Syntax

Nicholas LaCara

FLing · 5 September 2015

1 The Big Questions

Syntacticians are concerned primarily with understanding the structural relations in natural language. These structures ultimately derive word order, and word order makes a difference to interpretation:

(1) Lee will save Lyra.

(2) Lyra will save Lee.

- Sentences are more than just their lexical content. These two sentences have the same lexical material (words), but they do not mean the same thing.

We have to detect structure indirectly, though, and detecting these structures is one of the main jobs of a syntactician.

- We know that structure exists, too. Certain combinations of words are not possible:

(3) Pan ran up the tree.

(4) Up the tree Pan ran.

(5) Pan looked up the word.

(6) * Up the word Pan looked.

We also know that, surprisingly, this does not always have a lot to do with what words mean.

(7) Lyra ate.

(8) * Lyra devoured.

Further, the interpretation of some elements is influenced by structural concerns.

(9) Mrs. Coulter thinks [$_{TP}$ Will saw himself].

(10) * Will thinks [$_{TP}$ Mrs. Coulter saw himself].

There are some obvious questions that this endeavor entails:

- How are syntactic structures (phrases) created?
- What grammatical mechanisms relate constituents with one another?
- What mechanisms mediate phenomena such as displacement, binding, ellipsis, interpretation?

Much of this is done in the service of answering much larger questions about the language faculty:

- What syntactic structures are possible in natural languages? What structures are not?
- How do children acquire a complete adult grammar with only a limited, impoverished input?
- Assuming that humans have some innate capacity for linguistic competence, how much can be attributed to this capacity and how much is language specific? What, if anything, is universal?
- In what ways is syntax responsible for the relationship between sound and meaning?
- How do a limited, constrained set of rules and finite set of terms interact to generate an infinite set of expressions?
- How do we describe the system that makes this possible and not just the output of that system?

That said, the day-to-day life of a syntactician is often taken up with much more low-level activities:

- Categorizing structures and constituents.
- Understanding the structural relations between words.
- Identifying displaced constituents and determining how displacement is constrained.

Contents:

2 Some background on Minimalism and Generative Grammar

The most widely accepted modern paradigm for approaching theoretical syntax is the Minimalist Program (but see Section 6!). The Minimalist Program was not born in a vacuum, and it is important to understand the methodological and philosophical underpinnings that led to it since these still guide syntacticians today.

2.1 Competence and performance

Chomsky (1965) introduces a distinction between what a speaker knows when they know a language and what they actually do when they speak a language.

Linguistic Competence is, roughly speaking, what a speaker knows when they know a language. This is also referred to as I-language.

- Following Chomsky, when a child acquires a language, they create a mental system of rules for speaking that language. This list of rules is typically what syntacticians are interested in studying.
- Competence cannot be studied directly. Syntacticians try to understand what competence is by studying speakers intuitions about what is and is not a part of their language (more on this in Section 5).

Linguistic Performance, even more roughly, is what people actually do when they speak and how they use language.

- Performance is limited by numerous factors (memory, processing, fatigue, beer, *etc.*).
- Under the hypothesis that speakers really do have a system of rules about language in their minds, these factors are thought to be external to these rules. Thus, they are not truly part of the language faculty and should be factored out.

2.2 Descriptive and Explanatory adequacy

Syntacticians seek to model a speaker's linguistic competence. These models are referred to as grammars.

- Broadly speaking, such grammars should produce, or generate, all of the grammatical sentences of a language and none of the ungrammatical ones.
- Ideally, the model should tell us something about the language faculty itself.
- Chomsky (1964, 1965) discusses two concepts for measuring how well a grammar succeeds at these tasks.

Descriptive Adequacy is a measure of how closely a grammar mimics the competence of a native speaker.

- "A grammar can be regarded as a theory of a language; it is *descriptively adequate* to the extent that it correctly describes the intrinsic competence of the idealized native speaker. The structural descriptions assigned to sentences by the grammar, the distinctions that it makes between well-formed and deviant, and so on, must, for descriptive adequacy correspond to the linguistic intuition of the native speaker […] *a linguistic theory is descriptively adequate* if it makes a descriptively adequate grammar available for each natural language." (Chomsky 1965:24, emphasis original)

Explanatory Adequacy is a measure of distinguishing between theories beyond just the empirical data they explain. It is a metric for describing the linguistic system, not just the output of the system.

- "To the extent that a linguistic theory succeeds in selecting a descriptively adequate grammar on the basis of primary linguistic data, we can say that it meets the conditions on *explanatory adequacy* [...] Gross coverage of a large mass of data can often be attained by conflicting theories; for precisely this reason it is not, in itself, an achievement of any particular theoretical interest or importance. [...] On a much deeper and hence much more rarely attained level (that of explanatory adequacy), a grammar is justified to the extent that it is a principled descriptively adequate system, in that the linguistic theory with which it is associated selects this grammar over others, given primary linguistic data with which all are compatible." (Chomsky 1965:25, 26–27, emphasis original)

2.3 Principles and Parameters

The origins of Minimalism are rooted in the Principles and Parameters framework (P&P) and can often be seen as a reaction to the the main theory that preceded it: Government and Binding Theory.

- Central hypothesis of P&P: Humans have a language faculty that permits them to learn human language. This is Universal Grammar (UG).
- UG provides principles for constructing grammars. All grammars share these priniciples, and they constrain possible grammars.
- Alongside these, UG makes available a number of parameters. Individual parameters (*e.g.*, headedness, null subjects, *wh*-movement) are set during language acquisition, accounting for variation between languages

One of the main goals is to figure out what the principles and parameters are.

- The principles need to be constrained enough to allow acquisition of human languages, but they need to permit the full range of observed grammatical variation allowed by the parameters.
- Discovering these is necessarily an inductive process: We cannot observe the human language faculty directly, only the languages that it produces and the judgments speakers have.

2.4 Government and Binding

Government and Binding (GB) was the central theory that preceded Minimalism. It shares P&P assumptions.

- In particular, Government was the fundamental grammatical relation in GB (Hornstein et al. 2005:24).
- Government unified many of the different GB modules. It was used to do everything: Case assignment, θ-role assignment, trace licensing, establishing binding domains, determining the distribution of pro, and even ellipsis licensing.

GB included at least four levels of representation.

- D-structure was the place where all the words were put together, following the rules of X-bar Theory.
- This was transformed into S-structure, which was then sent to the interface levels Phonological Form (PF) and and Logical Form (LF). Further (covert) transformations changed S-structure into an LF.
- Movement operations introduced traces over the course of the derivation which were not part of the lexicon.
- Different modules applied at different levels of representation; *e.g.* Case at S-structure, binding at LF.

2.5 Rethinking

Putting it boldly, Minimalism begins by making the assumption that P&P is an adequate theory of grammar (Hornstein et al. 2005).

- P&P answers the question of how children acquire a complete grammar with an impoverished input: Principles constrain possible grammars, variation is determined by setting parameters.
- Assuming P&P is basically on the right track, Minimalism starts by questioning the results of GB.
  - Some questions are architectural: How many levels of representation are necessary? Can the relations between elements be reduced to more primitive relations? Are there any *ad hoc* components we can replace?

  - Some questions are more qualitative: If you are trying to discuss whether a theory is explanatorily adequate, then how does it compare to other theories in terms of naturalness or elegance? What counts as parsimonious? Is language optimal in design?

One of the main goals of Minimalism as a research program is to reduce the number of primitive syntactic relations to the fewest possible and remove unmotivated levels representation.

- This led ultimately the rejection of Government as a primitive relation. As a consequence, any theories/hypotheses that were dependent on Government needed to be reevaluated.
- The number of levels of syntactic representation have been significantly reduced.
  - PF and LF have been retained (since there is a relation between form and meaning).
  - There is no D- or S-structure anymore since they are theory internal. The conception has been replaced with a single continuous derivation. Any theoretical constructs referencing these levels must be reconsidered.
  - X-bar theory is not really compatible with the continuous derivation approach, and has been replaced with Bare Phrase Structure.
  - GB modules were often reduced to interface conditions.

Let's look at some of these results.

3 Minimalist technology

There is an oft-repeated saying: 'Minimalism is a program, not a theory'. It is worth considering what this means.

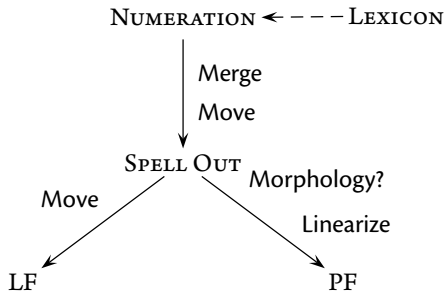
- As a research program, Minimalism has a number of goals, most of which center on reduction of theoretical machinery, reevaluation of proposed mechanisms, and reduction of theory-internal levels of representation.
- In that sense, the Minimalist Program is not a theory: A way of thinking or pursuing research does not make predictions in and of itself.
- That said, Chomsky (1993) first introduced the Minimalist Program over 20 years ago, and this way of thinking has resulted in a number of interrelated mechanisms, hypotheses, and principles that have become central to mainstream Generative syntax.
- When one refers to *Minimalism* or *Minimalist syntax*, they are often referring to this collection of ideas. Given the influence the Minimalist Program has on syntax – especially given how different it is from its forebears – this is hardly surprising.

The following are core ideas that underlie most current analyses and are widely assumed as background knowledge. These descriptions are, of course, cursory, and each of them is subject to its own controversies.

3.1 A Minimalist model of grammar

Like most approaches in the generative tradition, Minimalist syntax assumes multiple levels of representation. The most widely assumed basic model is the so-called Y-model, based on Chomsky 1995:

(11)



- The numeration is the set of lexical items to be used in the derivation. These are assembled by the operations Merge and Move.
- At some point in the derivation, the output of these operations is spelled out to the interface levels, LF and PF. Each branch of the derivation is distinct: Operations on the PF branch are not visible to LF, and *vice versa*. The part of the derivation before Spell Out is referred to as the narrow syntax.
- LF is the domain of interpretation (roughly, what feeds semantics). By hypothesis, it is assumed that syntactic operations can continue on the LF

1 Hornstein et al. (2005:Ch. 9) discuss some alternatives to this picture.

branch after Spell Out to check features. Overt movement must happen in the narrow syntax. Covert movement happens on the LF branch.[1]

- PF is the domain of pronunciation (what feeds phonetics and phonology). What happens on the PF branch is a bit more contentious. This is certainly the domain of linearization. Late insertion models of morphology argue that morphology occurs on this branch (see, *e.g.*, Halle and Marantz 1993).

3.2 Merge & Move

Merge and Move are traditionally the main operations of the computational system.

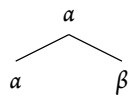
3.2.1 Merge

MERGE is an operation that takes two terms and combines them to create a new element.

(12) $\text{Merge}(\alpha, \beta) \rightarrow \{\alpha, \{\alpha, \beta\}\}$

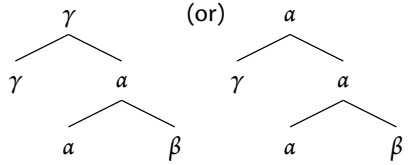
- One of the two terms, in this case α, projects a label. This term that projects the label is the HEAD.

The result of this operation can be represented informally in tree form:

(13)


Merge is recursive and can operate on its output.

- For instance, if Merge creates an object $K = \{\alpha, \{\alpha, \beta\}\}$, K is a legitimate term for further instances of Merge.
- Merge can form $L = \{\gamma, \{\gamma, K\}\}$ or $L = \{\alpha, \{\gamma, K\}\}$ depending on whether α or γ projects the label. By so doing, Merge derives larger structures as more lexical items are merged.

(14)


- This leads to bottom-up derivations (in the literal sense): The most deeply embedded elements in a tree are merged first. Compare this with the rewrite rules from, *e.g.*, Chomsky (1957), which are top-down.[2]
- Since Merge takes two terms, it can only create structures that are binary branching. This is probably a good result, since phenomena traditionally treated with *n*-ary branching show evidence of being non-symmetrical. Given its binary nature, Merge is also central to Bare Phrase Structure.[3]
- Understanding which term projects the label is still problematic, and we would like it to fall out from something rather than stipulating which element projects. This is important because heads select the labels of their complements when building trees.

3.2.2 Move

MOVE is an operation that is responsible for displacement. It is considerably more complicated than Merge (see Chomsky 1995:249–276).

- Minimalist syntacticians typically assume the COPY THEORY OF MOVEMENT.[4]

Move is a complex operation with (at least) two subparts:

i. *Copy* the term to be moved.

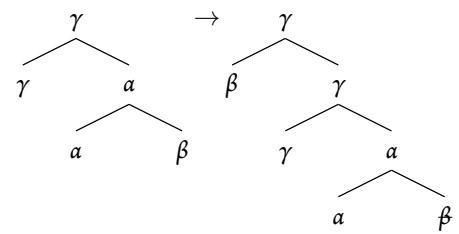
ii. *Merge* the copy in the target position, which is typically assumed to be the root of the derivation (the topmost node).

2 This approach is inherently derivational. Compare representational approaches like LFG, HPSG; GB D-structure.

3 Some of the evidence for lack of *n*-ary branching is covered in 601, so I will leave it aside here.

4 Chomsky (2001) distinguishes between EXTERNAL and INTERNAL Merge. External Merge, represented here, takes two independent terms and combines them. Internal Merge is when one of terms is a subpart of the other. This appears as movement, but is different from the more commonly assumed Copy Theory of Movement (Chomsky 1995). In particular, it invokes multidominant structures (Citko 2005).

(15)



The appearance of movement is the result of the fact that the original copy is typically left unpronounced.

- Under GB, Move was one operation (along with Bind) that derived S-structure from D-structure, and LF from S-structure. When an element moved, it would leave a trace in the position of the moved element in satisfaction of the Projection Principle.
- The Copy Theory of Movement avoids problems with traces that Minimalism seeks to avoid (*e.g.*, positing that traces are a primitive element, having to introduce new material over the course of the derivation).
- Dealing with the additional copies, though, is not a trivial matter. Exactly how a copy is left unpronounced is complicated, but it is probably not a properly syntactic issue; see Nunes 2004. Additionally, determining the proper semantic interpretation of copies is also an issue, though it has some benefits; see, for instance, Fox 1999.

Merge and Move, by hypothesis, are part of UG and thus invariant across languages.

3.3 Agree and Features

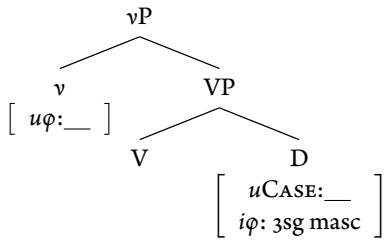
Features on heads play an important role in syntactic theorizing, generally driving the computation and movement.

3.3.1 Feature valuation

Features come in two varieties: INTERPRETABLE and UNINTERPRETABLE.

- Features are interpreted at the interfaces, and so uninterpretable features must be rendered interpretable for the derivation to converge.
- Uninterpretable features must be checked against interpretable features. In modern terms (Chomsky 2000, 2001), this is usually thought to happen via valuation: Heads bearing uninterpretable features enter the derivation without a value. Interpretable features are valued.
- For the derivation in (16) to eventually converge, the pronoun (represented by D°) must receive a value for Case. *v*° must have its *φ*-features valued.

(16)



The mechanism for valuing features is AGREE. Most syntacticians working under Minimalist assumptions assume this operation.

- An unvalued feature, or PROBE, searches for a matching valued feature, a GOAL, in its c-command domain.
- Thus, in (16), the *φ*-feature on *v*° causes *v*° to probe for a value in its c-command domain. It finds the *φ*-feature on D°.
- Case is assigned from *v*° to D° in the process.

3.3.2 Movement

Implementations of movement vary. Under the original Minimalist picture (Chomsky 1995), features divide further into STRONG and WEAK features.

- Weak features must be checked before LF.
- Strong features must be checked before PF.

Ultimately, checking was thought to involve movement of the interpretable element into a local configuration with the uninterpretable element.

- Thus, for an element to be checked, it needed to move, either overtly or covertly.
- Since there is no movement on the PF branch, as shown in (11), a strong feature must move before Spell-Out, or the derivation will crash. This leads to OVERT MOVEMENT.

- Since there is movement on the LF branch, weak features could move after Spell-Out since they only need to be checked at LF. This leads to COVERT MOVEMENT.

More recent thought, however, has attempted to dispense with the strong–weak distinction since the mechanisms driving it are stipulative.

- Instead, a Generalized EPP has been proposed (Chomsky 2000).
- Originally, the EPP was the feature on T° that required it to have a filled specifier.
- Under the more recent view, all features are essentially weak. The Generalized EPP feature is parasitic on other features and must be checked before Spell Out.
- If feature valuation with Agree is adopted, then there is no need for covert movement for feature checking (Hornstein et al. 2005:327).

If you have been paying attention, you may have realized that neither of these proposals answer the question why things move at all.

- Proposing feature-driven movement merely passes the question on: Why do some features need to be checked at PF or bear an EPP feature? We just don't know.

3.4 Economy

One of the motivations for Minimalism is trying to understand what should count as an economical derivation. Economy conditions constrain derivations, ruling out less optimal alternatives.

Many economy conditions are DERIVATIONAL in nature:

- INCLUSIVENESS: Do not introduced material that is not in the numeration.
- MERGE OVER MOVE: If you can merge a new item, then don't move one.
- LAST RESORT: A syntactic operation affecting α and β must check some feature of α or β.
- PROCRASTINATE: Don't move until you have to.
- SHORTEST MOVE: Smaller movements are better.

There are also REPRESENTATIONAL economy conditions, such as Full Interpretation, which requires that the output of the be legible at the interfaces.

Chomsky (2000, 2001) introduces PHASES, which are a domains for locality.

- Instead of Spell Out occurring 'at some point', Spell Out is triggered by certain heads over the course of the derivation.
- Spell Out prevents further operations from occurring, so material that is Spelled Out is no longer accessible for further operations (the PHASE IMPENETRABILITY CONDITION).
- This effectively limits the area in which computation can take place.

4 A convincing analysis: Extraction from ellipsis sites

Ellipsis is a well-studied phenomenon where a contiguous part of a sentence (marked here with Δ) is apparently missing.

- The meaning of this missing material is filled in by some antecedent (underlined here).
- This filling in has long been understood to be semantic-(ish) in nature – meanings are recovered, not phonologically identical material, which implies some LF component.

(17) Lyra rode a polar bear, but Will has not Δ.

(18) Mrs. Coulter kidnapped somebody, but I don't know who Δ.

An old question: Do ellipsis sites contain syntactic structure? There are two old hypotheses:

A. Ellipsis sites are phonologically null pronoun-like elements with no syntactic structure. They receive their interpretation from the antecedent at LF or in the semantics (Chao 1987, Lobeck 1995, Williams 1977).

B. Ellipsis sites contain syntactic structure that is unpronounced or deleted at PF. Having an antecedent with the same meaning is what allows this material to go unpronounced (Ross 1969, Sag 1976, Wasow 1972).

Spoiler alert: Hypothesis B, is the current consensus view. However, in the '80s and '90s, Hypothesis A was thought to be correct.

So what happened? Why did we change our mind? Let's look at SLUICING.

- Sluicing, as shown (18), involves a *wh*-word stranded next to an ellipsis site representing a missing TP.
- The question is: How does that *wh*-word get there? There are two options.
    a. It can be base-generated in that position. This is the only possibility if ellipsis sites do not have unpronounced structure.
    b. It can be moved into that position like any other *wh*-word. This is possible only if the ellipsis site contains unpronounced structure that the *wh*-word can move out of.

Thus if we can find evidence for movement, then ellipsis sites have syntactic structure.

4.1 Extraction

Evidence for movement has been known since Ross's (1969) seminal paper on Sluicing and was reinforced by Merchant's (1999, 2001) dissertation.

4.1.1 Case matching

First, in languages that display morphological case on DPs, the sluiced *wh*-word must match the case-marking it would have received had it remained *in situ*.

- In German many verbs assign an idiosyncratic case to their direct objects. The verb *schmeicheln*, 'to flatter', assigns dative case.
- A sluiced *wh*-element must receive the same case it would have received in the full sentence.

(19) Er will *jemandem* schmeicheln, aber sie wissen nicht, {*wer / *wen / wem} Δ
he wants someone.DAT flatter but they know not who.NOM / who.ACC / who.DAT

- If the sluiced *wh*-word gets case from a verb (which is not pronounced; Hyp. B) then this is explained straightforwardly.
    – If there is a verb in the ellipsis site that assigns case to the direct object 'wem', then we don't need to say anything special about how 'wem' gets dative case.
    – It gets it the same way any other direct object does (remember (16)) and then moves in the same way any other question word does. The TP it moves out of is simply left unpronounced.
- If the sluiced *wh*-word is base-generated next to the ellipsis site (Hyp. A), then we have to explain why case-marking behaves in a special way in sluicing contexts.
    – What assigns case to the *wh*-word? What mechanism makes it match the case in the antecedent? Why can't the case be different here?
    – We would have to significantly revise our approach to case assignment. Sluicing could be evidence that we need to do this, but we should want some independent evidence that our view of case assignment is wrong or that Hypothesis B is wrong.

4.1.2 P-stranding

The second consideration is preposition stranding, discussed by Merchant (2001).

- Languages that have obligatory PP pied-piping, like German (20a), reproduce this behavior in sluicing. When the correlate of the sluiced *wh*-element is contained in a PP, the sluiced *wh*-element must be in a PP (20b).
- Languages that allow preposition stranding, like Swedish, do not require prepositional pied-piping (21a). These languages also permit the sluiced *wh*-element to be a DP or a pied-piped PP when the antecedent contains a PP (21b).

(20) a. * Wem hat sie [mit *t*] gesprochen?
b. Anna hat [mit jemandem] gesprochen, aber ich weiß nicht, *(mit) wem Δ.
Anna has with someone spoken but I know not with who

(21) a. Vem har Peter talat [med *t*]?
who has peter spoken with?
b. Peter har talat [med någon]; jag vet inte (med) vem Δ.
Peter has spoken with somebody I know not with who

- Again, this matches what we observe in movement without ellipsis. The conditions on the sluiced *wh*-element are the same regardless of whether ellipsis occurs or not.
- If *wh*-words are base-generated next to the ellipsis site, then there needs to be some explanation for why languages like Swedish are different from German. Why must German base-generate PPs when the antecedent contains a PP? Why is Swedish allowed to base-generate DP instead? Why does this correlate with the ability to strand PPs overtly?

Again, we could try to come up with a way of explaining this under Hypothesis A, but we would also have to invent something for the case assignment facts too. Hypothesis B explains this data straightforwardly.

4.2 Congruence

Both case-matching and P-stranding point toward there being structure inside of ellipsis sites, and this supports Hypothesis B.

- Part of this has to do with our other commitments. We have an OK idea of how case-assignment works outside of ellipsis contexts. Do we want to change that around just to make it fit with Hypothesis A?
- Part of this has to do with other empirical observations: Even though we might not understand why preposition stranding is permitted in some languages, it looks like the generalization holds in ellipsis contexts in just the same way it does in normal circumstances.
- Case-matching and P-stranding – two fairly different phenomena – point the same way. We would have to revise our theory of case assignment *and* our theory of preposition stranding to allow for Hypothesis A. Hypothesis B does not require us to change anything.

Additional evidence for Hypothesis B has since been adduced from verb phrase ellipsis, which targets VPs instead of TPs.

- *Wh*-movement out of verb phrase ellipsis is possible (Schuyler 2001) and head movement out of verb phrase ellipsis sites is also strongly indicative of unpronounced structure (Goldberg 2005).

4.3 Why would/did we believe Hypothesis A?

Hypothesis B proposes that ellipsis is the non-pronunciation of syntactic material when there is an antecedent with the same meaning. This view is, on the face of it, hard to understand given the Y-model in (11).

- You will notice that PF and LF are not directly related.
- They both share the syntactic derivation as an input, but things that happen at PF cannot affect things at LF, and *vice versa*.
- If ellipsis is non-pronunciation (at PF), then the conditions on meaning at LF cannot trigger the non-pronunciation at PF. Similarly, an operation at PF cannot check and see if the conditions at LF hold.

This is why Hypothesis A was once so appealing.

- A null pronoun can be inserted in the syntax and would thus feed both PF and LF. The inability of PF and LF to see each other is sidestepped since they both share the narrow syntactic derivation.
- At PF this pronoun would receive no pronunciation, and at LF it would either receive an interpretation just like any other pronominal element or have some semantics copied onto it from the antecedent.
- But this just doesn't fit the data: Pronouns don't have syntactic structure out of which *wh*-movement can occur.

However, implementations of Hypothesis B retain some conceptual aspects of this idea.

- Merchant (2001) proposes that an element in the narrow syntax with a special semantics enforcing the antecedent requirement is responsible for triggering non-pronunciation at PF: the [E] feature.
- This is not a pronoun, but it allows us to maintain our commitment to the idea that LF and PF are not directly related.

5 Data and evidence in Syntax

The primary source of data and evidence in Syntax comes from informally collected intuitive judgments about whether utterances are grammatical or about their interpretation. Speakers are asked whether certain sentences are possible in their language or whether certain readings are available. Stars are assigned.

(22) a. Lyra often kisses Will.
b. * Lyra kisses often Will.

(23) Lee_i said that Iorek_k likes $\text{himself}_{k/*i}$.

One note on terminology: Syntacticians call these GRAMMATICALITY JUDGMENTS, but they are really more accurately referred to as ACCEPTABILITY JUDGMENTS.

- Grammaticality is only one component of acceptability that determines whether speakers accept a sentence or not (Chomsky 1965:11–12).
- Speakers do not have direct access to grammaticality. Linguists, when asked for judgments, can sometimes filter out other factors (*e.g.*, plausibility), but naïve speakers are not generally good at separating different aspects of acceptability.
- Separating out grammaticality from other factors is not always a trivial task. See, *e.g.*, psycholinguistics.

Importantly, not all judgments are categorical, and syntacticians, even in practice, do not assume that they are. As Chomsky (1965:11) put it, 'grammaticalness is, no doubt, a matter of degree'.

- Speakers often report feeling unsure about their judgments, thinking that the sentence is OK but not great, or bad but not terrible.
- Some speakers might think a sentence is completely grammatical while others do not accept it.
- Sometimes you get a mixture of these: Some speakers accept an example, others only think it's a bit off, some hate it.

Marginal judgments such as these often get marked with one or more question marks in the literature.[5]

(24) ? What did you wonder whether to read?

There is, unfortunately, no set way of dealing with marginal data, and what is done with these judgments varies.

- Most examples in theoretical syntax are judged in comparison to some other examples, and this contrast is used to detect differences in grammaticality (Sprouse et al. 2013). To that end, absolute grammaticality or acceptability is not always critical.
- Ideally, authors should be forthright about the reason for marginal data. If the examples are critical to the analysis, authors should take time to explain the (potential) source of marginality in an example.

There probably is not much of a problem here, at least for major languages. Since native speakers can check reported judgments when they hear talks or read articles, the accuracy of reported data is under frequent scrutiny.

5.1 Methodological criticisms

The informality of data collection in Syntax is a perennial source of criticism.

- One criticism is that data collection methods are not rigorous enough and judgments reported in the literature do not reflect linguistically valid tendencies.
- Another criticism is that judgments introduce author bias into the data when they aren't checked against the judgments of other speakers.

With regard to the first criticism, Phillips and Wagers (2007) suggest that informal intuitive judgments are typically sufficient for the types of questions with which syntacticians are concerned.

- Sprouse and Almeida (2013:225): "[T]he small, qualitative experiments that syntacticians have traditionally employed appear to be well calibrated to the phenomena that syntacticians have traditionally investigated, just as one would expect from a serious scientific field."

5 Unfortunately, there are no set standards in the field. Typically ?? is worse than ?. The mark *? is often used for things that are not grammatical, but not completely bad. The mark # should ideally be used for sentences that are grammatically well formed but anomalous in a given context (*i.e.* do not have a sensible reading), though * is often used for this. Other rarer markers exist, including % for cases where some speakers find a sentence grammatical and others do not, and ! for cases that pertain to non-standard dialects. I have even seen £ used to mark sentences that are only grammatical in British English.

As I noted above, the majority of data in theoretical syntax does not appear to be problematic.

- Phillips and Wagers (2007) note that "There are, of course, examples of errors and disagreements, and notorious cases of judgments that are subtle at best, but these are the exception rather than the rule."
- Sprouse and Almeida (2013) note that these are criticisms worth investigating and addressing, but there is little empirical evidence that the judgments reported in the literature are deeply flawed.

There have been recent attempts to determine whether the judgments reported in the literature are replicable.

- Sprouse et al. (2013), for example, collected all 1743 English data points published in *Linguistic Inquiry* from 2001 to 2010. They tested a random sample of 296 examples with 936 naïve speakers. Their results suggest a convergence rate of 95% between informal and formal methods.
- This suggests that informal methods do about as well as formal methods, though there is no clear way of deciding which is better.

5.2 Variation

One question I received is how syntactic theory deals with variation. The answer depends on what level you are looking at.

- Interspeaker and dialectal variation can be modeled in the same way we model different languages – with parameters or lexical variation.
- Intraspeaker variation is not typically handled very well. Our categorical theory is not very well-suited to handling apparently probabilistic data. There are some proposals out there; see, for instance, Abney 1996, Adger 2014 or Kroch 1994. Tom Roeper and Lisa Green are good people in the department to talk to about these issues.

6 Alternatives to Minimalism

There are a few alternatives to Minimalism in current active use, and you are likely to come across them in the wild.

- Perhaps the most notorious is Head-driven Phrase Structure Grammar, or HPSG (Pollard and Sag 1987). HPSG rejects the notion of transformation. Consequently, long-distance dependencies are not handled by movement. Rather, HPSG is unification-based: Dependencies and featural information are transmitted up the constituent structure.
- Lexical Functional Grammar, or LFG (Bresnan 2001), posits several levels of representation; for instance, constituent structure, functional structure and lexical structure (or c-structure, f-structure, and l-structure). Like HPSG, LFG rejects the notion of transformation, and levels of representation are not linked derivationally. C-structure is the only level of syntactic representation. Functional categories, like subject and object, are primitives and are mapped to elements of c-structure at f-structure. Functional categories are mapped to argument structure at l-structure. Consequently, LFG can be said to be non-configurational: Positions in the syntactic structure are not specific particular functions or arguments.[6]
- Outside of the Generative tradition, there has recently there has been a lot of discussion about Construction Grammars, or CxG. This is a diverse lot, but they are unified by the notion that constructions are the basic unit of syntactic representation. Constructions are pairings of forms with meanings, and everything from the double object construction to an individual word is a construction (there is no meaningful distinction between syntax and the lexicon). Thus, whereas most Generative approaches assume that constructions are the result of lexical composition, CxG argues that the constructions themselves are learned elements. CxG divides into roughly two groups, according to Adger (2012): Sign-based varieties and usage-based varieties. Sign-based varieties (*e.g.* Sag 2012) admit that the mind has some language-specific capabilities, but usage-based varieties (Goldberg 2006) reject the Generative notion that humans have a specialized innate capability for language.

Acknowledgments

Many thanks to the FLing organizers Caroline Andrews, Rodica Ivan, and Petr Kusily. Many thanks are also due to Mike Clauss, Georgia Simon, and Megan Somerday for the late-night conversations that inspired large chunks of these notes. I looked over Jason Overfelt's hadnout from last year's FLing, but I don't think I stole any of it (at least not intentionally). I conulted Hornstein et al. 2005 and Adger (2003) fairly ex-

6 Compare with Chomsky's Projection Principle.

tensively while writing these notes, and those authors deserve credit for their ideas that I repeated here. And of course, thanks everybody who submitted questions. I, regretfully, could not address all of them here.

References

Abney, Steven. 1996. Statistical Methods and Linguistics.

Adger, David. 2003. *Core Syntax: A Minimalist Approach*. Oxford University Press.

Adger, David. 2012. Constructions and Grammatical Explanation. *Queen Mary's Occasional Papers Advancing Linguistics* .

Adger, David. 2014. Variability and Grammatical Architecture. In *Linguistic Variation in the Minimalist Framework*. Oxford University Press.

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Blackwell Publishing.

Chao, Wynn. 1987. On Ellipsis. Doctoral Dissertation, University of Massachusetts Amherst.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

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

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The View from Building 20*, ed. Kenneth Hale and Samuel Jay Keyser, 1–52. MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by Step*, ed. R. Martin et al., 89–155. Cambridge, Mass: MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. In *Ken Hale: A Life in Language*, ed. Michael Kenstowicz, 1–52. Cambridge, Mass: MIT Press.

Citko, Barbara. 2005. On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge. *Linguistic Inquiry* 36:475–496.

Fox, Danny. 1999. Reconstruction, Binding Theory, and the Interpretation of Chains. *Linguistic Inquiry* 30:157–196.

Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Goldberg, Lotus. 2005. Verb-Stranding VP Ellipsis: A Cross-Linguistic Study. Doctoral Dissertation, McGill, Monteal, QC.

Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. Samuel Jay Keyser and Kenneth Hale, 111–176. Cambridge, Mass.: MIT Press.

Hornstein, Norbert, Jairo Nunes, and Grohmann. 2005. *Understanding Minimalism*. Cambridge University Press.

Kroch, Anthony. 1994. Morphosyntactic variation. In *Papers from the 30th regional meeting of the Chicago Linguistic Society*, volume 2, 180–201. Chicago: Chicago Linguistics Society.

Lobeck, Anne. 1995. *Ellipsis*. New York: Oxford University Press.

Merchant, Jason. 1999. The Syntax of Silence: Sluicing, Islands, and Identity in Ellipsis. Doctoral Dissertation, Univeristy of California, Santa Cruz.

Merchant, Jason. 2001. *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford Studies in Theoretical Linguistics. Oxford University Press.

Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Number 43 in Linguistic Inquiry Monographs. Cambridge, Mass: MIT Press.

Phillips, Colin, and Matthew Wagers. 2007. Relating structure and time in linguistics and psycholinguistics. In *Oxford Handbook of Psycholinguistics*, ed. M. Gareth Gaskell, 739–756. Oxford University Press.

Pollard, Carl, and Ivan A. Sag. 1987. *Information-Based Syntax and Semantics*. Center for the Study of Language and Information.

Ross, John Robert. 1969. Guess who? In *Papers from the 5th regional meeting of the Chicago Linguistic Society*, ed. R. Binnick, A. Davidson, G. Green, and J. Morgan, 252–286. Chicago Linguistic Society.

Sag, Ivan. 1976. Deletion and Logical Form. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In *Sign-Based Construction Grammar*, ed. Hans C. Boas and Ivan A. Sag. Stanford, Calif.: Center for the Study of Language and Information.

Schuyler, Tami. 2001. Wh-Movement out of the Site of VP Ellipsis. Master's thesis, University of California, Santa Cruz.

Sprouse, Jon, and Diogo Almeida. 2013. The empirical status of data in syntax: A reply to Gibson and Fedorenko. *Language and Cognitive Processes* 28:222–228.

Sprouse, Jon, Carson T. Schütze, and Diogo Almeida. 2013. A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001–2010. *Lingua* 134:219–248.

Wasow, Thomas. 1972. Anaphoric Relations in English. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Mass.

Williams, Edwin S. 1977. Discourse and Logical Form. *Linguistic Inquiry* 8:101–139.