise to restructuring effects with the goals of understanding why only a certain class of infinitival complements should fail to create a clause boundary, and why these effects should be exhibited by infinitives that are complements to a particular set of verb classes – namely modals, aspectuals, and motion verbs – across languages. Cinque's proposal is these so-called "restructuring verbs" are the realization of functional categories in the extended projection of a lexical verb. In other words, the modal, aspectual, or motion verb and the infinitival complement behave like a single clause because they are indeed part of a single clause.

One piece of evidence in favor of this view is that, when two restructuring verbs co-occur and the infinitival complement exhibits transparency effects, their relative order is quite rigid. We can see this in the examples in (2), where only the order in (2a) yields a grammatical sentence:

(2)   a.   Lo *tenderebbe   a voler*   fare   sempre lui.
           it   would–tend to want(INF) do(INF) always he
           'He'd tend to want to always do it himself.'
      b. *Lo *vorrebbe   tendere*   a fare   sempre lui.
           it   would–want tend(INF) to do(INF) always he
           'He'd want to tend to always do it himself.'

Cinque assumes that functional heads reflect a rigidly ordered hierarchy of functional projections, as proposed in Cinque (1999). If the two restructuring verbs are functional heads (*tendere* related to so-called predispositional aspect and *volere* to volitional modality), it follows that their relative order is rigidly ordered.

Cinque further argues that a subset of the restructuring verbs – *volere* 'want' and some of the aspectual verbs – are always functional categories, even in sentences like (3a) where there is no other overt verb. Following proposals by McCawley (1974), Ross (1976), and den Dikken et al. (1996), he analyzes such sentences as containing a nonovert lexical verb HAVE (or COME TO HAVE), as in (3b):

(3)   a.   Gianni vuole una bicicletta.
           Gianni wants a   bicycle
      b.   Gianni vuole HAVE/ COME TO HAVE una bicicletta.
           Gianni wants HAVE/ COME TO HAVE a   bicycle

Other restructuring verbs – including *sembrare*, 'seem', and the motion verbs – have both a functional and a lexical usage.

In sum, this article approaches the much debated issue of how to capture the differences among infinitival complements by combining a detailed study of the

nature of restructuring verbs with the fundamental idea of what has become known as the "cartographic approach" to syntax, namely the idea that Universal Grammar provides a richly articulated structure of the clause, which consists of a rigidly ordered set of functional heads.

34.2   From "'RESTRUCTURING' AND FUNCTIONAL STRUCTURE"

1   Introduction

In what follows I would like to show how the articulated functional structure of the clause suggested in Cinque (1999) may shed new light on the "restructuring" phenomenon (Rizzi 1976a,b; 1978) and perhaps afford a deeper understanding of it.

In the past twenty-five years, numerous analyses have been proposed to explain why certain phenomena that are otherwise clause-bound [such as Clitic Placement – see (1)] appear to be able to span over two clauses when the matrix verb is either a *modal*, an *aspectual*, or a *motion* verb and the complement is nonfinite [see the "climbing" of the clitic in (2)]:

(1)   a.   *Lo detesto [vedere t in quello stato] '(I) him detest seeing in that state'
      b.   *Lo ammetto [di conoscere t appena] '(I) him admit to barely know'
      c.   *Lo rinuncio [ad avere t per me] '(I) it give up having for me'
(2)   a.   Lo volevo [vedere t subito] '(I) him wanted to see immediately' (*modal*)
      b.   Lo finisco [di vedere t domani] '(I) it finish to see tomorrow' (*aspectual*)
      c.   Lo vengo [a prendere t domani] '(I) it come to fetch tomorrow' (*motion*)

Even if each of the proposed analyses captures one or another aspect of restructuring, it is fair to say that none of them manages to answer the two most basic questions that the phenomenon raises; namely, why it should exist at all, and why it should exist with those particular verb classes (modal, aspectual, and motion). The fact that one finds transparency phenomena comparable to Clitic Climbing language after language, and with the same set of verbs (or subsets thereof), suggests that the phenomenon is universal and should thus follow from some general property of UG. Here I would like to propose an analysis that derives its universality and answers at the same time the two basic questions just mentioned. The analysis is a natural extension of proposals made in Cinque (1999), where, on the basis of the relative order of functional morphemes in head position and of the corresponding classes of AdvPs, I suggested that the functional portion of the clause, in all languages, is constituted by the same, richly articulated and rigidly ordered, hierarchy of functional projections, a subset of which is shown in (3);

(3) $\text{MoodP}_{\text{speech act}} > \text{MoodP}_{\text{evaluative}} > \text{MoodP}_{\text{evidential}} > \text{ModP}_{\text{epistemic}} > \text{TP(Past)} > \text{TP(Future)} > \text{MoodP}_{\text{irrealis}} > \text{ModP}_{\text{alethic}} > \text{AspP}_{\text{habitual}} > \text{AspP}_{\text{repetitive(I)}} > \text{AspP}_{\text{frequentative(I)}} > \text{ModP}_{\text{volitional}} > \text{AspP}_{\text{celerative(I)}} > \text{TP(Anterior)} > \text{AspP}_{\text{terminative}} > \text{AspP}_{\text{continuative}} > \text{AspP}_{\text{retrospective}} > \text{AspP}_{\text{proximative}} > \text{AspP}_{\text{durative}} > \text{AspP}_{\text{generic/progressive}} > \text{AspP}_{\text{prospective}} > \text{ModP}_{\text{obligation}} > \text{ModP}_{\text{permission/ability}} > \text{AspP}_{\text{Completive}} > \text{VoiceP} > \text{AspP}_{\text{celerative(II)}} > \text{AspP}_{\text{repetitive(II)}} > \text{AspP}_{\text{frequentative(II)}}$

The verbs that enter the restructuring construction appear to correspond to distinct heads of (3), in the sense that each seems to lexicalize the content of one or another functional head. This is obvious for the various modal and aspectual verbs, but it is true for motion verbs as well.

In previous work (Cinque 2001, [2002], originally written and circulated in 1997, and Cinque 1998), I had suggested that this striking correspondence rendered the following hypothesis appealing: only those verbs that happen to match semantically the content of a certain functional head admit of two distinct possibilities. They are either regular verbs, heading a VP [in which case they take a fulfledged [sic] sentential complement (CP) – cf. (4a)], or functional verbs, directly inserted in the head position of the corresponding functional projection (cf. (4b):

(4) a. $[_{CP} \ldots [_{FP} \ldots [_{FP} \ldots [_{VP} V_{restr} [_{CP} \ldots [_{FP} \ldots [_{FP} \ldots [_{VP} V]]]]]]]]]$
     b. $[_{CP} \ldots [_{FP} \ldots [_{FP} V_{restr} [_{FP} \ldots [_{VP} V]]]]]$

Following the received opinion, I had also assumed that the presence or absence of transparency effects reduced to two mutually exclusive options: the obligatory *presence* of transparency effects in the monoclausal structure (4b) and the obligatory *absence* of transparency effects in the biclausal structure (4a).

Here, after arguing that *when transparency effects obtain*, "restructuring" verbs are functional verbs in a monoclausal configuration (sections 2–5), I will explore the stronger and at first sight more difficult claim that they are *always* functional verbs in a monoclausal configuration (*even in the variant that shows no transparency effects* – section 6). This implies that restructuring verbs have no other option but to enter structure (4b) (ultimately, a consequence of their corresponding to the semantic content of a distinct functional head). This also requires interpreting the differences between the variant with and the variant without transparency effects in a different manner (section 7).

2   The constituency issue

The analysis whereby, when transparency effects obtain, restructuring verbs are functional verbs (directly inserted under the corresponding functional heads) leads one to expect a constituent structure quite different from that of Rizzi (1976a, 1978). According to Rizzi's analysis, modal, aspectual and motion verbs can trigger a process of structural simplification (Restructuring), which turns an original biclausal configuration into a monoclausal one, forming a complex verb out of the complement and matrix verbs, as shown in (5):

(5) a. $[_{CP}$ io [verrò $[_{CP}$ a parlarti di questi problemi]]] RESTRUCTURING $\rightarrow$
         (I will come to talk-to-you about these problems.)
     b. $[_{CP}$ io $[_V$ ti verrò a parlare] di questi problemi]

As a result of this complex verb formation, the embedded verb is taken to no longer form a constituent with its own complements [cf. (5b)].

In the present analysis, instead, the expected constituent structure is (6), with the embedded verb still forming a constituent with its complements:

(6)   [_CP io [_AndativeP ti verrò [_VP a parlare [di questi problemi]]]]

This requires reassessing the arguments brought forth by Rizzi (1976a, 1978) in support of the constituency in (5b). He shows, for example, that when transparency effects obtain a number of operations apparently cease to apply to the sequence formed by the embedded verb and its complements, taking this to support the derived structure (5b). Let us consider these cases in turn.

2.1 Cleft Sentence Formation

As shown by the contrast between (7a and b), when the clitic has climbed to the matrix verb the embedded verb cannot be clefted together with its complement:

(7)   a.   E′ proprio a parlarti di questi problemi che verrà
          'It's just to talk to-you about these problems that he'll come.'
      b.   *E′ proprio a parlare di questi problemi che ti verrà

This would seem to follow from the constituency in (5b). Notice, however, that with other fronting rules (such as Focus Movement and Topicalization) no such restriction obtains:

(8)   a.   A PARLARE DEI SUOI PROBLEMI, ti verrà! Vedrai.
          'To speak about his problems (focus), he'll to-you come! You'll see'
      b.   PORTARE A CASA, lo voleva!   'Take home (focus), he it wanted'
      c.   Leggere a tutti, non lo potevo 'Read to everybody, I it couldn't'

As the latter constructions are no less valid constituency diagnostics than Cleft Sentence Formation, we must conclude that the embedded verb *does* form a constituent with its complement, just as (6) implies, and that the ungrammaticality of (7b) is due to some other reason (not dependent on constituency).
   [. . .]

2.2 Right node raising

As Rizzi (1976a, 1978) also notes, the embedded verb and its complement can be Right Node Raised only in the absence of tra[n]sparency effects. See the contrast between (13a and b):

(13) a.   Piero voleva – ma francamente adesso non so se vorrà ancora – parlarne con Gianni
          'P. wanted to – but frankly now I don't know if he still will – speak about it with G.'
      b.   *Piero ne voleva – ma francamente adesso non so se ne vorrà ancora – parlare con Gianni
          'P. about it wanted to – but frankly now I don't know if he still will – speak about it with G.'

Once again this would seem to follow from the assumption that in the presence of Clitic Climbing the embedded verb and its complement do not form a constituent. But this conclusion is not necessary. Another possibility exists, which is compatible with the idea that the embedded verb continues to form a constituent with its complements.

In the framework in which Rizzi (1976a, 1978) was working, Right Node Raising was considered a rightward movement rule (cf. Postal 1974: 125–128). More recently, Kayne (1994: 67f.), following Wexler and Culicover (1980: 298ff.), has proposed to reinterpret it as a deletion rule deleting under identity the lefthand copy of the "raised" phrase: *Piero voleva – ma francamente adesso non so se vorrà ancora – parlarne con Gianni*. The following contrasts between Italian and English indeed appear to support Kayne's reinterpretation of Right Node Raising. As noted in Napoli (1981: 846), Right Node Raising of the complement of an auxiliary is impossible in Italian. See (14):

(14) a.   *Mario ha – ma dirà di non avere – capito la lezione 'M. has – but he
              will say he hasn't – understood the lesson'
      b.   *Gianni allora era – ma non so se ancora oggi sarebbe – apprezzato per
              il suo autoritarismo 'G, then was – but I don't know whether today still
              he would be – appreciated for his authoritarianism'

Right Node Raising of the complement of an auxiliary is instead possible in English:

(15) Tony should have – and Pete probably would have – called Grace (Postal
      1974: 126)

Now, the two languages also differ with respect to the deletion of the complement of an auxiliary, as shown in (16):

(16) a.   Have you called John? Yes. I have_____.
      b.   Hai chiamato John? *Sì. Ho_____

If Right Node Raising involves deletion, the first contrast reduces to the second. No such reduction is possible under the Movement analysis of Right Node Raising. In the more restrictive deletion analysis, which crucially relates (14) to (16b), the ungrammaticality of (13b) can, then, be attributed not to the fact that *parlare con Gianni* fails to be a constituent but to the impossibility of deleting an infinitival complement in the presence of transparency effects. See (17), noted in Radford ([1977]: 113) (whatever the right analysis of this phenomenon is; see Depiante 1998 and section 7 below):

(17) a.   Gianni voleva parlare di questo, ma Piero non (*ne) voleva_____
              G. wanted to talk about this, but P. not (about-it) wanted
      b.   *Certe cose si possono fare, ma queste non si possono_____
              Certain things one can do, but these not one can
      c.   Gianni poteva andare a casa, ma non ha/*è voluto_____
              G. could go home, but not has/is wanted

      [. . .]

3   Monoclausality versus biclausality

In this section I examine some potential evidence (in addition to that recently discussed in Wurmbrand 1998, 2001) for the monoclausal nature of the construction when transparency effects obtain (sections 3.1–3.2), and consider in section 3.3 some of the apparent evidence for its biclausality, concluding that it is unconvincing.

   [. . .]

3.2   The relative order of "restructuring" verbs

If more "restructuring" verbs occur, their relative order appears to be quite rigid when transparency effects obtain. Although this is unexpected under biclausal analyses, it is to be expected in a monoclausal one in which "restructuring" verbs are 'functional' verbs directly inserted into the corresponding functional heads. This occurs because functional heads are themselves rigidly ordered.

   [. . .]

When *tendere* and *volere* 'want' co-occur, the order is rigidly *tendere* > *volere*, in turn suggesting the order $Asp_{predispositional} > Mod_{volitional}$:

(28)   a.   Lo tenderebbe a voler fare sempre lui 'He would tend to want to always do it he himself.'
   b.   *Lo vorrebbe tendere a fare sempre lui 'He would want to tend to always do it he himself.'

Putting together the various relative orders, one arrives at the order of verbs in (31), corresponding to the order of functional heads shown in (32):

   [. . .]

(31)   *solere* > *tendere* > *volere* > *smettere* > *continuare*
(32) $Asp_{habitual} > Asp_{predispositional} > Mod_{volitional} > Asp_{terminative} > Asp_{continuative}$

3.3   Apparent cases of transparency effects across CP

A strong case for the biclausal character of restructuring would seem to come from two instances of Clitic Climbing across what looks like a CP-boundary.

   The first is already discussed in Rizzi (1978: 151f.), where such cases as (33) are noted:

(33)   a.   [. . .] non ti saprei che dire 'I you wouldn't know what to tell'
   b.   ?Mario, non lo saprei a chi affidare, [. . .] 'M., I him wouldn't know to whom to entrust'
   c.   ??[. . .] proprio, non lo saprei come risolvere 'Really, I it wouldn't know how to solve'

As Rizzi himself (n. 38) observes (cf. also Napoli 1981: 855; Moore 1994: n. 3; Rooryck 1994: 420ff.; etc.), the productivity of the construction is, however, severely limited.

   [. . .]

The generalization appears to be that Clitic Climbing is allowed across a wh-phrase with *sapere* either if *sapere* means "know how" (33c) or if the sentence allows for a rhetorical reading without the wh-phrase; with *sapere* meaning "be able" (33a) is equivalent to *Non ti saprei dire niente* 'I to-you wouldn't be able to say anything', and (33b) to *Non lo saprei affidare a nessuno* 'I him wouldn't be able to entrust to anybody').

In either case, the verb embedding a wh-phrase is interpreted as a modal of mental ability (a notion often distinguished from physical ability in the languages of the world). This makes the verb a natural candidate for direct insertion under the root modal head of ability, like other restructuring verbs, an option not open to the verbs in (34) and (35) [e.g., *chiedersi* 'to wonder', *domandarsi* 'to ask oneself', *sapere*], whose interpretation is not one of mental ability. The only auxiliary assumption that needs to be made is that the root modal head of *mental* ability can take a single wh-CP-layer above its ordinary functional XP-complement (without full recursion of the extended functional projection).

In sum, the very selective nature of Clitic Climbing across a wh-CP in mental ability contexts and the interpretation of it just sketched render the argument based on (33) for the biclausal character of restructuring very dubious. If anything, the properties of (33) point, once again, to the functional nature of the verb, a modal (and to the monoclausal character of the construction).

[. . .]

4 The functional status of restructuring verbs in the presence of transparency effects

One consequence of the idea that (when transparency effects obtain) "restructuring" verbs are "functional" verbs directly inserted under the corresponding functional heads is that, like auxiliaries (cf. Pollock 1989), they should have no thematic roles to assign, and hence no arguments of their own. Despite certain appearances, this will prove a welcome (and correct) consequence.

[. . .]

For many speakers, myself included, *sembrare* allows Clitic Climbing [cf. (41a)] but, crucially, not if it takes a (dative) complement [cf. (41b)]:

(41)   a.   Gianni non lo sembra apprezzare abbastanza 'G. does not it seem to appreciate enough'
       b.   *Gianni non ce lo sembra apprezzare abbastanza 'G. doesn't to-us it seem to appreciate enough'

A comparable contrast concerning 'long' L-*tous* in French (also found only with "restructuring" verbs) is noted in Pollock (1978: 97f.) (I thank Richard Kayne for pointing this out to me):

(42)   a.   ?Elle a tous semblé/paru les avoir lus
             She seemed/ appeared to have read them all
       b.   *Pierre m'a tous semblé/paru les avoir lus
             She seemed/appeared to-me to have read them all

These contrasts, which are very sharp, seem to indicate that it is the presence of the dative complements of 'seem,' *ce* 'to us,' *me* 'to me,' which inhibits Clitic Climbing and 'long' L-*tous*, respectively.

[. . .]

6 The functional status of "restructuring" verbs in the absence of transparency effects

So far, following the traditional opinion, I have been assuming that the presence of one or more transparency effects is an unequivocal indication of the presence of a monoclausal configuration, while the variant without transparency effects indicates a biclausal one. Given their optionality, however, the variant without transparency effects tells us nothing about sentence structure. A restructuring verb could well be functional (directly inserted under a functional head in a monoclausal configuration) even when the clitic is on the embedded verb, *loro* has not climbed, or Long Object Preposing has not applied.

This opens up the theoretical possibility that restructuring verbs are always functional, even in the absence of transparency effects.

The existence of varieties where transparency effects are obligatory (such as most Southern Italian dialects) would already seem to suggest that restructuring verbs indeed are only functional. In this section, I consider some evidence supporting this first indication. We shall see that except for *sembrare* 'seem' and motion verbs (which also have genuine lexical usages), restructuring verbs are always functional, and hence necessarily enter a monoclausal configuration. This has the conceptual advantage that such verbs do not need to be marked in the lexicon as either lexical or functional, with the ensuing problem of having to account for the complete synonymy of the two uses and for what looks like a single subcategorization option (the uniform selection of either *di* 'of,' *a* 'to,' or *0* (cf. Rizzi 1978: 150). They need only be marked as functional.

In addition to this conceptual argument, there is some empirical evidence for their exclusively functional nature (see sections 6.1–3).

6.1 More on the relative order of restructuring verbs

In section 3.2 above, we observed that restructuring verbs come in a rigid order when transparency effects obtain. The same rigidity is, however, found even in the absence of transparency effects. See (83) and (84) (and Hernanz and Rigau 1984: n. 6 for the similar rigid ordering of restructuring verbs in the absence of transparency effects in Catalan):

(83) a. Suole provare a farle/provarle a fare da solo 'He uses to try to do them by himself'
      b. *Prova a soler farle/solerle fare da solo 'He tries to use to do them by himself'
(84) a. Soleva smettere di vederla/ ?smetterla di vedere ogni sei mesi 'He used to stop seeing her every six months.'
      b. *Smetteva di soler vederla/solerla vedere ogni sei mesi 'He stopped using to see her every six months'

This suggests that such verbs are only functional. If they were (also) lexical, taking a full-fledged CP complement, it would not be clear how they could determine the choice of the verb of their sentential complement. Note that the reason for the ill-formedness of (84b) can hardly be semantic. It would make perfect sense to "stop having the habit of doing something." Yet, the sentence is unacceptable.

[. . .]

6.3   Apparent lexical usages of *volere* and aspectual verbs

The idea that restructuring verbs are always functional would seem to be contradicted by certain prima facie lexical usages of *volere* 'want' and of some of the aspectual verbs. See (87) and (88):

(87)   Gianni vuole una bicicletta 'G. wants a bicycle.'
(88)   a.   Maria ha cominciato il romanzo 'M. began the novel.'
      b.   Mario ha finito il vino 'M. finished the wine.'
      c.   Il concerto sta cominciando/sta finendo/continua
          'The concerts is beginning/finishing/continuing'

In all such cases, the verb, unlike what happens with functional verbs, does not take a nonfinite verbal complement but rather a DP, object or subject, thus apparently qualifying as a simple transitive, or unaccusative, lexical verb. The appearances, however, are misleading, as there is evidence that (87) and (88) are structurally more complex than it looks. Den Dikken et al. (1996), following earlier proposals by McCawley [1974] and Ross [1976], provide syntactic arguments that in (87) 'want' does not directly take the DP as its object but takes an abstract verbal complement, whose head, roughly paraphrasable with HAVE, takes the DP as its object:

(89)   Gianni vuole [$_{XP}$ HAVE [$_{DP}$ una bicicletta]]

If this is so, *vuole* in (87) continues to be the functional verb seen so far, with *Gianni la vuole* 'G. wants it' a case of Clitic Climbing.

Similarly, Pustejovsky (1995) and Jackendoff (1997: 60ff.) (cf. also Rochette 1999: 159ff.), in order to account for the variable, and highly restrictive, interpretations that aspectual predicates show, depending on the nature of the object, have argued that they actually select an abstract verbal complement of activity, whose head is interpreted on the basis of the *qualia structure* of the object (differently from them, I assume here that they syntactically take an abstract verbal complement).

Although such special usages of 'want' and of phasal aspectuals deserve more careful investigation, it seems that they can be rendered compatible with the idea that such verbs are exclusively functional, part of the extended projection of another, overt or abstract, lexical verb.

6.4 Restructuring and lexical usages of motion verbs and *sembrare*

The case of motion verbs and *sembrare* 'seem,' which appear to have genuine usages as lexical verbs in addition to their functional usage, is different. We have seen that when these verbs take an internal argument (either a directional PP or a subject, for the former, and a dative PP for the latter) they cease to behave as restructuring verbs (e.g., they do not allow Clitic Climbing). See, in particular, [. . .] (41) and (42), respectively.

These data are still compatible with the idea that restructuring verbs are *always* functional if, when they take a complement, motion verbs and *sembrare* are actually different verbs, in fact, genuine *lexical* verbs. This appears to be confirmed by the fact that the case with and the case without a complement display a subtle difference in meaning.

Motion verbs, when they take a complement of their own and an optional adjunct clause [. . .] are interpreted literally as verbs of locomotion, part of whose meaning is the means of transportation [cf. (90a) below]. When they are used as restructuring verbs, instead, they are not verbs of locomotion for which one can ask the means of transportation – whence the ungrammaticality of the answer to (90b), where *come* 'how' can only ask 'the way he will (come to) paint the door.' They merely indicate that some distance is traversed before the action depicted by the lexical verb is carried out:

(90) a. A: Come verrà da te a dipingere la porta?   'How will he come by you to paint the door'
          B: In bicicletta   'With his bicycle'
      b. A: Come ti verrà a dipingere la porta?   'How will he come to paint your door?'
          B: *In bicicletta   'With his bicycle.'

Similar considerations hold for *sembrare* when it takes a dative argument versus restructuring *sembrare* without one. The former literally means that a certain state of affairs seems true to someone [hence the perfectly noncontradictory status of (91a)]. The latter is instead an evidential functional verb, which (mildly) commits the speaker to the truth of a certain state of affairs [whence the contradictory status of (91b)]:

(91) a. Gianni sembra a tutti apprezzarlo molto, ma io non credo che lo apprezzi
          'G. seems to everybody to appreciate it much, but I don't believe he appreciates it.'
      b. #Gianni lo sembra apprezzare molto, ma io non credo che lo apprezzi
          'G. seems to appreciate it much, but I don't believe he appreciates it.'

This, of course, does not exclude the possibility that the functional (restructuring) usage of motion verbs, and *sembrare*, have their ultimate basis in the lexical usages of these verbs (because of their semantics). But it shows that their functional and lexical usages should be kept distinct.

7 Presence versus absence of transparency effects: Syntactic contrasts

I have argued so far that restructuring verbs are always functional, appearing in a monoclausal configuration with their infinitival complement whether or not they show transparency effects.

[. . .]

7.1 The special status of *volere, sembrare,* and motion verbs

As seen above, such cases as *Gianni vuole restare* 'G. wants to stay' are structurally ambiguous even if *volere* is exclusively functional. That depends, as seen, on the additional possibility for *volere* to be followed by an abstract verb (OBTAIN), which itself takes the infinitival phrase as a complement:

(92) a. Gianni$_i$ vuole . . . [$_{VP}$ t$_i$ restare]
   b. Gianni$_i$ vuole . . . [$_{VP}$ t$_i$ OBTAIN [$_{CP}$ PRO$_i$ restare]]

This, we take it, is at the basis of the contrasts in (23)–(26) above. These are found with *volere* but in fact with no other modal or aspectual verb.

[. . .]

8.3 Variation in the membership of restructuring verbs

The often-made observation that the membership in the class of restructuring verbs varies across languages – and, within one language, even among speakers – would seem to go against the UG approach taken here and argue for an essentially lexical approach. This impression, however, is quite misleading. Consider, first, variation across languages. The idea that restructuring verbs correspond to distinct functional heads of a universal functional hierarchy does not per se entail that all languages should have a verb (a free morpheme) corresponding to *each* such head. It could well be that a language expresses a certain functional head via a bound morpheme (say, a suffix) or via no head category at all (but rather via an AdvP, arguably in the specifier of that head). Italian, for example, appears to instantiate the latter case when compared to Spanish (or French). Spanish has a restructuring usage of *acabar de* '(lit.) finish,' which seems to correspond to the so-called Retrospective Aspect (cf. Cinque 1999: 96–98, and references cited there):

(111) Lo acabo de ver
   (Lit.) Him (I) finish to see   'I have just seen him.'

The same aspect is rendered in French by the verb *venir de* [(Lit.) 'come from' (*Je viens de le voir* 'I have just seen him')]. In Italian, however (and English, for that matter), the only way to render such an aspect is by using the AdvP *appena* 'just' (in one of its uses) combined with the verb in the perfect form: *L'ho appena visto* 'I have just seen him.' Such lexical variation among Spanish *acabar de*, French *venir*

de, and Italian 0 (or rather *appena*) is of little significance from a UG point of view. It only obscures the fact that the three languages express one and the same functional head through different morphological means. Another case in point is the restructuring verb *faillir* in French ('to almost'), which renders the grammatical notion of "action narrowly averted," variously expressed in the languages of the world (see Kuteva 1998), and to which in Italian and English no restructuring verb corresponds but, instead, an AdvP (*quasi*/*almost*).

[. . .]

Apparently more serious for a UG approach is the fact that the same verb, with essentially the same meaning, is a restructuring verb in one language but not in another or, within one and the same language, for some speakers but not others.

One example is 'seem,' which is taken not to be a restructuring verb in Spanish (Zagona 1986: 232) or in Portuguese (Quicoli 1976: 215; Pizzini 1981: 427, n. 24) but is a restructuring verb, at least for many speakers, in Italian (cf. note 27 above). Even if true, this fact is not necessarily troublesome. It could mean that in Spanish and Portuguese *parecer* has only the lexical usage seen in section 6.4 above (alternatively, it could be that the speakers Zagona and Pizzini based their conclusion on, as opposed to other speakers, are as the Italian speakers that do not have *sembrare* as a restructuring verb).

34.3 Questions pertaining to Cinque (2004)

1 Cinque's example (8c) is *Leggere a tutti, non lo potevo* ('to-read to all, neg it I-could' = 'I couldn't read it to everybody'). In it, the pronominal clitic *lo*, which originates as the object of *leggere* ('read'), is stranded to the left of the matrix modal after topicalization of the embedded infinitival VP. This kind of example directly supports a) the existence of remnant movement, as well as b) the idea that (Italian) topicalization involves (or at least can involve) movement (as opposed to external merge into Spec,Top).

   Yet English topicalization does not allow sentences like **Thinking about I didn't know what he was (though I did know what he was talking about)*. Why might that be?

2 Cinque notes that Italian does not allow a counterpart of English Right Node Raising sentences that split auxiliary from participle, as in *John hasn't, but he really should have, apologized for his mistake*. He argues strongly that this Italian–English contrast supports the idea that in the English example the participial phrase in the left-hand part of the sentence has been deleted, rather than moved, the reason being that Italian is independently known to disallow deletion of the complement of such auxiliaries. (Cinque's conclusion is further supported by Chalcraft 2006.)

   To what extent does this reasoning carry over to English sentences like *John hasn't yet bought, but he really should consider buying, the new brand of smartphone that became available last week*?

3   Cinque shows that the arguments that Rizzi (1976, 1978) had adduced in favor of "restructuring" are not compelling and that such restructuring arguably does not apply at all in Italian. Is this just a choice that Italian made, or, as seems more likely, is such restructuring actually not made available at all by the human language faculty? In the latter case, what exactly is it about restructuring in Rizzi's sense that runs afoul of UG?

4   Cinque notes an Italian counterpart of the contrast that one finds in English between *They tend to want to criticize everybody* and **They want to tend to criticize everybody*, which he interprets as indicating that *tend* corresponds to a functional head higher in his hierarchy than the functional head to which *want* corresponds. Discuss the implications of the apparently parallel contrast between *They have a tendency to want to criticize everybody* and *?*They want to have a tendency to criticize everybody*.

5   In the face of Italian examples in which a clitic climbs into the matrix across a *wh*-phrase, such as . . . *non ti saprei che dire* ('. . . neg you I-would-know what to-say' = 'I wouldn't know what to say to you'), Cinque takes such matrix verbs to be compatible with a complement containing a single *wh*-layer (hosting, in this example, *che* ('what')). How serious a weakening is this of his general claim that clitic climbing reflects monoclausality (vs. Koopman and Szabolcsi 2000)? How is monoclausality to be understood in light of Terzi (1996)? How serious is the tension between his approach here and Cinque's (1999) suggestion that all sentences in all languages contain the entire array of functional heads?

6   Italian allows clitic climbing with the matrix verb *seem* in a raising context, but not if *seem* takes a dative complement of its own (Cinque's example (41)): *Gianni non (*ce) lo sembra apprezzare abbastanza* ('John neg (*to us) it seems to-appreciate enough' = 'John doesn't seem (*to us) to appreciate it enough'). Cinque's account is that clitic climbing requires a functional matrix verb, but functional verbs cannot take (dative) complements.

     English has something that looks similar, in the contrast given in *John can't seem (*to me/us) to lift the piano*. On the other hand, the dative is possible in *It seems (to me/us) that John can't lift the piano*. Assume that the English *can't seem to* example involves (remnant) movement of 'John can't' out of the embedded sentence and into the matrix, past 'seem.' Evaluate the relative plausibility/adequacy of accounting for the English restriction against **can't seem to me/us* . . . à la Cinque and accounting for it via an intervention effect due to the dative.

7   In taking the parallelism between *John/*the house wanted to get photographed* and *John/*the house willingly got photographed* to indicate that *want* needn't have a subject theta-role associated with it (because *willingly* doesn't), Cinque implicitly sets aside an alternative view according to which *willingly* is in fact itself associated with a subject theta-role. Explore some of the consequences of such an alternative view.

8   In his section 5.1, Cinque gives strong arguments for the optionality of clitic climbing in certain cases in Italian. How might such optionality be expressed in a phase-based account of clitic climbing?

9 In 6.1, Cinque notes that the (closest) Italian counterpart of English *used to*, as in *They used to work hard*, cannot be embedded (in Italian) under either *provare* ('try') or *smettere* ('stop'). He then in effect notes (for Italian) a contrast like that between **They stopped using to work hard* and *They stopped being in the habit of working hard*. In English there is also a sharp contrast in the present tense that can be seen especially clearly under negation: *They aren't in the habit of working hard* vs. **They don't use to work hard*, despite the apparent semantic equivalence of *(?)They didn't use to work hard* and *They weren't in the habit of working hard*.

   Against this background, what might be the significance of *There used to be more problems than there are now* vs. **There was in the habit of being more problems than there are now*? (Extra credit: How does this link up to question 4 above?)

10 In 6.3, Cinque suggests that (Italian counterparts of) sentences like *Mary began the novel* must contain a syntactically present embedded verb. How much of an effect will that proposal, assuming it to be correct, have on the work on "coercion" by Pylkkänen et al. (2009)? (Extra credit: How might Kayne 2006 be used to account for the obligatory absence of a prepositional complementizer in the presence of a silent verb, as in: *Mary began to read the novel* vs. *Mary began (*to) the novel*?)

11 Cinque suggests that Italian restructuring verbs with no overt complement at all (corresponding to English *I tried*) are followed by a silent counterpart of *do it*, which seems plausible. He then goes on to account for the absence of clitic climbing with this silent *do it* (as illustrated by *Gli ho provato *(a parlare)* ('to-him I-have tried (to speak)')) in terms of the idea that *do it* lacks the internal structure to provide a source for the raised clitic (here *gli*). What might (thinking of Kayne 2002) be the relevance of the contrast in English between ?*I tried to do it / that several times, to speak to him in private* and **Him I tried to do it / that several times, to speak to in private?* (Extra credit: How exactly might this approach provide an account of *Has John begun to build his house? Yes, he has begun (to)* vs. *Has there begun to exist a solution? Yes, there has begun *(to)*?)

12 What is the significance of the fact that in Italian (and various other languages, though not English) the complement of modal verbs has an infinitival suffix *-r-* that in some ways (cf. Raposo 1987) looks nominal?

13 English lacks pronominal clitic movement of the Italian sort and therefore also lacks clitic climbing of the Italian sort (as well as auxiliary shift and long NP-movement of the Italian sort). Conversely, English modals show special behavior in ways that do not have exact counterparts in Italian. That there is nonetheless a link here between the two languages is suggested by the fact that English modals almost correspond to a proper subset of the verbs that in Italian allow clitic climbing.

   Need, though, allows (though does not require) modal-like behavior, as in *He needn't leave yet, Need we leave so soon*, without Italian having any direct counterpart with clitic climbing. Italian normally expresses such sentences with *have* + nominal *need*, as in *Non ha bisogno di chiamarci* ('neg he-has need of calling-us' = 'He needn't call us'), in which case clitic climbing is prohibited: **Non ci ha bisogno di chiamare*.

From the perspective of Cinque's analysis, how might the English contrast between *There needn't be any solution* and **There has no need to be any solution* (vs. *They have no need to talk so much*) lead to an account of the impossibility of clitic climbing with *avere bisogno* ('to-have need')?

14   Across Romance languages, one of the core clitic climbing verbs is the one corresponding to Italian *volere* ('to-want'). Yet in English the verb *want* has none of the special properties that characterize English modals. To what extent is this fact paradoxical?

15   (Extra credit) Choose one variety of English that allows double modals. To what extent does it look like Italian? To what extent do its properties weaken Chomsky's (1957) claim that English modals are not verbs?

16   How exactly does Henry's (1995) discussion of Belfast English *for* bear on Cinque's functional hierarchy?

References

Aissen, J., and D. Perlmutter. 1976. Clause reduction in Spanish. In H. Thompson et al., *Proceedings of the 2nd Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society, pp. 1–30. [Reprinted in Perlmutter (1983).]

Bech, G. 1955. *Studien zum deutschen Verbum infinitum*. Tübingen: Max Niemeyer.

Burzio, L. 1986. *Italian Syntax: A Government-Binding Approach*. Dordrecht: Reidel.

Chalcraft, F. 2006. Right node raising as ellipsis: evidence from (what the) British do. *Snippets* 12: 7–8.

Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.

Cinque, G. 1998. On clitic climbing and other transparency effects. Talk given at New York University, New York City, and Massachusetts Institute of Technology, Cambridge, February and March 1998.

Cinque, G. 1999. *Adverbs and Functional Heads*. A Cross-Linguistic Perspective. New York and Oxford: Oxford University Press.

Cinque, G. 2001. Restructuring and the order of aspectual and root modal heads. In G. Cinque and G. Salvi (eds), *Current Studies in Italian Syntax. Essays Offered to Lorenzo Renzi*. Amsterdam: Elsevier, pp. 137–155.

Cinque, G. 2002 [1997] The interaction of passive, causative, and "restructuring" in Romance. In C. Tortora (ed.), *The Syntax of Italian Dialects*. New York: Oxford University Press, pp. 50–66.

den Dikken, M., R. Larson, and P. Ludlow. 1996. Intensional 'transitive' verbs and concealed complement clauses. *Rivista di Linguistica* 8: 331–348.

Depiante, M. 1998). On the Interaction Between 'Ellipsis' and Restructuring in Spanish and Italian. Ms., University of Connecticut, Storrs.

Evers, A. 1975. The guillotine principle. In A. Kraak (ed.), *Linguistics in the Netherlands 1972–1973*, Amsterdam: Van Gorcum, p. 147.

Henry, A. 1995. *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford Studies in Comparative Syntax. New York and Oxford: Oxford University Press.

Hernanz, M. L., and G. Rigau. 1984. Auxiliaritat i reestructuració. *Els Marges* 31: 29–50.

Jackendoff, R. 1997. *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

Kayne, R. S. 1989. Null subjects and clitic climbing. In O. Jaeggli and K. J. Safir (eds), *The Null Subject Parameter*, Dordrecht: Kluwer, pp. 239–261. [Reprinted in Kayne (2000, Ch. 4).]

Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. S. 2000. *Parameters and Universals*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2002. Pronouns and their antecedents. In S. D. Epstein and T. D. Seely (eds), *Derivation and Explanation in the Minimalist Program*, Oxford: Blackwell, pp. 133–166. [Reprinted in Kayne (2005, Ch. 6).]

Kayne, R. S. 2005. *Movement and Silence*. New York and Oxford: Oxford University Press.

Kayne, R. S. 2006. On parameters and on principles of pronunciation. In H. Broekhuis, N. Corver, R. Huibregts, U. Kleinhenz, and J. Koster (eds), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, Berlin and New York: Mouton de Gruyter, pp. 289–299. [Reprinted in Kayne (2010, Ch. 4)].

Kayne, R. S. 2010. *Comparisons and Contrasts*. New York and Oxford: Oxford University Press.

Koopman, H., and A. Szabolcsi. 2000. *Verbal Complexes*. Cambridge, MA: MIT Press.

Kuteva, T. 1998. On identifying an evasive gram: Action narrowly averted. *Studies in Language* 22: 113–160.

McCawley, J. 1974. On identifying the remains of deceased clauses. *Language Research* 9: 73–85. [Reprinted in McCawley (1979, pp. 74–85).]

McCawley, J. 1979. *Adverbs, Vowels, and Other Objects of Wonder*. Chicago: University of Chicago Press.

Moore, J. 1994. Romance cliticization and relativized minimality. *Linguistic Inquiry* 25: 335–344.

Napoli, D. J. 1981. Semantic interpretation vs. lexical governance: Clitic climbing in Italian. *Language* 57: 841–887.

Perlmutter, D. 1983. *Studies in Relational Grammar*. Chicago: University of Chicago Press.

Pizzini, Q. A. 1981. The placement of clitic pronouns in Portuguese. *Linguistic Analysis* 8: 403–430.

Pollock, J.-Y. 1978. Trace theory and French syntax. In S. J. Keyser (ed.), *Recent Transformational Studies in European Languages*. Cambridge, MA: MIT Press, pp. 65–112.

Pollock, J.-Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 10: 365–424

Postal, P. 1974. *On Raising*. Cambridge, MA: MIT Press.

Pustejovsky, J. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.

Pylkkänen, L., A. E. Martin, B. McElree, and A. Smart. 2009. The anterior midline field: Coercion or decision making? *Brain and Language* 108: 184–190.

Quicoli, C. 1976. Conditions on clitic-movement in Portuguese. *Linguistic Analysis* 2: 199–223.

Radford, A. 1977. *Italian Syntax: Transformational and Relational Grammar*. Cambridge: Cambridge University Press.

Raposo, E. 1987. Case theory and Infl-to-Comp: The inflected infinitive in European Portuguese. *Linguistic Inquiry* 18: 85–109.

Rizzi, L. 1976a. Ristrutturazione. *Rivista di Grammatica Generativa* 1: 1–54.

Rizzi, L. 1976b. La MONTEE DU SUJET, le *si* impersonnel et une règle de restructuration dans la syntaxe italienne. *Recherches Linguistiques* 4: 158–184.

Rizzi, L. 1978. A Restructuring Rule in Italian Syntax. In S. J. Keyser, ed., *Recent Transformational Studies in European Languages*. Cambridge, MA: MIT Press, pp. 113–158. [Reprinted in Rizzi (1982, Ch. 1).]

Rizzi, L. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Rochette, A. 1999. The selection properties of aspectual verbs. In K. Johnson and I. Roberts (eds), *Beyond Principles and Parameters. Essays in Memory of Osvaldo Jaeggli*. Dordrecht: Kluwer, pp. 145–165.

Rooryck, J. 1994. Against optional movement for clitic climbing. In M. L. Mazzola (ed.), *Issues and Theory in Romance Linguistics*, Washington, DC: Georgetown University Press, pp. 417–443.

Ross, J. 1976. To have have and to not have have. In M. A. Jazayery, E. C. Polom, and W. Winter (eds), *Linguistic and Literary Studies in Honor of Archibald Hill*, Lisee, Holland: de Ridder, pp. 263–270.

Terzi, A. 1996. Clitic climbing from finite clauses and tense raising. *Probus* 8: 273–295.

Wexler, K., and P. Culicover. 1980. *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.

Wurmbrand, S. 1998. *Infinitives*. Doctoral dissertation, MIT.

Wurmbrand, S. 2001. *Infinitives: Restructuring and Clause Structure. Studies in Generative Grammar* 55. Berlin: Mouton de Gruyter.

Zagona, K. 1986. Las períphrasis de gerundio y la restructuración. *Revista Argentina de Lingüística* 2: 232–244.

35

Deriving Greenberg's Universal 20 and Its Exceptions

Guglielmo Cinque

2005

35.1 Introduction

It is well known that languages exhibit default word orders, both at the level of the clause (e.g., SVO, SOV, VSO) and of the noun phrase (e.g., Demonstrative Adjective N, or N Adjective Demonstrative). Just as not all possible word orders are attested within a single language, we also find that not all possible word orders are attested across languages, as typological work in the Greenbergian tradition has shown. For example, within the noun phrase only certain word orders are attested crosslinguistically, an observation made in Greenberg's (1963) Universal 20:

> "When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite."

Cinque's article aims to account for the limits on word order variation attested across languages in the nominal domain. Its goal is to derive the possible word orders and rule out the impossible ones in a principled way, by deriving them from general properties of the syntactic component of human language.

Deriving Greenberg's Universal 20 refines the empirical generalization made by Greenberg's statement concerning the order of postnominal elements, including orders in which the noun is neither final nor initial. The article points out that, of the 24 mathematically possible orders, only 14 are actually attested. Assuming that this empirical finding is representative of the possibilities allowed by UG, a descriptively adequate theory of syntax must derive all and only these 14 attested

orders. Cinque's article argues that such a theory contains the following components:

(1) a structural hierarchy of functional projections in which Dem > Num > Adj > N (where > stands for "is structurally higher than");
(2) a restriction on movement, stating that the only admissible (non-focus-related) movements involve phrases that contain the (overt) NP, possibly carrying along (or "pied-piping") other material along with it.

Assuming that the linearization of hierarchical structures maps asymmetric c-command onto precedence (Kayne 1994), a syntactic component enhanced with properties (1) and (2) will be able to derive all and only the attested orders, as desired (for a different perspective on this pattern, see Abels and Neeleman 2009).

Cinque further notes that some of the orders attested crosslinguistically are very common, while others are rare. He suggests that the variation in frequency be captured in terms of different degrees of markedness of the various kinds of movement and pied-piping configurations involved in the derivation of each order. In particular, absence of movement and movement of an NP plus pied-piping of the "*whose picture*" type are unmarked. In contrast, movement of an NP without carrying material along is more marked, and movement of an NP plus pied-piping of the "*picture of who*" type is even more marked. Cinque leaves open the question of what underlies the differences in markedness.

In light of work on the parallelism between the noun phrase and the clause (for an overview, see Bernstein 2001 and Longobardi 2001), Cinque's account of word order variation in the nominal domain naturally leads to the question of whether the same restrictions on movement also apply in the clausal domain. If so, this provides support for an approach to verb movement in terms of XP movement, as independently pursued in Koopman and Szabolcsi (2000), Nilsen (2003), Müller (2004), and Jayaseelan (2010), among others.

Deriving Greenberg's Universal 20 illustrates how work in comparative syntax in the generative tradition can deepen our understanding of generalizations discovered in the Greenbergian typological tradition.

35.2 From "DERIVING GREENBERG's UNIVERSAL 20 AND ITS EXCEPTIONS"

1   A derivational account of Greenberg's Universal 20

Greenberg's (1963) Universal 20,[1] under its most sensible interpretation (see Hawkins 1983: 117ff.), states (a) that in prenominal position the order of demonstrative, numeral, and adjective (or any subset thereof) conforms to the order Dem > Num > A, and (b) that in postnominal position the order of the same elements (or any subset thereof) conforms either to the order Dem > Num > A or to the order A > Num > Dem.

Forty years later, the first part of this statement remains (virtually) unchallenged, while the second part has proven both too restrictive and too permissive. Some studies, for example, have uncovered the existence of postnominal orders that are excluded by Greenberg's formulation. For Gabra (Cushitic), Luo (Nilotic), and Logoli (Bantu), Heine (1981) reports the order N Num A Dem, which conforms neither to Dem > Num > A nor to A > Num > Dem. Hawkins (1983:119), citing Hyman (1979:27), mentions the existence in Aghem (Bantu) of the order N A Dem Num, which again conforms neither to Dem > Num > A nor to A > Num > Dem; and he reports, also citing Hyman (1981:31), that Noni (Bantu), in addition to N Dem Num A, displays the order N Dem A Num, again unexpected under Greenberg's formulation.

On the basis of these facts, Hawkins (1983:119–120) concludes that the second part of Greenberg's Universal 20 must be abandoned, and that for the postnominal order of demonstrative, numeral, and adjective essentially every combination is possible. His proposed revision of Greenberg's Universal 20 is quoted in (1).
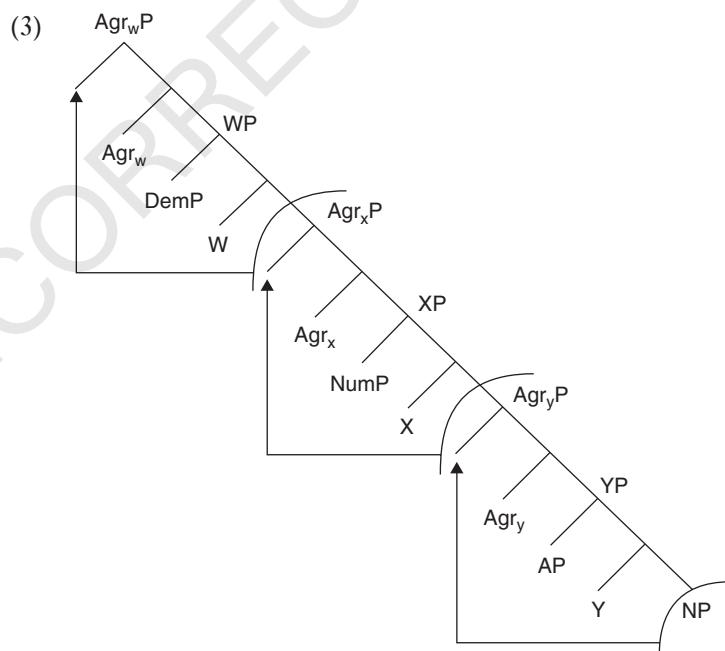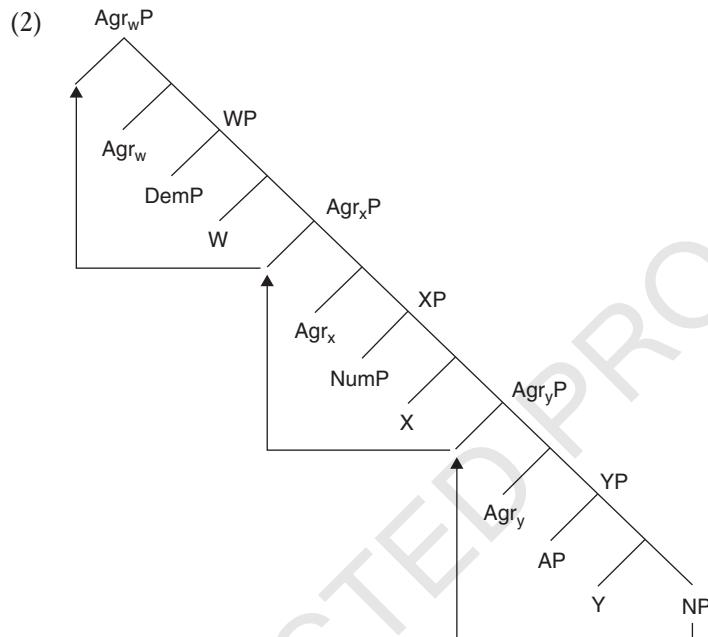
(1) "When any or all of the modifiers (demonstrative, numeral, and descriptive adjective) precede the noun, they (i.e., those that do precede) are always found in that order. For those that follow, no predictions are made, though the most frequent order is the mirror-image of the order for preceding modifiers. In no case does the adjective precede the head when the demonstrative or numeral follow."(Hawkins 1983:119–120, (20′))

Given that certain postnominal order possibilities (namely, *N Num Dem A and *Num N Dem A) are (still) unattested, as far as I was able to determine (see below), and given that even the actually attested orders differ significantly, as we will see, in the percentage of languages that instantiate them, I will not follow Hawkins, or Croft and Deligianni, in their conclusion that postnominally anything goes; instead, I will propose a refinement of an analysis I suggested in Cinque 1996, 2000 to derive Greenberg's basic generalization – one that may also derive its exceptions, and the different degree of markedness of the various orders.

The analysis I suggested in those works aimed at deriving the essential left–right asymmetry in word order possibilities found prenominally (one) and postnominally ((at least) two), starting from the idea (actually forced by Kayne's (1994) Antisymmetry Theory) that generating modifiers symmetrically to the left and to the right of the N could not easily account for the absence, prenominally, of the order A Num Dem.[5] This asymmetry could instead be made sense of, I submitted, if all orders are derived by moving (or not moving) the NP around the modifiers, base-generated prenominally in the fixed order Dem Num A.

If nothing moves, the unique (Merge) order found prenominally (the Dem Num A N order) surfaces. As for the two postnominal orders, they arise via the two ways in which the NP raises: either alone, from specifier to specifier (Spec to Spec) of Agr(eement) projections found above each of the functional projections hosting

adjectives, numerals, and demonstratives, to give the order N Dem Num A (as in (2)), or by moving successively to each such Spec and pied-piping the category that dominates it, in a "roll-up" fashion that reverses the order of the modifiers, to give N A Num Dem (as in (3)).

(2)

(3)

This was a simplification in that that analysis generated only the orders in (4) and, taking partial movements into consideration, those in (5) – 6 out of the 24 orders that are the mathematically possible combinations of the four elements Dem Num A N (factorial $4 = 4 \times 3 \times 2 \times 1 = 24$).

(4) a. Dem Num A N
     b. N Dem Num A
     c. N A Num Dem
(5) a. Dem Num N A
     b. Dem N Num A
     c. Dem N A Num

The crucial question then is, of the 24 orders, which ones are actually attested? And, if more are attested than the 6 indicated, how can the attested ones be derived in this system without also deriving the unattested ones?

(6) shows all the 24 orders. The "✓" and "*" preceding them indicate whether the order is attested or unattested, respectively. This indication is based on the typological (or other) sources available in the literature on the order of N, demonstrative, numeral, and adjective (that I have been able to find). The "∅" and references following some of the orders indicate that in those references, the order in question is explicitly claimed not to be attested. (Orders enclosed in a box are by far the most common. See the remark in Hawkins's revised formulation of Greenberg's Universal 20 in (1) about the mirror image of the prenominal order being the most frequent order in postnominal position. This is indeed matched by the relatively few languages instantiating the order N Dem Num A, as already noted in Greenberg's remark quoted in footnote 10 below.)

(6)	a.	✓	Dem Num A N	(very many languages)
	b.	✓	Dem Num N A	(many languages)
	c.	✓	Dem N Num A	(very few languages)
	d.	✓	N Dem Num A	(few languages)
	e.	*	Num Dem A N	(∅ – Greenberg 1963, Hawkins 1983)
	f.	*	Num Dem N A	(∅ – Greenberg 1963, Hawkins 1983)
	g.	*	Num N Dem A	(∅ – Lu 1998:183; but see footnote 26)
	h.	*	N Num Dem A	(∅ – Greenberg 1963, Lu 1998:162)
	i.	*	A Dem Num N	(∅ – Greenberg 1963, Hawkins 1983)
	j.	*	A Dem N Num	(∅ – Greenberg 1963, Hawkins 1983)
	k.	✓	A N Dem Num	(very few languages)
	l.	✓	N A Dem Num	(few languages)
	m.	*	Dem A Num N	(∅ – Greenberg 1963, Hawkins 1983)
	n.	✓	Dem A N Num	(very few languages)

o.	✓	Dem N	A	Num	(many languages)
p.	✓	N	Dem	A	Num	(very few languages – possibly spurious; see footnote 27)
q.	*	Num A	Dem	N	(∅ – Greenberg 1963, Hawkins 1983)
r.	✓	Num A	N	Dem	(very few languages)
s.	✓	Num N	A	Dem	(few languages; but see footnote 32)
t.	✓	N	Num	A	Dem	(few languages)
u.	*	A	Num	Dem	N	(∅ – Greenberg 1963, Hawkins 1983)
v.	*	A	Num	N	Dem	(∅ – Greenberg 1963, Hawkins 1983)
w.	✓	A	N	Num	Dem	(very few languages)
x.	✓	N	A	Num	Dem	(very many languages)

Keeping to the idea that no symmetric base-generation of modifiers is possible, and that postnominal orders are only a function of the raising of the NP (or of an XP containing the NP), it seems possible to derive all the attested orders, without also deriving the unattested ones. What we have to assume is this:

(7)   a.   Merge order: [. . . [$_{WP}$ Dem . . . [$_{XP}$ Num . . . [$_{YP}$ A [$_{NP}$ N]]]]]
      b.   Parameters of movement:
           (i)   No movement (unmarked), or
           (ii)   Movement of NP plus pied-piping of the *whose picture* type (unmarked), or
           (iii) Movement of NP without pied-piping (marked), or
           (iv)   Movement of NP plus pied-piping of the *picture of who* type (more marked still).
           (v)   *Total* (unmarked) versus *partial* (marked) movement of NP with or without pied-piping (in other words, NP raises all the way up, as in (6d,l,p,t,x), or just partially, as in (6b,c,k,n,o,r,s,w), around its modifiers).
           (vi)   Neither head movement nor movement of a phrase not containing the (overt) NP is possible (except perhaps for focus-related movements of phrases to a DP-initial position).

Let's consider how these assumptions manage to derive the attested orders, and fail to derive the unattested ones (in the computation of markedness, I take the markedness induced by partial movement to be less severe than the markedness induced by movement without pied-piping, which is in turn less severe than that induced by movement with pied-piping of the *picture of who* type; see below for some discussion).

(6a)   (Dem Num A N) is derived if nothing moves (7bi). (No marked option: very many languages.)

(6b)   (Dem Num N A) is derived from Dem Num A N if NP raises one notch, around A, either with (vacuous) pied–piping of the *whose picture* type (7bii) (unmarked) or without pied–piping (7biii) (marked). (Despite the marked-ness of partial movement, it includes the unmarked case of pied–piping: many languages.)

(6c)   (Dem N Num A) is derived if NP moves two notches, around A and Num (i.e., partially – marked option) without pied–piping ((7biii) – marked option). (Two marked options: very few languages.)

(6d)   (N Dem Num A) is derived if NP moves three notches, around A, Num, and Dem (i.e., all the way up) without pied–piping ((7biii): marked). (One marked option: few languages.)

(6e)   (Num Dem A N) cannot be derived through (7). NP has not moved, and the modifiers to its left are in the wrong Merge order (cf. (7a)).

(6f)   (Num Dem N A) cannot be derived through (7). Raising of NP without pied–piping implies a wrong Merge order of the modifiers (Num Dem A N) (see (7a)). Raising of NP with pied–piping of the *picture of who* type either of [Dem N] or of [Num Dem N] also implies a wrong Merge order (either Num A [Dem N] or A [Num Dem N]).

(6g)   (Num N Dem A) cannot be derived through (7). Raising of NP without pied–piping implies that the Merge order is Num Dem A N, which is a wrong order. Raising of NP with pied–piping of the *whose picture* type again implies a wrong Merge order of the modifiers (Num A Dem N), with N first raising around Dem and [N Dem] then raising around A. Raising of NP with pied–piping of the *picture of who* type (raising of [Num N] two notches) also implies a wrong Merge order of the modifiers (Dem A Num N).

(6h)   (N Num Dem A) cannot be derived through (7). Raising of NP without pied–piping implies wrong Merge order (Num Dem A N). Raising of NP with successive pied–pipings of the *whose picture* type also implies a wrong Merge order (A Dem Num N). Raising of NP without pied–piping around Dem and Num, followed by raising with pied–piping around A, would derive (6h), but, again, from a wrong Merge order (A Num Dem N). (Similarly if NP were to move around Num and pied–pipe it to the left of A and then move on without further pied–pipings. The Merge order in this case would be Dem A Num N – again, the wrong order.)

(6i)   (A Dem Num N) cannot be derived through (7). NP has not moved, and the modifiers to its left are in the wrong Merge order (see (7a)).

(6j)   (A Dem N Num) cannot be derived through (7). NP has moved one notch, but the two modifiers to its left are in the wrong Merge order (see (7a)). (6j) could also arise via raising of NP with pied–piping of the *picture of who* type of either Dem N or A Dem N around Num, but both derivations presuppose a wrong Merge order (A Num Dem N and Num A Dem N, respectively).

(6k)   (A N Dem Num) has a well-formed, though marked, derivation with rais-ing of NP plus pied–piping of the *picture of who* type of the lowest modifier (A), followed by raising of [A N] without pied–piping around both Num and Dem. (Two marked options: very few languages.)

(6l)   (N A Dem Num) has a derivation in which NP raises past A, followed by pied–piping of the *whose picture* type past Num, followed by raising of [N A] without pied–piping (marked) past Dem. (One marked option: few languages.)

(6m)   (Dem A Num N) cannot be derived through (7). NP has not moved, and the modifiers to its left are in the wrong Merge order (see (7a)). (See footnote 2 for discussion of the apparent existence of some such cases.)

(6n)   (Dem A N Num) has a derivation with partial (marked) raising of NP plus pied–piping of the *picture of who* type of [A N] (marked) around Num. (Two marked options: very few languages.)

(6o)   (Dem N A Num) has a derivation from (7a) involving partial (marked) raising of NP plus pied–piping of the *whose picture* type, vacuously, and nonvacuously (of [N A]) around Num. (One marked option: many languages.)

(6p)   (N Dem A Num), if genuine (see footnote 27), may be especially marked, as its derivation from (7a) would seem to involve raising of NP with successive pied–pipings of the *whose picture* type around A and Num (alternatively, a single raising of the *picture of who* type of [A N] around Num) and then extraction of the sole NP around Dem.

(6q)   (Num A Dem N) cannot be derived through (7). NP has not moved, and the modifiers to its left are in the wrong Merge order (see (7a)).

(6r)   (Num A N Dem) has a derivation with partial (marked) raising of NP plus pied–piping of the *picture of who* type of A and Num ([Num A N]) (marked) around Dem. (Two marked options: very few languages.)

(6s)   (Num N A Dem) has a derivation with partial (marked) raising of NP around A, followed by raising plus pied–piping of the *picture of who* type of [Num N A] (marked) around Dem. (Two marked options: few languages (but see footnote 32).)

(6t)   (N Num A Dem) has a derivation with raising of NP without pied–piping around A and Num (marked), followed by raising plus pied–piping of the *whose picture* type of [N Num A] around Dem. (One marked option: few languages.)

(6u)   (A Num Dem N) cannot be derived through (7). NP has not moved, and the modifiers to its left are in the wrong Merge order (see (7a)).

(6v)   (A Num N Dem) cannot be derived through (7). Raising of NP without pied–piping implies a wrong Merge order of the modifiers (A Num Dem N) (see (7a)). Raising of NP with pied–piping of the *picture of who* type either of [Num N] or of [A Num N] also implies a wrong Merge order (either A Dem [Num N] or Dem [A Num N]).

(6w)   (A N Num Dem) has a derivation from (7a) with raising of NP plus pied–piping of the *picture of who* type of A around Num (marked), followed by raising of [A N Num] around Dem. (One marked option: few languages.)

(6x)   (N A Num Dem) has a derivation from (7a) involving raising of NP with successive pied–pipings of the *whose picture* type all the way up. (No marked option: very many languages.)

The fact that all N-final orders that do not respect the order Dem Num A ((6e), Num Dem A N; (6i), A Dem Num N; (6m), Dem A Num N; (6u), A Num Dem N) are very clearly unattested can indeed be taken to indicate that it is the raising of NP (or of an XP containing it) that is responsible for word order variation within the DP (perhaps, more generally, that it is the raising of the lexical part of a phrase that is responsible for word order variation within its "extended projection").[30]

This offers a way to make sense of the fact that only to the right of N are more orders possible (indeed, those deriving from the different modes in which NP, or an XP containing it, raises). It also offers a way to derive, at least in part, the different degrees of markedness of each order (and, I take it, the ensuing differences in the numbers of languages that instantiate each one).

Although I know of no clear independent reason why movement *without* pied-piping should count as more marked than movement *with* pied-piping (of the *whose picture* type) (whence the respective numbers of languages instantiating each order), it seems natural that those orders whose derivation crucially involves pied-piping of the *whose picture* type should be less marked (and be instantiated in more languages) than those involving pied-piping of the *picture of who* type. The different degree of markedness of the two types of movement appears to be suggested independently by contrasts like the following in English (and corresponding contrasts in other languages): *Whose pictures did you see yesterday?* versus ?*Pictures of who did you see yesterday?*; *Now I know whose picture he saw yesterday* versus **Now I know a picture of who he saw yesterday*.

[. . .]

3   Further questions and implications

Two of the crucial components of the analysis suggested above are that [Dem . . . [Num . . . [A . . . [N]]]] is the (universal) structure of the DP resulting from Merge, and that word order variations within DP across languages are fundamentally a function of how (different phrases containing) the NP move up the structure. At least, such assumptions were seen to be able to derive the actually attested orders of the four elements crosslinguistically (without also deriving the unattested ones). Other plausible options, such as raising just the N, and/or moving phrases not including the NP, were seen not to yield the same results.

Should this general approach to crosslinguistic word order variation within DP be confirmed, it will be natural to ask whether the quite extensive crosslinguistic word order variation within the clause and other phrases should not also be treated in terms of movement of "extended" phrases (necessarily) containing the "lexical" projection (VP, AP, PP, etc.), rather than in terms of head movement. This question remains to be explored.

[Dem . . . [Num . . . [A . . . [N]]]] is but a fragment of the internal structure of the DP. If we were to add universal quantifiers, ordinals, numeral classifiers, and relative clauses (RCs) – setting aside the fact that *A* is just an abbreviation for an ordered sequence of adjectives (Cinque 1994, Scott 2002, and references cited there), and ignoring Case, Number, possessors, demonstrative reinforcers, various types of determiners, functional adjectives like *other* and *same* (Kayne 2005: sec. 2.1),

diminutives/augmentatives, complements,[34] and so on – we would have 8 elements, whose mathematically possible combinations number (factorial 8 =) 40,320.

The actually possible combinations would (luckily) be much fewer if the 8 elements entered a fixed hierarchical structure resulting from Merge, and if variations of this structure could only arise via upward movements of phrases containing the NP.

Notes

Num refers throughout to cardinal numerals, not to Number (singular, plural, etc.) or to ordinal numerals.

1   "When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite" (p. 87).

5   For example, one could assume the two following symmetric base structures, the right-branching (ia) and the left-branching (ib):

(i)   a.   [Dem [Num [A [N]]]]
      b.   [[[[N] A] Num] Dem]

This would account for the mirror-image order of Dem Num A found to the right of the N. In order to account for the other postnominal order (the same as that found prenominally: N Dem Num A), one would have to either assume the left-branching structure [[[[N] Dem] Num] A] (possibly at odds with the natural relative semantic scope of these elements) or admit the possibility that N(P) raises to the left of Dem in the base-generated structure (ia). In either case, though, it would not be clear how the unwanted order A Num Dem N could be excluded, as nothing principled in such a system would seem to prevent the symmetric right-branching structure [A [Num [Dem [N]]]] once [[[[N] Dem] Num] A] is allowed, or, alternatively, the raising of N(P) to the right of Dem in the base-generated structure (ib). To exclude them, specific ad hoc principles would have to be introduced. The principled unavailability of such symmetric solutions is precisely one of the main consequences of Antisymmetry Theory (see Kayne 1994).

30 Allowing for movement of both N and NP, and especially for remnant movement of phrases not containing the NP (or containing only its trace), would wrongly permit the derivation of most of the unattested orders. See below for a constraint on remnant movement, proposed by Kayne (2005), that actually appears to ban such possibilities.

34 Complements require at least a brief discussion. They do not seem to be part of the NP that raises. So, for example, in Semitic, except for construct state genitives, which are found between the initial N and its modifiers, prepositional complements are stranded at the end of the DP, and are not dragged along by the NP in its "roll-up" movements (see Cinque 2000, Shlonsky 2004, and references cited there). Similarly, in Romance, the NP can raise

across (certain classes of) adjectives, stranding its PP complements (see Cinque [2003]). This nonadjacency of complements follows from Kayne's analysis of the (overt and covert) prepositions that introduce them. In Kayne 2000, 2002, 2004, prepositions are argued to be heads merged higher up in the extended projection of the NP (or outside the DP altogether), attracting their "complements" and forcing (in VO languages) the remnant to raise to their left, thus making them final in the DP. Interestingly, complements of the N in OV languages are generally DP initial, before Dem. This is the case in, for example, Turkish (Jaklin Kornfilt, pers. comm.); Hindi (Anoop Mahajan (pers. comm.) notes that, more markedly, they can also occur after the N); and Malayalam (though, as K. A. Jayaseelan (pers. comm.) points out, they are necessarily introduced as predicates of relative clauses). In other words, they seem to involve attraction to the left of P, but no movement of the remnant.

35.3   Questions pertaining to Cinque (2005)

1   In what way(s) is Cinque's account of Greenberg's generalization dependent on Kayne's (1994) antisymmetry proposal?

2   In what way(s) is having demonstratives, adjectives, and numerals generated prenominally essential to his account?

3   How easy or hard would it be for Cinque to recast his account in terms of head movement rather than phrasal movement? (Extra credit: Bring in Cinque 2010.)

4   Cinque has "Dem Num Adj N" directly reflecting the merge order of these elements, with other merge orders being disallowed by UG. To what extent is this view of Cinque's similar or dissimilar to Baker's (1988) UTAH hypothesis? To the proposal in Cinque's own 1999 book?

5   In the next-to-last footnote, Cinque mentions *all those four new jobs* as suggesting that *all* be merged above Dem. What might be the significance of the fact that in English the preceding is less natural than *all of those four new jobs* and especially than *all four of those new jobs*?

6   The *of* of *all (four) of* . . . seems not to be found in other Germanic languages or in Romance languages. Find another instance of *of* that is (largely) limited to English. (Extra credit: Find two additional ones.)

7   In his detailed survey, Cinque by and large does not include either definite articles or indefinite articles. Do you think that their inclusion would have changed Cinque's conclusion? Why, or why not?

8   English has *a long book* but *too long a book*. How might Cinque integrate such alternations into his general proposal? (Extra credit: How might one account for the fact that colloquial English allows *too long of a book* but not *a long of book*?)

9   English has *three good weeks* but also *a good three weeks*. How might Cinque try to integrate this apparent violation of strict prenominal order? What interpretive clues might be relevant?

10   Another instance in English of nonstrict order is *the exact same book / the same exact book*. What might Cinque say about these? In what way might an

adverb like *exactly* be relevant? (Extra credit: Is there any connection here to phrases like *a mere linguist*?)

11   What is the significance to Cinque's paper of notions like "many languages," "few languages," "very few languages"? In what way might it matter that a large number of languages have gone extinct, whether recently or long ago? Similarly for the large number of languages that will exist in the future, but do not yet exist.

12   In Basque and in Hebrew the numeral "one" follows N, while the other numerals precede N. To what extent is this kind of fact compatible with Cinque's proposal? (Hint: Bring in Barbiers 2005, 2007 and Kayne 2009.)

13   Find one way in which Turkish articles differ from English ones. Find one way in which the Turkish indefinite article strikingly resembles the English one.

14   Greek sometimes has phrases looking like "the red the book." How might Cinque attempt to integrate them into his analysis?

15   Gascon has bare plurals of the Spanish sort, but when there's a prenominal adjective that adjective must be preceded by a preposition corresponding to "of." What might Cinque propose for this kind of fact? (Extra credit: This is similar to phenomena discussed by Leu 2009. Find one point of similarity and discuss its significance.)

16   English allows in some cases pairs like *(?)all that day* and *that whole day*. How might these be related? Which would then more closely reflect the merge order? Would *whole* be an adjective or a quantifier word like *all*? Give your reasons.

17   Find (at least) one language whose counterpart of *all the books* looks more like "the all books" and discuss its importance.

18   Koopman (2003, 2005) has suggested, on the basis of her work on Maasai, that nouns themselves are derived from (reduced) relative clauses. Assume she's (on the) right (track) and discuss the possible consequences for Cinque.

19   Cinque notes that classifiers of the East Asian type generally appear between Numeral and Adjective and suggest that they be merged at that point. To what extent would it affect Cinque's proposal if classifiers turned out to be a subtype of noun?

20   If Cinque is right, all moved phrases within DP (apart from focus movement to a pre-D position) have to contain N. Why might that be required? (Hint: See the next question.) (Extra credit: How does this restriction on DP-internal movement link up to Cinque 1999?)

21   Cinque suggests that DP-internal NP-movement (or in some cases Chomsky's 2000 Agree) serves to license modifiers. Find three languages in which numerals display less agreement than adjectives and discuss the importance of that.

References

Abels, K., and A. Neeleman. 2009. Universal 20 without the LCA. In J. M. Brucart, A. Gavarró, and J. Solà (eds), *Merging Features: Computation, Interpretation, and Acquisition*, New York and Oxford: Oxford University Press, pp. 60–79.

Baker, M. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Barbiers, S. 2005. Variation in the morphosyntax of one. *Journal of Comparative Germanic Linguistics* 8: 159–183.

Barbiers, S. 2007. Indefinite numerals one and many and the cause of ordinal suppletion. *Lingua* 117: 859–880.

Bernstein, J. B. 2001. The DP hypothesis: Identifying clausal properties in the nominal domain. In M. Baltin and C. Collins (eds), *The Handbook of Contemporary Syntactic Theory*, Oxford: Blackwell, pp. 536–561.

Chomsky, N. 2000. Minimalist inquiries: The framework. In R. Martin, D. Michaels, and J. Uriagereka (eds), *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, Cambridge, MA: MIT Press, pp. 89–155.

Cinque, G. 1994. On the evidence for partial N movement in the Romance DP. In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, and R. Zanuttini (eds), *Paths Towards Universal Grammar. Studies in Honor of Richard S. Kayne*, Washington, DC: Georgetown University Press, pp. 85–110.

Cinque, G. 1996. The antisymmetric programme: Theoretical and typological implications. *Journal of Linguistics* 32: 447–464.

Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. New York and Oxford: Oxford University Press.

Cinque, G. 2000. On Greenberg's Universal 20 and the Semitic DP. In Laura Brugè (ed.), *University of Venice Working Papers in Linguistics* 10:2, Venice: University of Venice, Dipartimento di Scienze del Linguaggio, pp. 45–61. [An abridged version appears in L.-O. Delsing, C. Falk, G. Josefsson, and H. A. Sigurðsson (eds), (2003), *Grammar in Focus: Festschrift for Christer Platzack 18 November 2003*, Lund: University of Lund, Department of Scandinavian Languages, vol. 2, pp. 243–251.]

Cinque, G. 2003. The dual source of adjectives and XP- vs. N-raising in the Romance DP. Paper presented at the Incontro annuale di dialettologia, Padua, 26 June 2003; at the CASTL Conference, Tromsø, 2–4 October 2003; and at NELS 34, State University of New York at Stony Brook, 7–9 November 2003.

Cinque, G. 2010. *The Syntax of Adjectives. A Comparative Study*. Cambridge, MA: MIT Press.

Croft, W., and E. Deligianni. 2001. Asymmetries in NP word order. Ms., University of Manchester.

Greenberg, J. H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (ed.), *Universals of Language*, Cambridge, MA: MIT Press, pp. 73–113.

Hawkins, J. 1983. *Word Order Universals*. New York: Academic Press.

Heine, B. 1981. Determination in some East African languages. In G. Brettschneider and C. Lehmann (eds), *Wege zur Universalienforschung: Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler*, Tübingen: Gunter Narr, pp. 180–186.

Hyman, L. M. (ed.). 1979. *Aghem Grammatical Structure. Southern California Occasional Papers in Linguistics* 7. Los Angeles: University of Southern California.

Hyman, L. M. 1981. *Noni Grammatical Structure. Southern California Occasional Papers in Linguistics* 9. Los Angeles: University of Southern California.

Jayaseelan, K. A. 2010. Stacking, stranding, and pied-piping: A proposal about word order. *Syntax* 13: 298–330.

Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. S. 2000. A note on prepositions, complementizers, and word order universals. In R. S. Kayne, *Parameters and Universals*, New York: Oxford University Press, pp. 314–326.

Kayne, R. S. 2002. On some prepositions that look DP-internal: English *of* and French *de. Catalan. Journal of Linguistics* 1: 71–115.

Kayne, R. S. 2004. Prepositions as probes. In A. Belletti (ed.), *Structures and Beyond. The Cartography of Syntactic Structures* 3, New York: Oxford University Press, pp. 192–212.

Kayne, R. S. 2005. Some notes on comparative syntax, with special reference to English and French. In G. Cinque and R. S. Kayne (eds), *The Oxford Handbook of Comparative Syntax*, New York: Oxford University Press, pp. 3–69.

Kayne, R. S. 2009. The English indefinite article *one*. Handout of talk given at the University of Cambridge.

Koopman, H. 2003. Inside the noun in Maasai. In A. Mahajan (ed.), *Syntax at Sunset 3: Head Movement and Syntactic Theory. UCLA Working Papers in Linguistics* 10, Los Angeles: UCLA Linguistics Department, pp. 77–115.

Koopman, H. 2005. On the parallelism of DPs and clauses: Evidence from Kisongo Maasai. In A. Carnie, H. Harley, and S. A. Dooley (eds), *Verb First: On the Syntax of Verb-initial Languages. Linguistik Aktuell/Linguistics Today* 73, Amsterdam and Philadelphia: John Benjamins, pp. 281–302.

Koopman, H., and A. Szabolcsi. 2000. *Verbal Complexes*. Cambridge, MA: MIT Press.

Leu, T. 2009. From Greek to Germanic: Poly-(*in)-definiteness and weak/strong adjectival declension. In J. M. Brucart, A. Gavarró, and J. Solà (eds), *Merging Features: Computation, Interpretation, and Acquisition*, New York and Oxford: Oxford University Press, pp. 293–309.

Longobardi, G. 2001. The structure of DPs: Some principles, parameters, and problems. In M. Baltin and C. Collins (eds). *The Blackwell Companion to Syntax*, Oxford: Blackwell.

Lu, B. F. 1998. *Left-right asymmetries of word order variation: A functional explanation*. Doctoral dissertation, University of Southern California, Los Angeles.

Müller, G. 2004. Verb-second as vP-first. *Journal of Comparative Germanic Linguistics* 7: 179–234.

Nilsen, Ø. 2003. *Eliminating positions: The syntax and semantics of sentence modification*. Doctoral dissertation, Utrecht University.

Scott, G.-J. 2002. Stacked adjectival modification and the structure of the nominal phrases. In G. Cinque (ed.), *Functional Structure in DP and IP*, New York: Oxford University Press, pp. 91–120.

Shlonsky, U. 2004. The form of Semitic noun phrases. *Lingua* 114: 1465–1526.