Handbook of Lexical Functional Grammar

Edited by

Mary Dalrymple

Empirically Oriented Theoretical Morphology and Syntax

language science press

Empirically Oriented Theoretical Morphology and Syntax

Chief Editor: Stefan Müller

Consulting Editors: Berthold Crysmann, Laura Kallmeyer

In this series:

1. Lichte, Timm. Syntax und Valenz: Zur Modellierung kohärenter und elliptischer Strukturen mit Baumadjunktionsgrammatiken.

2. Bîlbîie, Gabriela. Grammaire des constructions elliptiques: Une étude comparative des phrases sans verbe en roumain et en français.

3. Bowern, Claire, Laurence Horn & Raffaella Zanuttini (eds.). On looking into words (and beyond): Structures, Relations, Analyses.

4. Bonami, Olivier, Gilles Boyé, Georgette Dal, Hélène Giraudo & Fiammetta Namer. The lexeme in descriptive and theoretical morphology.

5. Guzmán Naranjo, Matías. Analogical classification in formal grammar.

6. Flick, Johanna. Die Entwicklung des Definitartikels im Althochdeutschen: Eine kognitiv-linguistische Korpusuntersuchung.

7. Zinova, Yulia. Russian verbal prefixation: A frame semantic analysis.

ISSN: 2366-3529

Handbook of Lexical Functional Grammar

Edited by

Mary Dalrymple

language science press

Mary Dalrymple (ed.). 2022. *Handbook of Lexical Functional Grammar* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog/book/312

ISBN: no digital ISBN
no print ISBNs!
ISSN: 2366-3529
no DOI
Source code available from www.github.com/langsci/312
Collaborative reading: paperhive.org/documents/remote?type=langsci&id=312

Cover and concept of design: Ulrike Harbort
Fonts: Libertinus, Arimo, DejaVu Sans Mono, Source Han Serif ZH, Source Han Serif JA
Typesetting software: XƎLaTeX

Language Science Press
xHain
Grünberger Str. 16
10243 Berlin, Germany
langsci-press.org

Storage and cataloguing done by FU Berlin

Freie Universität Berlin

Contents

Chapter 1

Clause structure and configurationality

Avery D. Andrews
The Australian National University

LFG differs strongly from Mainstream Generative Grammar in basing its theory of clause structure on overt surface appearance, as would be input to a parser, rather than the outputs of a derivational process that might produce these structures. This leads to a number of differences, such as a much smaller number of functional projections, and more emphasis on a typology of overt structures, including the inclusion of special provisions for 'nonconfigurationality'. In this chapter, we examine LFG analyses resulting from this perspective from the beginning of LFG in Bresnan (1982b) through to the theory as presented in Bresnan et al. (2016).

1 Introduction

Because LFG is based on using phrase-structure rules (PS rules, with a substantial involvement of universal principles) to provide a direct description of overt structure, with more abstract levels such as f-structure determined by annotations on these rules, they carry a major burden in describing the organization of clause-structure. In particular, it is not possible to invoke 'movement' mechanisms to get things into their surface positions. Rather, with one plausible minor exception,[1] the PS rules have to put everything in the exact positions where they are found overtly, albeit with the possible help of filtering by other components of the grammar.

Partly for this reason, in LFG. the treatment of clause structure has been from the beginning closely involved with the concept of 'nonconfigurationality', a term coined by Hale (1981) to refer to situations where linear order does not determine grammatical relations in any clear way, and where, in addition, referring

[1] 'Second position' items, as discussed below.

Avery D. Andrews. Forthcoming(a). Clause structure and configurationality. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

expressions and other clausal constituents are sometimes discontinuous. Transformational Grammar and its more direct descendants, sometimes called 'Mainstream Generative Grammar', avoided having a problem with this by proposing that underlying clause structures were transformed into overt ones by the application of 'scrambling' rules (in later work, sometimes relegated to the 'phonology'). But when phrase structure rules are to be used for providing a direct characterization of overt structures (with filtering by other components), this is not possible.

Another relevant issue is the position of subjects. If a language appears to have verb phrases that exclude an apparent 'subject' NP argument, then the PS rules have to provide a position for NP external to VP, while if a putative subject is freely intermixed with other arguments, then we probably do not want to have a full-sized VP containing the verb and other arguments, but rather have the verb and the arguments appear directly under S.[2] The theory should then plausibly provide two possibilities along the lines of (a) and (b) below, where (a) puts an NP in front of a VP, while (b) has no VP:[3]

(1) a. S $\longrightarrow$ NP VP
    b. S $\longrightarrow$ NP* V NP*

A comma could also be included in the first rule to allow the NP to precede or follow the VP, as in Makua (Stucky 1983).

A further general consequence of the archicture is that because many phenomena including agreement, case-marking and anaphora can be largely or entirely described in terms of the more abstract level of f-structure, the sources of evidence for phrase structure are more limited than they are in Mainstream Generative Grammar. We cannot, for example, easily use coreference phenomena to motivate phrase structures in which one object in a double object construction c-commands another, but would need a very extensive (and therefore fragile) argument to show that other levels such as f-structure are not sufficient.

In this chapter, I take a predominantly historical approach to clause structure in LFG, on the basis that a reader might want to engage in literature from any

[2] As we shall see, languages sometimes have a smaller verbal phrase containing the verb and certain other material, but not, normally, the object; this is sometimes treated as a VP, and sometimes as a different kind of verbal phrase, often symbolized as $\overline{\mathrm{V}}$.

[3] We don't use the 'ID/LP' notation of Gazdar (1982), first applied to LFG by Falk (1984), to allow the daughters of S to appear an any order in (1b) (S $\longrightarrow$ NP*, V) because of the plausibility of interpreting this as a possibly null string of NPs, either before or after one V. There are issues worth looking into further here, but not in this chapter.

time from the early 1980s to the present, and therefore find useful some discussion of what kinds of proposals were being made at different times. I will divide the history of clause structure in LFG into three periods so far, with the possibility of a new one starting now. In the first, from the beginnings of LFG in the early 1980s to the early 1990s, some version of the X-bar theory was assumed, but there was little explicit discussion about exactly what that version was. The 90s constitute a transitional period, in which both the 'extended projections' from Minimalism and ideas from Optimality Theory are taken on. The third period plausibly begins with Bresnan's (2001) theory of structure-function mappings, which can be seen as a consolidation of the work of the transitional period, based on a division between 'endocentric' and 'exocentric' structure, the former obeying the X-bar theory with functional structures, the latter not, along with some principles derived from Optimality Theory, such as Economy of Expression. This approach has persisted with little alteration through Bresnan et al. (2016) to the present. Since it is the result of multiple analyses of different languages by a number of people, I will call it the '2001 Synthesis'. More recently, a fourth period may have begun with Lowe & Lovestrand (2020) and Lovestrand (2022), a thorough revision of the underlying phrase-structure theory making greater use of architectural ideas of LFG rather than simply applying some version of mainstream X-bar theory. However there hasn't yet been substantial work on a variety of clause structures in this new approach.

In this chapter, I consider early LFG in the first section, the transitional period in the second, and the 2001 Synthesis in the third. Then, in the fourth section, I review some of the earlier and transitional systems in light of the 2001 Synthesis, and discuss the revisions that are thereby motivated, and conclude with a speculation about S derived from a modification in the new X-bar framework made in Lovestrand (2022).

2 Early LFG

In early LFG, it was assumed that some version of the X-bar theory was correct, but no attempt was made to seriously formalize or revise the proposals that were standard at the time. Bresnan (1982a: 354-356),[4] which also appeared in the foundational LFG collection (Bresnan 1982b), developed a fairly permissive theory of 'structure-function mappings', many provisions of which have persisted until now, which constrain how c-structure nodes can be annotated to produce f-structures, heavily influenced by Ken Hale's ideas about 'configurational' vs

[4]Pp. 296-299 in the version of the paper that appears in Bresnan (1982b).

'non-configurational' languages. The other papers in the 1982b collection tended to conform to these ideas without discussing them explicitly.

In these papers, the languages treated as configurational (where grammatical relations are largely coded by phrase-structure position) were English, French, Russian and Icelandic, all of them analysed as SVO, while the one analysed as nonconfigurational was Malayalam, predominantly verb-final. Most of the SVO languages were analysed with distinct S and VP rules like these:

(2) a. S $\longrightarrow$ NP VP
$(\uparrow \textsc{subj})=\downarrow$ $\uparrow=\downarrow$

b. VP $\longrightarrow$ V NP …
$\uparrow=\downarrow$ $(\uparrow \textsc{obj})=\downarrow$

But for Icelandic, Andrews used a flat S rule with subject first (consistent with Bresnan's (1982a: 354/296)[5] schema, which he does not however cite or discuss).

From that time to the present, a major concern of LFG authors has been to marshal arguments that c-structure relationships were neither a necessary nor sufficient basis for assigning grammatical relations in the general manner suggested by Chomsky (1965) and persisting through to Chomsky (1981), which appeared in the early days of LFG, and beyond. So considerable attention was paid to arguing for the nonexistence of a VP node in languages that appeared not to have a VP constituent, and a contrastive sketch of configurational English versus nonconfigurational Warlpiri constitutes the first chapter of the high-level LFG textbook (Bresnan et al. 2016).

In the following two subsections, we consider first Malayalam, and then Warlpiri, the other nonconfigurational language to which considerable attention was developed in this period; discussion of Warlpiri was included in Bresnan (1982a) and several other chapters in the Bresnan (1982b) collection.[6]

2.1 Malayalam

Nonconfigurational Malayalam, an essentially verb-final language, was analysed by Mohanan (1982) using the clausal category S, introducing its arguments as an unstructured, possibly empty, string of NPs. It was proposed to have a verb-final S rule, with the NPs introduced by a rule whose first version was (Mohanan 1982: 507):

(3) S $\longrightarrow$ NP* V

[5] The page before the '/' is for the journal version, the one after for the page in the 1982 volume.

[6] Warlpiri is discussed in Simpson & Bresnan (1983), but the focus of that article is grammatical relations rather than clause structure as such.

Annotations for this version of the rule were not specified. Later (Mohanan 1982: 542–643), a series of annotations are proposed associating specific grammatical functions with cases and other properties such as animacy, for example:

(4) (↑ OBJ)=↓
(↓ CASE)=ACC
(↓ ANIM)=+

This is one of the alternative annotations to the NP in rule (3), allowing any NP in the series to have any grammatical function, subject to filtering by other constraints of LFG.

The flatness of the structure given by (3), with its absence of a VP dominating objects, is motivated by several arguments, one of which is the fact that the bearers of the grammatical relations can appear in any order. But the most important one, according to Mohanan (1982: 526-533), is the workings of a clefting phenomenon which allows all and only direct daughters of the S to be clefted, but not subconstituents of anything, such as possessors or objects of prepositions. The construction is effected by suffixing *-aanə* ‘is’ to the last word of the clefted constituent, and *-ṯə* ‘is’ to the verb. Some relevant examples are:

(5) Malayalam (Mohanan 1982: 528-529)

a. Kuṭṭi iṉṉale ammakkə anaayey-aaṇə koṭuṯṯa-ṯə.
child yesterday mother.DAT elephant-is gave-it
‘It was an elephant that the child gave to the mother yesterday.’

b. kuḷaṯṯil wecc-aaṇə jooṇinte kuṭṭi aanaye ṇuḷḷia-ṯə.
pond at-is John.GEN child elephant pinched-it
‘It was at the pond that John’s child pinched the elephant.’

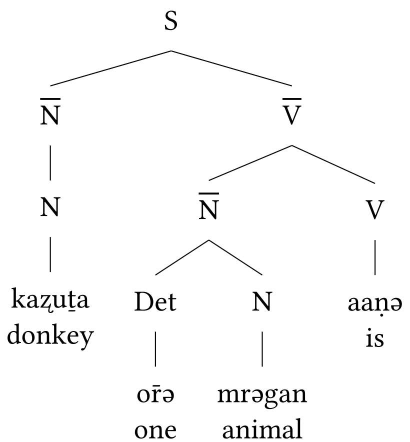
c. *kuḷaṯṯil aaṇə weccə jooṇinte kuṭṭi aanaye ṇuḷḷia-ṯə
pond is at John.GEN child elephant pinched-it

In (a), the object is clefted, in (b) the PP, but when we try to cleft the object of the PP (or, not shown, the possessor of the object), the result is bad. The proposed generalization is that you can cleft a direct constituent of S, but not a subconstituent of S, which precludes the existence of a VP sitting on top of the verb and its object.

There are two further significant elaborations in Mohanan’s analysis. First, it turns out that Malayalam is not actually strictly verb final: in main clauses, the verb can be followed by additional NPs, but this normally requires putting a heavy nuclear pitch on the verb, wiping out the word melodies on the following NPs, and lengthening the vowel of the verb, evidently with some kind of contrastive meaning (Mohanan 1982: 511). This is furthermore not possible with certain kinds of subordinate clauses. Mohanan suggests an analysis involving a 'scrambling rule', which applies to the S rule, but this is not an option that is available in the LFG formalism, and these NPs need to be introduced in their surface positions, presumably with annotations connecting them to discourse functions (Zaenen forthcoming [this volume]).

The second elaboration is that there is a kind of verb phrase, but it contains only the verb and certain additional elements, such as NPs and PPs used to form Copula Constructions and Complex Predicates (Mohanan 1982: 513-524). An example is:

(6)



'The donkey is an animal.' (Mohanan 1982: 513-544)

These are plausibly VPs (maximal projections of V, Malayalam having only one phrasal projection layer), which are however very restricted in what kinds of constituents and bearers of grammatical relations they can introduce. On the other hand, some superficially similiar complex verbal constituents in other languages do not appear to contain any maximal projections, and so can be analysed as V^0 nodes with adjoined 'non-projecting' lexical nodes (Toivonen 2001), also discussed in Belyaev forthcoming [this volume]. One example is Complex Predicate constructions in Japanese (Ishikawa 1985, Matsumoto 1996), discussed in Andrews forthcoming(b) [this volume], and another is Warlpiri preverbs, considered shortly below.

Another important characteristic of Malayalam and other nonconfigurational languages is that all or most arguments of verbs can be freely omitted, and understood as if they were represented by pronouns. Mohanan (1982: 544) discusses this briefly, and provides a few examples including:

(7) Malayalam (Mohanan 1982: 544)
Roṭṭi ewiṭe? Kuṭṭi ṯiṉṉu.
bread where? Child ate.
'Where is the bread? The child ate it.'

The LFG treatment of this kind of phenomenon uses lexical rules of 'anaphoric control', developed in Bresnan (1982a), which optionally add a pronominal f-structure specification to the lexical entries of verbs. In the above case, this would be ($\uparrow$ obj pred)='pro'. In Malayalam, anaphoric control applies to the grammatical relations subj, obj, and obj$_\theta$ ('indirect object'; obj2 in the original). Anaphoric control is not restricted to nonconfigurational languages, and is subject to numerous variations in different languages. In English, anaphoric control is predominantly used with subjects of nonfinite verbs, but has in addition some more limited versions. For example, an inanimate subject can be omitted but understood as if 'it' refers to something that the group of people being addressed are looking at:

(8) Looks bad! [said by one of a group of people staring at an engine with smoke coming out of it]

The typological range of such constructions and their semantics deserves further investigation in LFG.

2.2 Warlpiri

A more extreme form of nonconfigurationality was addressed in the comprehensive analysis of Warlpiri provided by Simpson (1983), later published with substantial revisions as Simpson (1991), and also discussed by Nordlinger forthcoming [this volume]. Warlpiri differs from Malayalam (and Japanese) in a number of ways:

(9) a. In finite clauses, there is no constraint on NPs coming after the verb.
b. NPs can be discontinuous, with different components appearing separated by other constituents of the clause.
c. There is an 'auxiliary', obeying a complex 'second position' constraint.

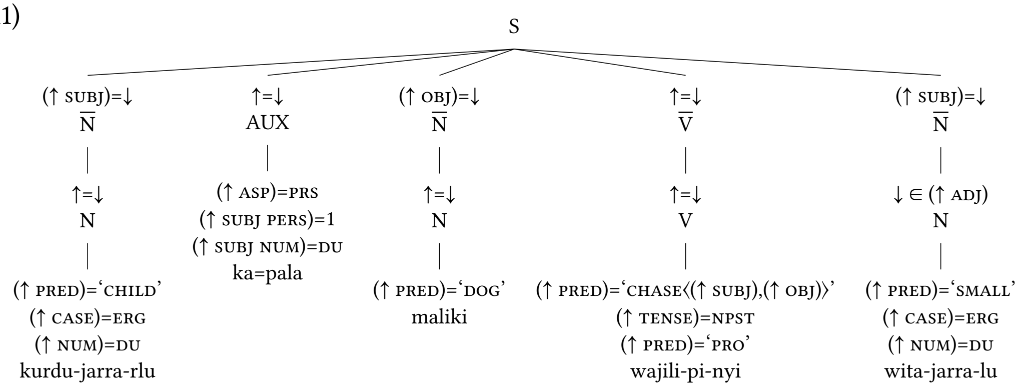
All of these phenomena are illustrated in this example:

(10) Warlpiri
Kurdu-jarra-rlu ka-pala maliki wajilipi-nyi wita-jarra-rlu.
Child-DU-ERG PRS-DU dog(ABS) chase-NPAST small-DU-ERG
'(The) two small children are chasing the dog.'
'(The) two children are chasing the dog and they are small.'

The theoretically most interesting point is (9b).

Hale (1981) described the two nominals 'child' and 'small' in (10) as being interpretable in two ways, one being the 'merged' interpretation, shown in the upper gloss, where the two components are interpreted in the same way as a normal NP in English, and the other being the 'unmerged' interpretation, shown in the lower gloss, in which the second nominal is interpreted as a secondary predicate giving additional information.

To capture the merged interpretation, Simpson proposed that an NP (in her analysis, for Warlpiri, a $\bar{\text{N}}$, the language having no evidence for either a specifier level for lexical projections, or any kind of DP) could expand to an adjunct alone with no head (as well, as of course, to a single head), so that two independently introduced components of an f-structural NP-correspondent could merge, as reflected in the annotations on the tree for example (10):[7]

(11)


(Simpson 1991: 283)

The annotations on the first and last $\bar{\text{N}}$s allow them to unify into a single f-substructure, the value of SUBJ:

[7] Note that Simpson assumed a rule or convention that would copy agreement features from the f-structure correspondent of an NP down to those of its ADJ-members. There is furthermore a problem with the positioning of AUX that we will consider shortly.

(12)
$$\left[\begin{array}{ll} \text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{`CHILD'} \\ \text{ADJ} & \left\{\left[\begin{array}{ll}\text{PRED} & \text{`SMALL'}\\ \text{CASE} & \text{ERG}\end{array}\right]\right\} \\ \text{CASE} & \text{ERG}\end{array}\right] \\ \text{PRED} & \text{`CHASE}\langle(\uparrow \text{SUBJ}),(\uparrow \text{OBJ})\rangle\text{'} \\ \text{ASP} & \text{PRS} \\ \text{TENSE} & \text{NPST} \\ \text{OBJ} & \left[\text{PRED `DOG'}\right] \end{array}\right]$$

Similarly to Malayalam, Warlpiri also makes extensive use of anaphoric control, although many arguments not expressed as NP will receive morphological registration in the AUX constituent, which we will discuss shortly.

There are two further characteristics of Simpson's analysis that interact with each other, and have been important in later developments. The first is that similarly to Malayalam, Simpson analyses Warlpiri as having an inner VP, symbolized as $\overline{\text{V}}$, containing the verb and certain other elements, especially 'preverbs', as discussed in Simpson (1991: 111). However none of these items can contain complex phrasal constituents, and it is therefore probably better to treat them as non-projecting words adjoined to V.

The second additional feature of Warlpiri, already seen in the Warlpiri tree structure (11), is the 'AUX' constituent. This was postulated for Warlpiri[8] in the classic article of Hale (1973: 310), as a constituent containing three kinds of constituents, all optional. First comes a 'complementizer', which has a variety of functions, later called the 'augment' by Laughren (2002). We will follow this usage. Next comes the 'base', which is one of the tense-aspect markers *-ka* 'present imperfective' or *-lpa* 'past imperfective'.[9] Finally come agreeement markers, for subject and object. Hale (1973: 312) proposed that if the augment+base sequence was less than two syllables in length, then the auxiliary could not appear in initial position, but only after some other, evidently first, element of the clause.[10]

Simpson (1991: 83) proposes that the underlying position of the AUX is initial, as specified by this rule:

(13) $\text{S} \longrightarrow (\text{AUX}) \quad \alpha \quad (\alpha)^*$
$\alpha = \overline{\text{V}}, \overline{\text{N}}, \text{Particle}$

[8] AUX as a node-type was widely proposed at that time for the analysis of many other languages, including English, as discussed extensively for example by Akmajian et al. (1979).

[9] Hale treated the future marker *kapi* as a base, but Legate (2008) shows that it is actually an augment.

[10] With the exception that certain items, such as topics set off with a pause, were seen as appearing outside the basic clause structure, allowing the AUX to appear in apparent third position if these items were included.

(↑ ASPECT) ↔ (↑ TENSE)
Assign (↑ G)=↓ freely
(where G stands for any grammatical function)

The second line tells us that α can be any of three kinds of constituents, while the third adds the information that ASPECT is specified if and only if TENSE is, a move that has the effect of requiring a verb to be present if an AUX is, by mechanisms we will not consider here. Finally, the last two lines allow constituents to be annotated freely as either heads (↑=↓) or arguments bearing any grammatical function, providing a high degree of nonconfigurationality, including generating multiple $\overline{\text{N}}$ nodes with the (↑ SUBJ)=↓ annotation, allowing NP-splitting.

AUX is then put into second position in most examples by first allowing all AUXs to be classified as enclitics, but obligatorily so for the ones with monosyllabic bases (Simpson 1991: 69). Then the clitics are postposed to after the first phonological unit by a rule of sentence-phonology:

(14) Enclitization Rule:
]AUX [α] [α]* → α+AUX [α]*
(the ‘]’ in front of AUX represents that the AUX has enclitic status)

It is perhaps worth noting that for all examples where AUX appears in second position, the trees are also written with AUX in second position, including the effects of the Encliticization Rule in the diagram. In later work, various aspects of this proposal are questioned and revisions proposed, as we will see later, in section 3.2 of this chapter.

A final observation I will make about Warlpiri concerns the treatment of discontinuous NPs. The LFG analysis permits an NP to be split into any number of separated components, all of which can contribute to a single f-structure with a nominal PRED-feature, subject to no constraints of any kind. There is considerable work showing that this appears to be false, including Schultze-Berndt & Simard (2012), Schultze-Berndt (2022) and Louagie & Verstraete (2016). Rather, discontinuous NPs appear to be associated with a range of specific discourse functions (and to furthermore be rather rare, probably not more than 1% of NPs), and examples where an NP is split into three or more components at the same level in f-structure (e.g. demonstrative, modifier(s) and head noun) do not appear to be attested in the literature so far. Unfortunately, none of this recent literature discusses Warlpiri, but I am aware of no triply split NPs in Simpson (1983, 1991), Nash (1986) or Laughren (1989),[11] nor in the discussion of discontinuous NPs in

[11]Who also provides an example of discontinuous participial VPs, which Simpson (1991) argues are nominalized.

Latin and Classical Greek provided by Devine & Stephens (2000). Therefore, the proposition that discontinuous constituents are limited to two components is a proposition worth further investigation.

Similarly to Mohanan, Simpson provides some arguments for a flat structure and no VP, but they are more complex than Mohanan's cleft argument, one involving coreference, another involving nonfinite constructions, and will not be discussed here.

Concluding our discussion of the first period, we find a basic distinction between configurational and nonconfigurational encoding, the former associated with SVO languages, usually associated with a VP, the latter with verb-final or verb-anywhere languages, often with flat structure. It was usual to assume some kind of X-bar theory, without being very specific about the details. There were however some intimations of later developments, such as Falk's (1984) analysis of the English Auxiliary system, in which, influenced by Jackendoff (1977), he treats auxiliaries as a lexical category M, taking VP as a complement and the subject as a specifier. This can be seen as an early version of the idea of the functional projection IP, with its binary branching auxiliary structure as opposed to the flat trinary structure NP AUX VP proposed for S by Akmajian et al. (1979), which is taken up in the third period, the 2001 Synthesis (section 4).

3 The transitional period

The characteristics of the transitional period are (a) the introduction of the concept of 'functional categories' from the GB and Minimalist frameworks, a feature which has remained; (b) considerable experimentation with ideas from Optimality Theory, which appears to have fallen off to some degree, although is still being explored (Kuhn forthcoming [this volume]). The dating of the period is difficult, since use of functional categories could be said to have been anticipated by Falk (1984), while drafts of what I take to be the initiation of the third period, Bresnan (2001), were available to some workers as early as 1996 (Nordlinger 1998: 15). But I will here take it to begin with Kroeger (1993),[12] where the functional categories I and C are adopted from Government-Binding theory, and continue until the publication of Bresnan (2001). Many of the features of what I will in the next section call the 2001 Synthesis are present in the analyses of the transitional period, to the point that some discussions could be put in either section. But here we take an essentially chronological view surveying phenomena that lead to the

[12] The Stanford PhD thesis upon which the book was based is from 1991.

2001 Synthesis, presenting the resulting system in section 4, together with some new analyses as well as possible updates to older ones.

3.1 Kroeger (1993) on Tagalog

Tagalog is a verb-initial language with preverbal discourse positions, and, according to Kroeger, no evidence for a VP, and some evidence against (we will consider a different view below), but evidence for some other predicate phrases, namely, PP, AP and NP. Kroeger analyses these patterns by taking from Chung & McCloskey (1987) the idea of a special category 'S' that can constitute a predication ('small clause') without providing TAM information, and combining this with the notion of a 'functional projection' IP, where the S appears as the complement of I. He also departs from Chung and McCloskey to allow S to expand to a lexical predicate and multiple arguments, rather than only to a subject NP and a predicate phrase.

For clauses with an aspect-marked verb, this verb appears in the I head of the functional projection IP (INFL for Kroeger), while the arguments and adjuncts appear in free order under S, although there are some tendencies (p. 111):

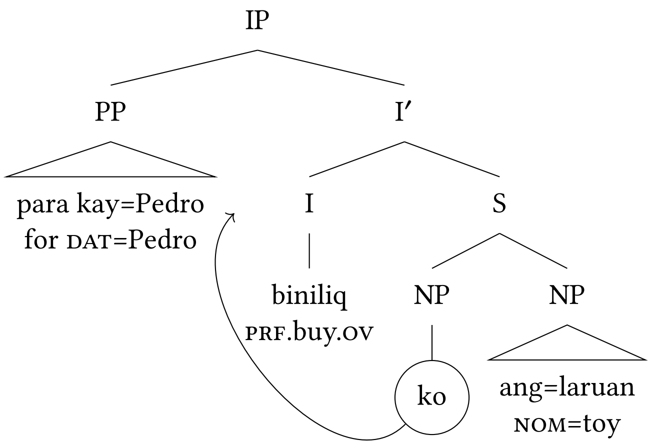
(15) a. the 'Actor' (non-nominative Agent marked with *ng*, or *ni* if a proper name) tends to come first.
b. the 'Nominative' (marked with *ang*, or *si* if a proper name) tends to come after the other arguments.
c. "heavier" NP's tend to follow "lighter" NPs.

Kroeger does not actually give a structure for a sentence with multiple NP arguments: the closest is one with a clitic Actor and a focussed adjunct in the specifier of IP. Clitics are however subject to very interesting positional restrictions which in this case put the clitic *ko* after the SPEC of IP, as indicated in this example (Kroeger 1993: 129), where the original tree has INFL rather than just I:[13]

(16) Tagalog
a. [Para kay=Pedro] ko binili ang=laruan.
for DAT=Pedro 1SG.GEN PRF.buy.OV NOM=toy
'For Pedro I bought the toy.'

[13]The clitic rule, discussed in Kroeger (1993: 119-123), is: "Clitics appear immediately after the first daughter of the smallest maximal projection that contains them" (but there are some apparent exceptions). The 'object-focus' suffix glossed OV indicates that the Patient of the verb is the 'grammatical subject', traditionally called the 'focus' in Philippine linguistics, analysed by Kroeger as the SUBJ grammatical function in his LFG analysis.

b.

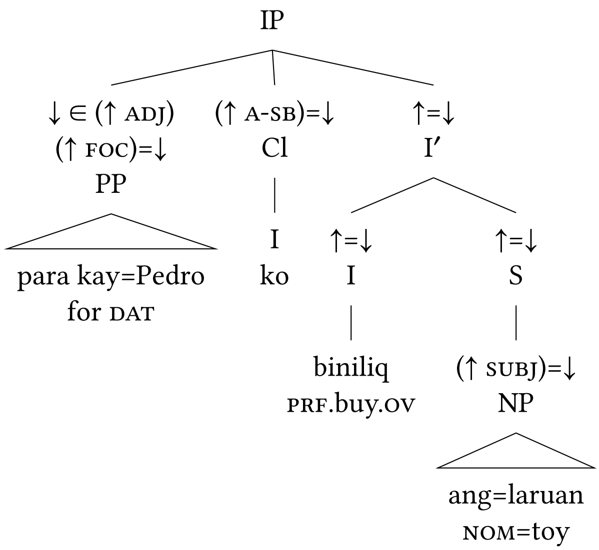


In this tree, the *ko* is ascribed to a position under S where a full NP argument could appear (initial in accord with the ordering tendencies noted above), with the arrow indicating some kind of clitic displacement to after the first constituent of IP. Another thing to note is the use of prime notation rather than bars, so I′ instead of $\bar{\mathrm{I}}$. In this chapter, I will use whichever notation is employed by the original author.

Since functional projections have both the head and the complement annotated with ↑=↓, there is no problem with assembling the f-structure for a c-structure such as (16b). The initial PP is an instance of what Kroeger and much subsequent work has called 'Adjunct Fronting', which applies to the bearers of non-core GFs, that is, adjuncts, oblique arguments and adverbials (Kroeger 1993: 43). We can analysis this by allowing SPEC of IP to receive one of the non-term grammatical functions, together with some sort of focus-like discourse function. I can't find an explicit statement of this in the literature, but it appears to be an implication of the discussion in Kroeger and other sources such as Gerassimova & Sells (2008) that the construction is clause-bounded, since only subjects are said to be extractable from subordinate clauses, and only from ones that are themselves subjects (Kroeger 1993: 210,215-221).

So we can propose an annotated structure like (17) below for the example, with the clitic *ko* placed overtly in its second position, without concern here for what constraints put it there, an issue discussed extensively by Kaufman (2010), but too complex to attempt to provide an updated account of here. We will notate it as 'A-SB', for the non-subject Agent in Philippine languages, following the choice of Manning (1996) for the Agent in syntactically ergative languages such as Inuit:

(17)



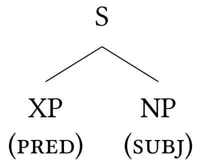
Given appropriate lexical entries, this will produce the following f-structure:[14]

(18)
$$\left[\begin{array}{ll} \textsc{foc} & \left[\begin{array}{ll} \textsc{pred} & \text{‘for’} \\ \textsc{obj} & \left[\textsc{pred}\ \text{‘Pedro’}\right] \end{array}\right] \\ \textsc{subj} & \left[\begin{array}{ll} \textsc{pred} & \text{‘toy’} \\ \textsc{case} & \textsc{nom} \\ \textsc{spec} & \textsc{def} \end{array}\right] \\ \textsc{pred} & \text{‘buy}\langle(\uparrow \textsc{a-sb})\ (\uparrow \textsc{subj})\rangle\text{’} \\ \textsc{asp} & \textsc{prf} \\ \textsc{a-sb} & \left[\begin{array}{ll} \textsc{pred} & \text{‘pro’} \\ \textsc{pers} & 1 \\ \textsc{num} & \textsc{sg} \\ \textsc{case} & \textsc{gen} \end{array}\right] \\ \textsc{adjuncts} & \{[\ ]\} \end{array}\right]$$

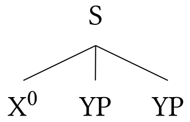
In addition to verbal clauses, there are clauses with adjectival, nominal and prepositional predicates. Kroeger argues that these show a different pattern, where some phenomena of clitic placement are said to show that the main predicate can either appear on its own as first daughter of S, or as head of a phrase that contains its complements, with the subject final under S in the former case, as indicated by these (somewhat abbreviated) structures (Kroeger 1993: 133):

[14] This structure uses the older treatment of discourse functions such as FOCUS as grammatical functions in f-structure, along the lines of Bresnan (2001: 97-98) or Bresnan et al. (2016: 97). Kroeger (1993) does not provide any specific f-structures. For contemporary views, see Zaenen forthcoming [this volume].

(19) a.



b.



Unfortunately, Kaufman (2010: 259-260), working within the Minimalist framework, finds that the clitic facts cited by Kroeger[15] do not appear to be representative, in ways that undermine Kroeger's analysis. Since this is of some interest for the history of the subject, I think it is worth considering the examples, in the hope that it will be further investigated in LFG.

What Kroeger says is that with nominal, adjectival and propositional phrasal predicates, a personal pronoun clitic can appear either at the end, or after the predicate word, illustrated here for PP:

(20) Tagalog

a. Galing sa=Maynila siya.
   from DAT=Manila 3SG.NOM
   'He is from Manila.'

b. Galing siya sa=Maynila.
   from 3SG.NOM DAT=Manila
   'He is from Manila.'

But with a verbal main predicate, the sentence-final position is impossible:

(21) Tagalog

a. ??Hinangkan ng=nanay ako.
   PRF.kiss.DV GEN=mother 1SG.NOM
   'I was kissed by mother.'

b. Hinangkan ako ng=nanay.
   PRF.kiss.DV 1SG.NOM GEN=mother
   'I was kissed by mother.'

This is to be explained by:

[15] Originally from Schachter & Otanes (1972) and Sityar (1989).

(22) a. A principle to the effect that the clitics are placed after the first constituent in the domain they apply to, which is the IP.

b. The two constituent structures in (19) are available for nonverbal predicates, but only the flat one of (19b) for verbal predicates.

However, Kaufman finds that there is no significant difference between the clitic final position for verbal and nonverbal predicates: both are pretty bad. He also argues that Kroeger's generalization about where the clitic goes is insufficient, and proposes something different, well beyond the scope of this chapter. This leaves the flat rule (19b) motivated by the evidence, but not (19a), for Tagalog.

Nevertheless, there is motivation for structures of the general form of (22a) elsewhere in the Austronesian language family: Dalrymple and Randriamasimanana use it in their XLE grammar of Malagasy,[16] and Liu (2017) presents an LFG analysis of Squliq Atayal arguing on various grounds for this structure. Finally, Kaufman & Chen (2017) review a rather long tradition of argumentation in Austronesian historical syntax for the position that structures with a clause-initial predicate phrase and a following subject are the original form of the 'Philippine type' of which Tagalog is the most often discussed exemplar.

It is perhaps worth emphasizing that the 'subject' in Philippine languages is not the classic subject of western European languages with its strong association with semantic Agent properties, but rather the 'Pragmatic Peak' of Foley & Van Valin (1984), drawing heavily on earlier work by Paul Schachter and Edward Keenan (Keenan 1976; Schachter 1977), or the 'g-subject' of Manning (1996). These have an association with topic-like pragmatic functions, but not with agentivity. Indeed, the constructions with patient as subject tend to be more common than those with agent, and are closer in form to the proposed diachronic original, as discussed by Kaufman & Chen (2017).

A final point is that above IP, arguably the domain of clitic positioning, there are projections treated by Kroeger as CP, and irrelevant to clitic positioning (resulting in 'third position' phenomena, and this general approach is also adopted by Kaufman (2010).

3.2 Warlpiri: Nonconfigurationality in Australian Languages

Austin & Bresnan (1996) update Simpson's (1983, 1991) analysis of Warlpiri to use the functional projection IP to house the material constiting the auxiliary, and also give an extended treatment of Jiwarli. Two important differences between Warlpiri and Tagalog with respect to I are:

[16] http://users.ox.ac.uk/~cpgl0015/pargram/; argumentation is however not provided.

(23) a. In Warlpiri, the verb does not appear in I.[17]

b. I nodes in Warlpiri that meet a certain condition, to be discussed immediately below, cannot appear initially, at least in a phonologically independent clause.

c. Most items, including the verb, can appear in front of the auxiliary material (contents of I). This analysis proposes two mechanisms for how this happens: NPs appearing in SPEC of IP, and a prosodic inversion operation for verbs and preverbs.

The nature of condition (b) calls for some discussion.

As mentioned above, earlier work from Hale (1981) to Simpson (1991) proposed that the AUX had to appear in second position if the augment+base was monosyllabic, but Laughren (2002) shows that the bisyllabicity condition is not correct, on the basis that the complementizer *yi-* 'for, since' followed by a null base can appear initially, as long as the entire auxiliary, including agreement markers, is bisyllabic (all the other augments are bisyllabic). This is also the case for the present imperfective base *-ka*, but apparently not for the past imperfective base *-lpa*.[18] So I suggest that the actual condition is a combination of phonology and morphology:

(24) In order for I to be overtly initial in Warlpiri, its contents must:

a. be at least bisyllabic

b. be phonologically well-formed as a word (initial *-lp* clusters are not allowed)

c. have an overt augment+base (either augment or base is sufficient, as long as something appears).

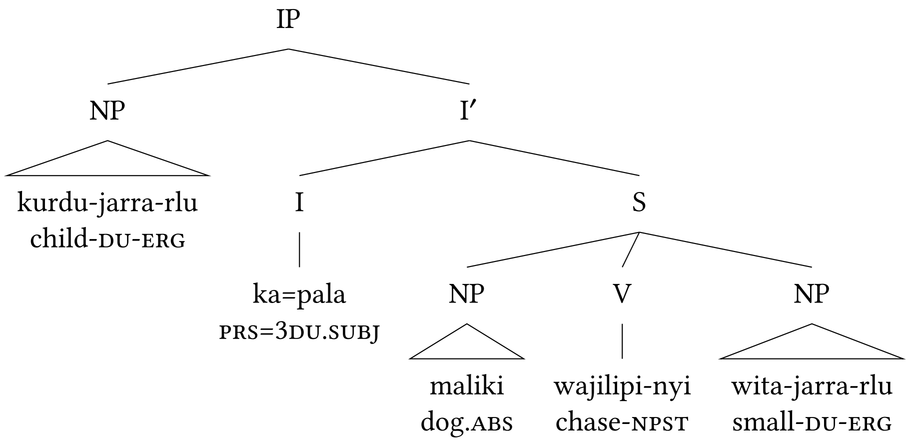
I will call an auxiliary that meets these conditions 'heavy', and one that doesn't, 'light'. So our basic generalization is that only heavy auxiliaries can be initial in the sentence.

With this issue considered, we examine the basic sentence structure that Austin & Bresnan propose for Warlpiri, with an initial NP in the SPEC of IP position, and auxiliary material in the following I:

[17] Legate (2008) proposes in her Mimimalist analysis that the verb can be attracted to I, but in LFG, there is no advantage to be obtained by allowing it to appear there, as we will see below.

[18] But the existence of possible exceptions is discussed in Laughren (2002: 125, footnote 19).

(25)



This illustrates the first mechanism whereby the auxiliary material can appear in second position, but is not plausible for cases when the verb is in first position, because a lexical category should not be able to occur in SPEC position in Warlpiri (Austin & Bresnan 1996: 226).

To deal with verb-initial sentences, Austin and Bresnan propose the rule of prosodic inversion from Halpern (1995), which moves the contents of I to a position after the verb, or, sometimes, after the initial part of a complex verb. This division of labor permits the inversion rule to apply to a considerably more restricted range of cases than Simpson's Encliticization, removing the need for it to swap the auxiliary around multiword phrases.

Examples of multiword phrases are modifier+modified nominal constructions with case marking, on either only the last or both elements, and also coordinate NPs (auxiliaries in boldface to make the examples easier to follow):

(26) Warlpiri

a. Kurdu(-ngku) wita-ngku=**ka** maliki wajilipi-nyi
child(-ERG) small-ERG=PRS dog(ABS) chase-NPST
'The small child is chasing the dog.'
(Nash 1986: 159-160, citing Hale (1981))

b. Karnta-ngku manu ngarka-ngku=**pala** kurdu nya-ngu
woman-ERG and man-ERG=3.DU.SUBJ child.(ABS) see-PST
'The man and the woman saw the child.' (Nash 1986: 177)

Since these multi-word NPs can be generated in a position before I, we do not need to have any rule putting the auxiliary after them.

However, sentences with something other than an NP appearing before the auxiliary pose some tricky problems. If all and only the things found in this position were verbs, we could suggest that a V could optionally be adjoined to I, appearing in front of the auxiliary material. But this proposal faces two problems. One is that if the preverb appears in its normal position before the verb, then the auxiliary can appear between them, as long as it doesn't contain an augment (but polysyllabic auxiliaries with base *ka* are fine.):[19]

(27) Warlpiri (Austin & Bresnan 1996: 227)

a. Rambal-luwa-rnu=**rna=rla=jinta** marlu-ku
mistake-shoot-PST=1SG SUBJ-3SG DAT-DD kangaroo-DAT

b. Rambalpa=**rna=rla=jinta** luwa-rnu marlu-ku
mistake=1SG SUBJ-3SG.DAT-DD shoot-PST kangaroo-DAT

'I shot at a kangaroo and failed.'

If the auxiliary intervenes, a preverb whose stem ends in a consonant must end in the stem-extender *-pa*, which it can do anyway. This indicates that one requirement for intervention is that the preverb must be construed as in some sense being an independent word.

Another is that for those 'productive' preverbs that can appear after as well as before the verb, the auxiliary material seems to almost obligatorily appear after just the verb rather than after the whole verb+preverb combination when the preverb comes second (Simpson 1991: 117):

(28) Warlpiri (Nash 1986: 52)
Yani=**rli** wurulypa
go-NPST=12 seclusion
'Let's go and hide.'

Nash observes that the ordering *yani-wurulypa=rli* occurs once in a text, but seems much less common than the other possibilities, while Simpson characterizes it as "hardly ever found, and it is usually rejected by speakers". This is a problem for any analysis which puts V in front of I in c-structure, unless we assume a category other than V to dominate the V+Preverb order.[20]

[19] The suffix *-jinta* glossed DD indicates that this is a 'failed effect' construction discussed in Hale (1973: 336), in which the object is marked dative, indicating that the action indicated by the verb did not succeed.

[20] On the other hand, Laughren (2002: 100) provides such an example without comment, but with a heavy auxiliary which could not be inserted into the verb.

So the conclusion is that an inversion rule along the lines of Simpson's encliticization is needed, but applying in a narrower range of cases, more consistent with being a morphological or phonological operation. Austin and Bresnan assume that it is the 'prosodic inversion' of Halpern 1995, which only applies as a last resort. This is supposed to explain why the clitics can't be inserted into phrasal units as in (27), but there is a problem here in that the structure in which the phrasal NP is sitting in SPEC of IP is different from one where it is initial in S right after I, so it is not clear that a 'last resort' restriction can apply in a well-defined manner.[21] A further problem is that it appears to be fine for a heavy auxiliary to appear after the verb (Laughren 2002: 97), so a last resort restriction won't work.

What I suggest is an inversion rule which can be formulated like this:

(29) I V/Pvb
1 2 ⇒ 2+1
Subject to restrictions that need further investigation.

The category restriction is sufficient to prevent I from being inserted into an NP (a restriction documented at considerable length by Laughren), and another restriction, not formalized here, states that an auxiliary with an augment cannot be inserted into a verb.

A final problem discussed by Laughren (2002) and Legate (2008) concerns evidence that there is more than one functional projection dominating S, in spite of no evidence of two distinct head positions being occupied in the same clause. This is the interaction of topicalization and focus in questions. In (30) below, we see the auxiliary between the topic and a question word:

(30) Warlpiri (Legate 2008: 34)

Kuturu-ju	ka=npa=nyanu	nyarrpara-wiyi	marda-rni?
nullanulla-TOP	PRS.IPFG=2SG.NOM=ANAPHOBJ	where-first	have-NPST

'Where do you have this nullanulla of yours?'

And in (31a-c) below, we see that a potentially interrogative word can be interpreted as either interrogative or indefinite if it appears right after an auxiliary, but only interrogative if before, while (31d) shows that if a potential question word appears further into the clause, after the auxiliary and the verb, it can only be interpreted as indefinite:

(31) Warlpiri

[21] And it is furthermore clear that it is in general not impossible for second-position clitics to be inserted into otherwise intact NPs; this is for example rather common with *-que* 'and' in Latin.

a. Kaji=ka=rna nyarrpara-kurra ya-ni.
NFACTC=PRS.IPFV=1SG where-ALL go-NPST
'I might go somewhere.'/'Where might I go?' (Legate 2008: 17)

b. Nyarrpara-kurra kaji=ka=rna ya-ni.
where-ALL NFACTC=PRS.IPFV=1SG go-NPST
*'I might go somewhere.'/'Where might I go?' (Legate 2008: 17)

c. Nyiya=rlangu kaji=ka=rlu nyina
what-for.example NFACTC=PRS.IPFV=3PL.OBJ be.NPST
wampana-piya-ju.
spectacled.hare.wallaby-like-TOP
'What ones for example might be like this spectacled hare wallaby?' (Legate 2008: 18)

d. Kaji=lpa=ngu wanti-yarla nyiya-rlangu
NFACTC=PST.IPFV=2SG.OBJ fall=IRR what-for.example
milpa-kurra.
eye-ALL
'If something fell into your eyes ...'
*'What might have fallen into your eyes?' (Legate 2008: 18)

I tentatively suggest the following analysis. The auxiliary appears in a fixed position, which Austin and Bresnan call I, although C would also work. The interrogative/indefinite pronouns are interpreted as interrogative if they appear 'external to S', indefinite if 'internal'. 'Internal to S' means that they appear inside the lowest S in a stack of S's to which things have been adjoined (and are therefore not themselves adjoined), 'external to S' outside of the lowest in such a stack, so either adjoined to a S or in some higher projection. In Warlpiri, there are two ways in which this can happen: they can appear in Spec of CP, giving rise to (31b-c) above, or adjoined to S, giving rise to the interrogative interpretation of (31a), where an internal position is also possible, giving rise to the indefinite interpretation. But in the case of (31d), the pronoun can only be internal to S, so only an indefinite interpretation is possible. However, for this to be the case we need a bit more, namely a restriction on adjunction to S, that it can only add a question-focus, which is easy to arrange with appropriate annotations.[22] Finally, in the case of (30), the Spec of CP position is occupied by the topic, so adjunction to S, and consequent position right after the auxiliary, is the only possibility for an interrogative reading (the only one that makes sense in the context).

[22] I do not know what happens if there is more than interrogative word; typologically, there are various possibilities.

The essential difference between the present LFG analysis and Legate's is that in the LFG analysis, the auxiliary appears in one position, and interrogatives in two, one one on either side of the auxiliary, while in Legate's, interrogative appear in one position, while auxiliaries can appear overtly in two. I am aware of no clear theory-independent empirical evidence distinguishing between these possibilities; they are each motivated by what appears to work out best given the resources of the theory.

3.3 Russian

Although Russian has sometimes been presented as nonconfigurational, King (1995) argues that it is configurational, but with provisions that make the word order considerably more flexible than in English. She provides first a Government Binding (GB) analysis, and then an LFG one, which leans heavily on the GB analysis for data and associated discussion.

She analyses Russian clauses as having CP, IP and VP layers, with two bar levels in each. The outer level of the VP introduces subjects, and there is one further layer over CP, which is available only in main clauses. This is for 'external topics', which have an initial XP, set off by a pause, with possibly an anaphoric pronoun later in the clause (King 1995: 202):

(32) Russian

a. Gleb, ja ego ne ljublju.
   Gleb, I him not like
   'Gleb, I don't like him.'

b. Opera, net drugogo vida musykal'nogo iskusstva, kotoryj privlekal by k sebe takoe vnimanie.
   Opera, not other type musical art, which attract would to itself such attention.
   'Opera, there is no other kind of musical art which would attract such attention to itself.'

She analyses this with an 'expression phrase' rule (Banfield 1982; Rudin 1985) as follows, outside the X-bar system (similarly in her GB analysis):

(33) $$\text{EP} \longrightarrow \underset{(\uparrow \text{E-TOP})=\downarrow}{\text{XP}} \quad \underset{\uparrow=\downarrow}{\text{CP}}$$

King (1995: 105) suggests that these external topics do not fall under the X-bar system in Government-Binding theory, and does not attempt to assimilate them to X-bar theory in LFG either, but I suggest that perhaps the rule in (33) could be replaced by 'Chomsky-adjunction' to CP,[23] with some kind of further restriction, perhaps essentially semantic, preventing them from occurring in embedded positions (King 1995: 106):

(34) CP $\longrightarrow$ XP CP
($\uparrow$ E-TOP)=$\downarrow$ $\uparrow$=$\downarrow$

Russian is not a full pro-drop language, so that an NP coreferential with an E-TOP is normally expressed.

In C we find the complementizer *što* and the question-marker *li*, most frequent in embedded questions, while in SPEC of CP we find question words: all question words in multiple wh-questions (King 1995: 215), unlike in English where only one appears. This is illustrated in the following example (King 1995: 216, $\uparrow$=$\downarrow$ annotations omitted):[24]

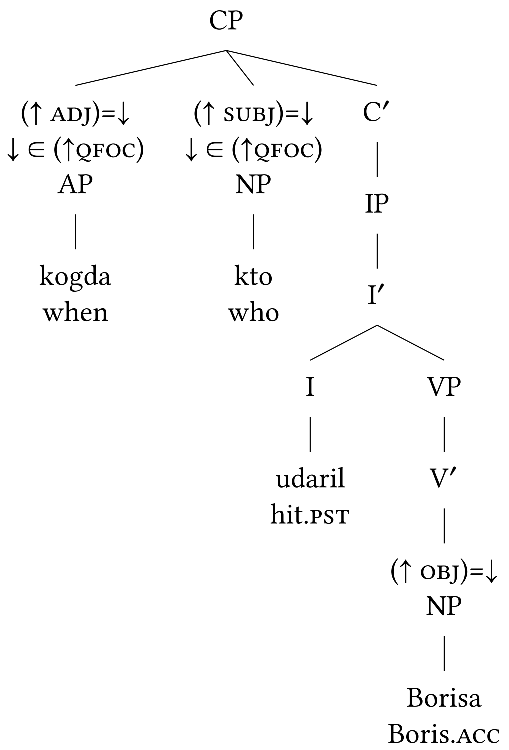
(35) Russian

a. kogda kto udaril Boris-a
when who(NOM) hit Boris-ACC
'Who hit Boris when?'

[23]Meaning that it has one CP node as sister, another as mother, with identical feature-composition.

[24]In the tree, the annotation over the first constituent should be $\downarrow \in (\uparrow$ ADJUNCTS), but in the structure I give the original.

b.

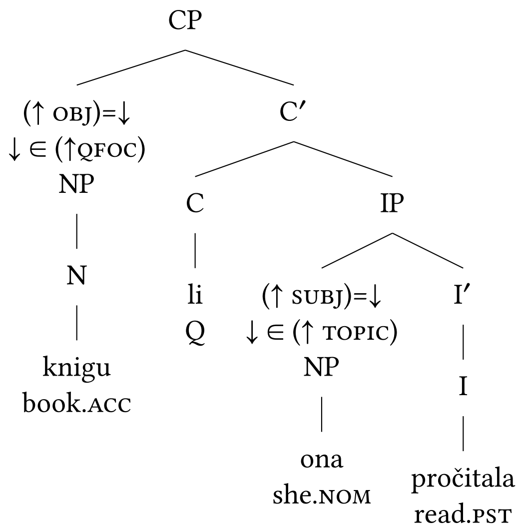


Yes-no questions can be marked either by intonation, or with the marker *li*, which appears in second position, after either an XP or the verb. An XP in front of *li* is interpreted as a focus, with the body of the question presupposed (King 1995: 236–237, King 1994), and items in SPEC of IP are Topics, as indicated by the annotation:

(36) Russian:

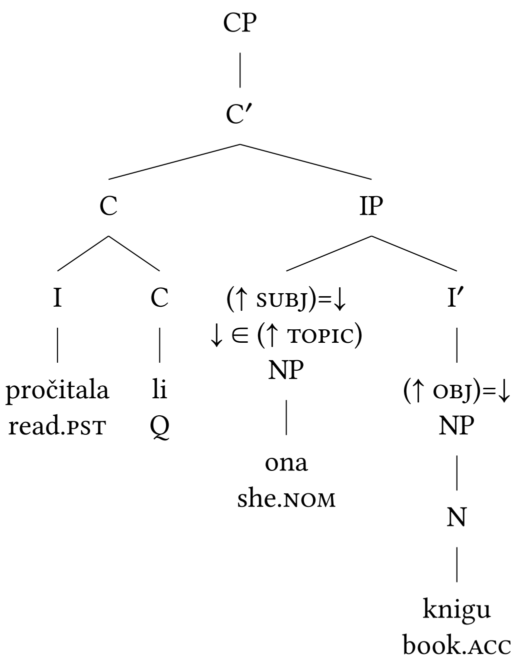
a. Knigu li ona pročitala?
book.ACC Q she.NOM read.PST
'Was it a book that she read?'

b.



If the yes-no question has no focus, then the verb appears before *li*, and King proposes that the verb is adjoined to C:

(37)



This solution avoids the issue of putting a nonmaximal projection in specifier position, and is independently motivated by the absence of the focus-presupposition articulation that appears when there is an XP in the specifier position. Formally, it could also be applied to Warlpiri, but without the additional motivation, insofar as is now known, and the problem of insertion of the auxiliary into combinations of verb and preverb would remain unsolved.

Moving down into IP, what we find here is ‘internal topics’ normally followed by contrastive foci, but with the possibility of some topics, especially pronouns, to come after the foci. Both contrastive foci and post-focus topics bear grammatical functions in their clause, and therefore do not cooccur with resumptive pronouns. The ordering phenomena are also connected to the issue of ‘emotive’ vs. ‘non-emotive’ sentences, an important topic in Russian syntax that does not get much discussion in King’s LFG analysis (although more in the GB one). After considering various proposals, including recursive right-branching, King’s final proposal is to use the ID/LP rule format (King 1995: 208), previously mentioned in footnote 3, where the ‘<’ symbol in (b) means that the item in front of < must appear before the one after, if both occur in the structure:

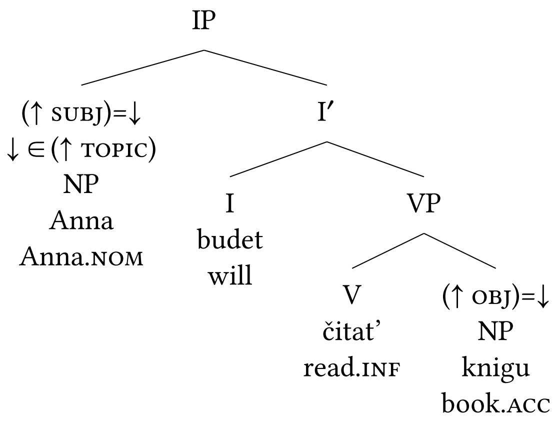
(38) a. $\mathrm{IP} \longrightarrow \begin{array}{cc} \mathrm{XP}^{*}, & \mathrm{I}' \\ (\uparrow \textsc{gf}){=}\downarrow & \uparrow{=}\downarrow \\ \downarrow \in (\uparrow \textsc{df}) & \end{array}$

b. TOP < CFOC, XP < I′

A constraint putting the I′ after the XPs has been added, and ‘DF’ (discourse function) is assumed to comprise ordinary topics (TOP but not ETOP), and contrastive foci (CFOC). This formulation allows for considerable flexibility in word order, even though the language is fundamentally configurational. The big difference between King’s IP rule for Russian and Bresnan’s for English is that in Russian, any GF that is also topic or focus can appear in Spec of IP, while in English, only subjects can. (King 1995: 133) notes that preverbal subjects tend to appear less markedly topicalized than other preverbal items, and discusses some possible reasons for this, including the tendency for arguments to be ordered consistently with the thematic hierarchy, which would put Agentive subjects, the most frequent kind, at the top, and therefore tending to be first.

In I itself appear finite verbs, either main verbs as in example (36), or the future auxiliary *budet* ‘be’, used to form imperfective futures, with the main verb appearing as head of VP:

(39)



This leads on to the structure of VP.

A somewhat unusual feature of King's analysis is that it introduces subjects not only as SPEC of IP, but also as SPEC of VP, along with other governed GFs that appear in V′ as usual, leading to this rule (King 1995: 209):[25]

(40)
a. $\mathrm{VP} \longrightarrow \underset{(\uparrow \textsc{subj})=\downarrow}{(\mathrm{XP})} \quad \underset{\uparrow=\downarrow}{\mathrm{V}'}$

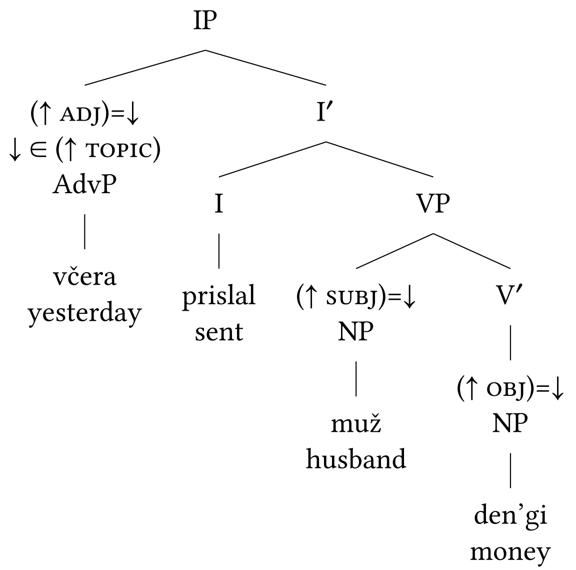
b. $\mathrm{V}' \longrightarrow \underset{\uparrow=\downarrow}{\mathrm{V}} \quad \underset{(\uparrow \textsc{gf})=\downarrow}{\mathrm{XP}^{*}}$

The evidence for this comes from various kinds of sentences where the subject is not also a topic, discussed more in King's GB analysis than in the LFG version.

One kind of example is 'thetic sentences', which assert that something happened, with no NP or other constituent singled out as the topic. The order in such examples is VSX*, as illustrated in the example below, as answer to the question 'what happened yesterday' (King 1995: 101):

[25] King omits the Kleene star on the complements in (b), presumably as a typographical error.

(41)



'Yesterday (my) husband(FOC) sent(FOC) (me) money(FOC).'

The placement of the subject in SPEC of VP rather than SPEC of IP causes it to be interpreted as Focus rather than Topic, leading to a thetic interpretation of the clausal material excluding *včera* 'yesterday'. We will see later in Section 5.3 that this analysis can be assimilated to that of other languages within the 2001 Synthesis presented in that section, by having the complement of I be (configurational) S rather than VP with internal subject.

3.4 German

German as analysed by Choi (1999)[26] resembles Warlpiri in arguably having only one functional projection,[27] but differs in a number of respects:

(42) a. Verbs can appear in the functional projection (and, often, must), but otherwise appear finally in VP.

b. There is no NP-splitting.

c. There is some evidence for an S node, although this is challenged by the later work of Berman (2003).

[26] A revised version of her 1996 Stanford dissertation of the same title.

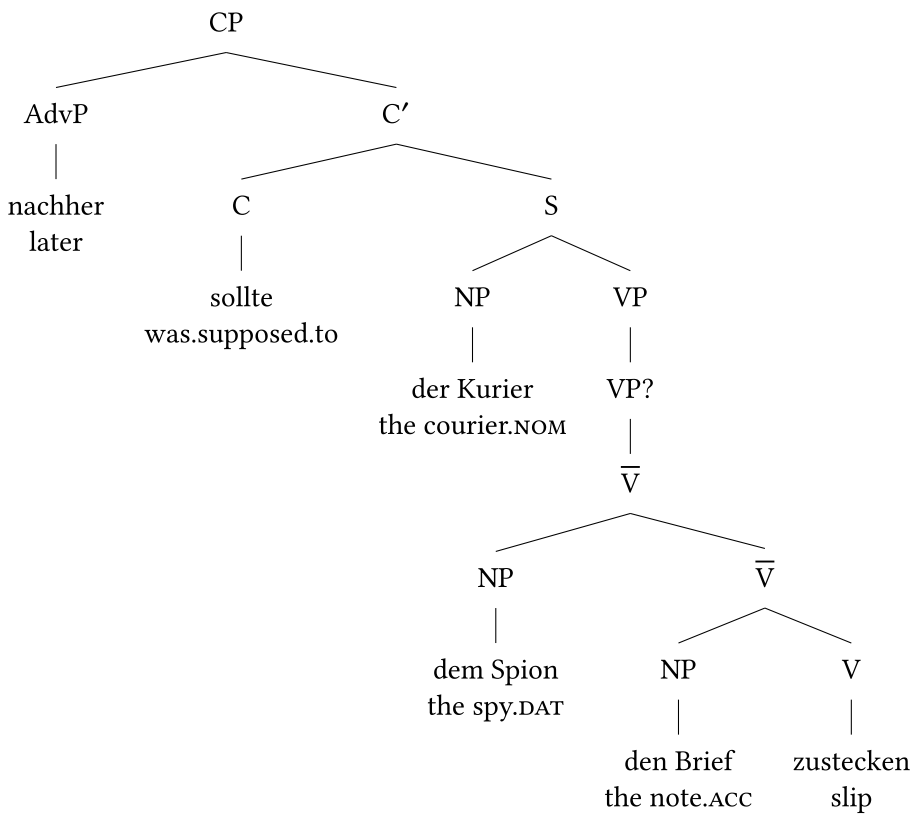
[27] A possibly confusing factor is that the German problem in Bresnan et al. (2016: 375-379) assumes that the auxiliary *haben* in final position is an occupant of I, but makes no argument for this analysis, which Choi explicitly rejects (Choi 1999: 33). Berman (2003: 31-32), citing Choi, also discusses the lack of evidence for I.

The functional projection in these works is labelled ‘C′’,[28] and houses complementizers in subordinate clauses, and the topmost (main) verb in matrix clauses. Otherwise, the verb appears clause-finally, and the so-called ‘auxiliaries’ are treated as main verbs taking VP complements. Full clauses consist of an S with a NP VP structure. A sample main clause with the auxiliary *sollte* is:

(43) German

a. Nachher sollte der Kurier dem Spion den Brief zustecken.
later was.supposed.to the courier.NOM the spy.DAT the note.ACC slip
‘Later, the courier was supposed to slip the letter to the spy.’

b.



[28]Presumably because it contains some traditional complementizers, such as *dass* ‘that’, although this choice is essentially arbitrary.

This structure is extrapolated from Choi (1999: 19, ex. (7a)) on the basis of later examples such as Choi (1999: 27, ex. (20)). The nested $\overline{\mathrm{V}}$ nodes are postulated to introduce the complements of V, a feature of the 2001 Synthesis which appears to be arbitrary, as we will discuss in section 4.

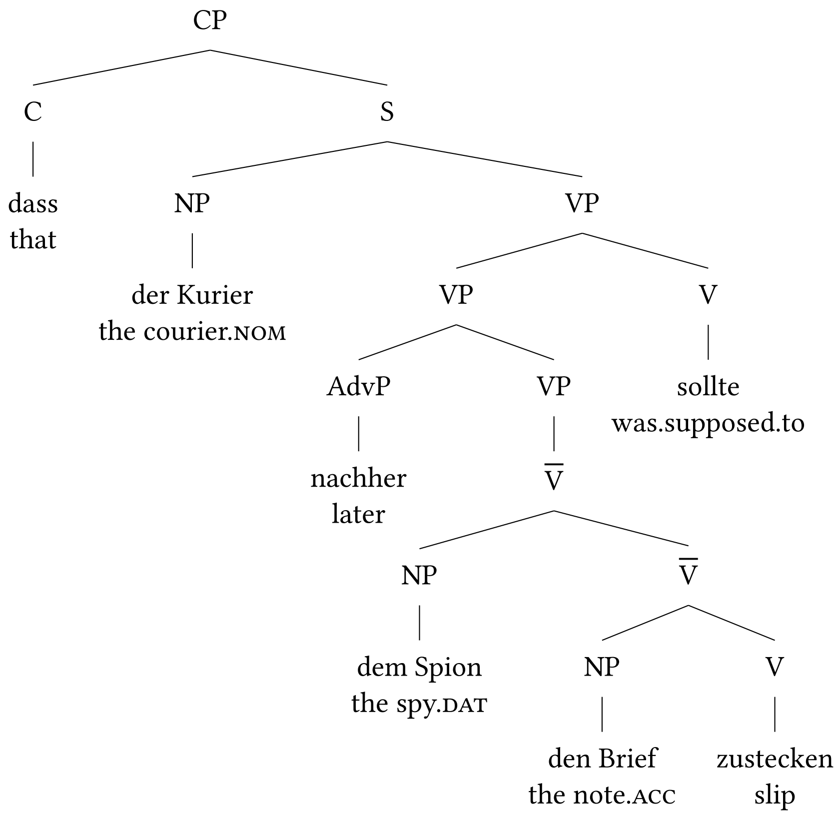
The question-marked VP solely dominated by another VP is motivated by the fact that the verbs traditionally called 'auxiliaries' in German (*sein* 'be', *haben* 'have', *werden* 'become', and the modals such as *sollen* 'should', among other meanings) appear syntactically to be ordinary verbs taking VP complements, and the question-marked VP could plausibly be taken to be the complement of the auxiliary *sollen* in C, whose VP is appearing immediately over the one with the question-marks. On the other hand, LFG for some time has been strongly oriented towards structure minimization principles, and the upper VP, which would be annoted '↑=↓' is not doing anything, and is therefore highly likely to be omitted, and indeed seems to be omitted by Choi in the somewhat later abbreviated structure (17) on p. 26.

A typical subordinate clause structure is:

(44) German (Choi 1999: 27, ex. 20)

a. dass der Kurier nachher dem Spion den Brief zustecken sollte
   that the courier.NOM later the spy.DAT the note.ACC slip was.supposed.to
   'that the courier was supposed to slip the note to the spy later'

b.



Here we see the complementizer occupying C, and the VP complement to the auxiliary *sollte*, at the end.

Although this analysis is in accord with the 2001 Synthesis, to be discussed in the next section, a later analysis by Berman (2003) rejects certain aspects of it, especially the arguments for putting the subject under S, as we will consider in section 5.4 below.

3.5 Korean

The languages we have seen so far have one or two functional projections over S or equivalent, but Choi (1999) presents Korean as having none: S expands to NP followed by VP. She finds no evidence for I or C, since the functions of these projections are expressed by formatives on the verb, removing the need for any phrase structure positions to house them, and no other kinds of evidence for their existence.

She provides three arguments for VP (Choi 1999: 43–47), of which I will give two. One of them is that there is a contrastive focus-marking particle *nun* which can be attached to either an object NP or the verb to make either the attached-to constituent or the entire VP the focus, but not the entire clause including the subject. Illustrating the two readings for marking on the object, we have:

(45) Korean
Mary-ka chayk-un ilk-nun-ta
Mary-NOM book-TOP read-PRS-DCL
'Mary reads nothing but books.'
(Mary does nothing but read books)
(Choi 1999: 45, example 52; note that Choi glosses the marker as 'TOP' even though its function here is described as contrastive focus.)

Another argument is phonological: syllable-initial obstruents become voiced after a vowel in a phonological phrase, and this happens between an object and the following adverb *caypalli*, but not a subject; here, the segments voiced for this reason are italicized:[29]

(46) Korean
a. Cwuni-ga kong-ul *j*aypalli *j*a*b*-a.
Cwuni-NOM ball-ACC quickly catch-PRS
'Cwuni catches balls quickly.'
b. Kong-ul Cwuni-ga caypalli *j*a*n*-a.
ball-ACC Cwuni-NOM quickly catch-PRS
'Balls Cwuni catches quickly.'

This treatment contrasts with that of Sells (1994), who proposes that Korean has an 'inner VP' (similar to the combinations of verb and preverb in Warlpiri) which can be plausibly analysed as a V^0 with adjoined non-projecting words, but no S vs. VP distinction. Instead, following Fukui (1986), all arguments are attached by phrase structure rules expanding VP to XP and VP, recursively. Sells's argument for the VP seems convincing, but not those for the binary branching structures for the arguments (Sells 1994: 353, fn. 2). Later, in Bresnan (2001), a branching VP like that of Sells was accepted, but with no serious attempt to show that it was superior to the more traditional flat VP assumed by Choi.

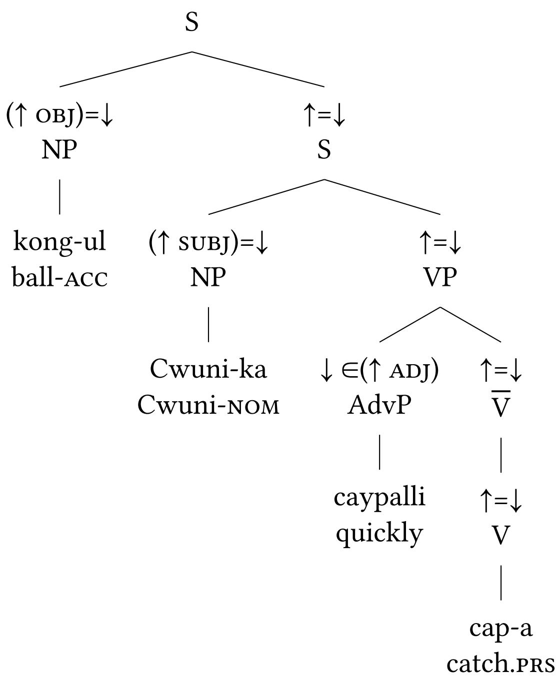
Another important characteristic of Choi's phrase structures is the absence of verbal functional projections. She considers an analysis in which tense and

[29]Choi uses underlining to indicate the non-phonemic voicing, rather than different segmental phonetic symbols.

the declarative markers are treated as inhabitants of I and C, respectively, and morphologically fused with the verb, but rejects such analyses on the basis of violating the LFG principle that inflected words are inserted under single terminal nodes in the c-structure. One could propose that these projections do exist, but are fused with the verb via lexical sharing, but then there would be the problem of the nonexistence of any evidence for the syntactic autonomy of the two components, of the kind that the mechanism of lexical sharing was devised in order to explain.

As a consequence of the absence of C and I, we cannot use SPEC positions of these projections to house preposed items to provide a treatment of scrambling as found in (46b). Choi does not in fact present any c-structures for scrambled sentences in Korean, but states (Choi 1999: 9) that this is to be the structure for scrambling, and illustrates them for German (Choi 1999: 127, ex. (17a)) with left-adjunction to S. This illustrates the principle that LFG does not propose a functional projection if there is no material that can occupy the head of that projection (in any structure; it is allowed for the head position to be unoccupied in some structures).

Furthermore, since there is no IP, we can't use SPEC of IP to house the subject, so Choi proposes S expanding to NP and VP. Therefore, the structure of (46b) is:

(47)


Caypalli is not one of the adverbs listed by Sells as restricted to immediately preverbal position, so we introduce it here as a daughter of VP rather than adjoined to V. The general question of VP versus $\overline{\text{V}}$ is a difficult one, which the new phrase structure theory of Lowe & Lovestrand (2020) might allow us to solve by eliminating, but is beyond the scope of this chapter.

4 The 2001 synthesis

And now we turn to the system presented in Bresnan (2001), foreshadowed at various points in the preceeding discussion, and largely unchanged in its successor (Bresnan et al. 2016), which will be the source of our page reference citations. We have already introduced many elements of this proposal, so it is time to develop it more systematically. The basic ingredients, some of which are present in all of the intermediate stage analyses, are:

(48)
a. the 3-level X-bar theory of Chomsky (1970), with one lexical and two phrasal levels, with the option for a language to have only two levels (one lexical and one phrasal), as in Warlpiri;
b. the modified system of category features (for nouns, verbs, etc.) from Jackendoff (1977) and Bresnan (1982a);
c. functional (extended) projections in the version of Grimshaw (2000), in which the extended projections share category features with their complements, which facilitates keeping the number of phrasal projections to 2. These are normally called I (as in Warlpiri), or C (as in German);
d. the existence of a category S, outside of the X-bar system, which can either be nonconfigurational, as proposed for Tagalog and Warlpiri, or configurational as will be proposed for Welsh and some minor constructions in English;
e. principles of structure-function mapping that limit what kinds of grammatical functions can be introduced in what positions;
f. the claim that phrases in the X-bar system are ‘endocentric’, while S is ‘exocentric’.

A new feature of the 2016 version relative to the 2001 version is the ‘nonprojecting words’ of Toivonen (2001), discussed in Belyaev forthcoming [this volume], which have a category feature but are adjoined to another phrase without projecting anything themselves. Another feature of both versions that is not widespread in X-bar theory is the treatment of multiple complements. Bresnan et al. (2016: 127) observe that there are two possible treatments of complements, (49a) below with a nested structure, (49b) with a flat one:

(49) a. X′ ⟶ X′, YP
b. X′ ⟶ X, YP*

(49b) is what has been assumed in most of the literature, (49a) what Bresnan (2001) and Bresnan et al. (2016) propose, on the basis of supporting a 'flexible definition of an endocentric complement' (Bresnan 2001: 118, Bresnan et al. 2016: 123).[30] Option (a) constitutes the choice that Choi makes for German in example (43), and works well when the complements can be ordered freely, but it is not clear to me how it is to account for the ordering restrictions in double object constructions, where the OBJ tends to precede an OBJ_{θ}; the relevant restrictions can be easily stated, either with conventional phrase structure rules as in most early LFG, or with the ID/LP format (briefly mentioned in footnote 3). I suggest this is an issue best left for future investigation.

A further theme that interacts with the X-bar principles is a tendency to reduce the complexity of c-structure to a minimum. Two of the more important conditions are:

(50) a. A c-structure position is not postulated unless there is a class of items that can fill it.
b. In any specific structure, all nodes are optional.

(50a) prevents us from postulating a functional projection such as 'T' for topic, or 'E' for 'evidential', unless we can find an class of items that plausibly appear in their head positions, while (50b) allows us to leave out items in specific cases, as will be discussed below.

The optionality of c-structure positions is an aspect of an important more general principle, Economy of Expression (Bresnan et al. 2016: 90):

(51) **Economy of Expression**:
All syntactic phrase structure nodes are optional and are not used unless required by independent principles (completeness, coherence, semantic expressivity).

A consequence of this principle is that the traditional principle of 'S as the 'initial symbol' in a phrase-structure derivation of a sentence' is abandoned: a 'sentence' can phrase-structurally be an S, an IP, or a CP, depending on the language

[30] Bresnan et al. (2016), footnote 50 refers to footnote 41, but that appears to be be irrelevant; the relevant definition appears on the page following footnote 41.

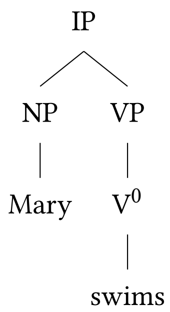
and details of the particular sentence. I suggest that this involves a shift from a traditional ‘syntactic’ notion of sentence-hood to a more ‘semantic one’, since glue semantics (Asudeh forthcoming [this volume]) can connect these multiple superficial syntactic structures to a single semantic type. Further reductions in the complexity of c-structure are achieved by the reworking of X-bar theory developed in Lovestrand & Lowe (2017), but these have not yet been applied to a substantial typological variety of clause structures so as to produce results with differently organized overall structure, and therefore will not be discussed here.

The c-structure principles interact with a set of structure-function mapping principles, which constrain the relationship between the c-structures and the f-structures. The principles (Bresnan et al. 2016: 105-109) assert that:

(52)
a. C-structure heads are f-structure heads.
b. Complements of functional categories are f-structure co-heads.
c. Specifiers of functional categories are the grammaticalized discourse functions, such as SUBJ, FOC, TOP.
d. Complements of grammatical categories are the non-discourse argument functions, such as OBJ, OBJ$_\theta$ (but not SUBJ).
e. Constituents adjoined to phrasal constituents are optionally nonargument functions, either adjuncts or nonargument discourse functions.

This principle does not apply to S, whose daughters can bear any grammatical function.

Some simple effects of these principles apply in the structure for ‘Mary swims’:[31]

(53)


(Bresnan et al. 2016: 120)

[31] The Bresnan (2001) version has S instead of IP, but this is rejected due to the lack of independent evidence for S in English main (indeed, finite and most nonfinite) clauses in English.

The NP 'Mary' can be a subject because it is external to the VP in SPEC of IP position, while the other nodes will bear the ↑ = ↓ equation and so correspond to the single f-structure that the NP's f-structure correspondent is SUBJ of. Turning to Economy, unfilled heads and intermediate level nodes that dominate nothing but their head are eliminated, and the IP, VP, and V nodes will all correspond to the same f-structure.

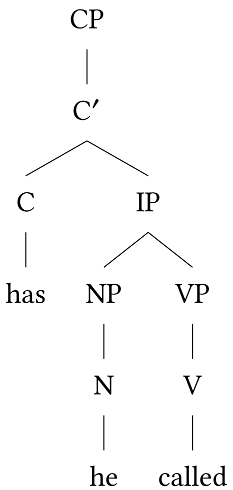
The CP level appears in complement clauses, where the complementizer is head of C, and in certain other structures such as questions, where we get a CP level with an auxiliary verb filling C, a kind of analysis that is strengthened by the fact that an auxiliary can replace the complementizer *if* in a somewhat archaic/ solemn variant of conditional clauses:

(54) Has he called?

(55) a. If he had called, I would have answered.
b. Had he called, I would have answered.

For (54), given the preceeding, the plausible structure is:

(56)



The IP provides the location of the subject, which has no alternative locations in English finite clauses.[32] The conditional clause in (55b) has the same structure, with the preposed auxiliary replacing the overt complementizer that appears in (55a).

[32] But, as we will see shortly, certain nonfinite 'small clauses' arguably use S rather than IP. The exact nature of the connection between 'finiteness' and IP deserves further investigation.

In the systematization of Bresnan et al. (2016: 103), the functional projections are distinguished from the lexical ones by having values 1 or 2 for a feature F, whose value is unspecified for lexical projections. This implies that the choice of C or I in the 1 level languages such as Warlpiri and German is arbitrary, with a further consequence that any tendency in two-level languages to express some things in C and others in I is probably functional in origin. It is of course also possible for there to be no verbal functional projections at all, as argued by Choi for Korean, and is plausibly also the case for Malayalam and Japanese.

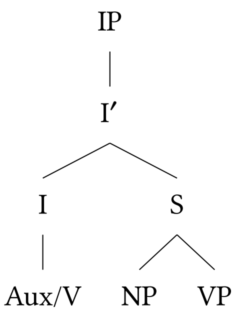
5 Applications to languages

In this section we consider the application of the 2001 synthesis to various languages, starting with Welsh, and then reviewing some of the previous ones which call for comment. Malayalam and Korean fit without further discussion, and so are omitted here.

5.1 Welsh

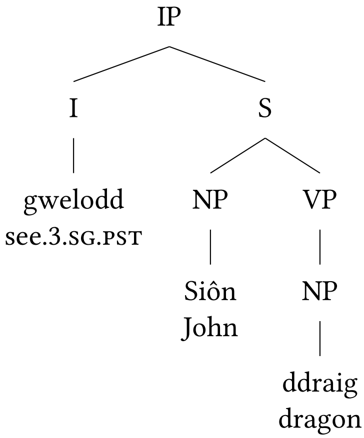
The basic structure for a main clause is:

(57)

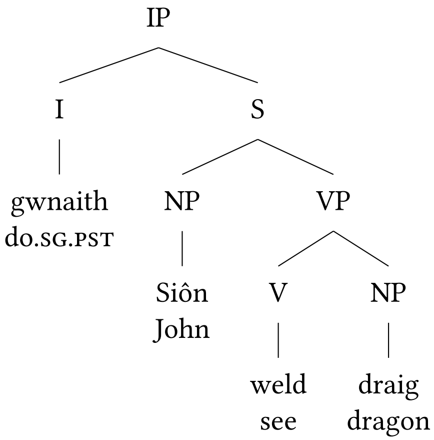


If there is no auxiliary, the finite verb appears at the front of the sentence, in the I position, as shown in (58a) below, very similarly to King's (1995) analysis of Russian. But if there is an auxiliary, the auxiliary appears in I, the verb initially in V of VP, again similarly to Russian, as shown in (58b), but with the VP under S dominating the subject, rather than a two-level VP (Bresnan et al. 2016: 131-133):

(58) a.



b.



One issue here is how a VP can appear sitting over only an NP in (58a). This goes against the idea that an endocentric category needs to have a head (while the presence of a VP is motivated by the range of things that can appear in the position after the subject, where overt VPs sometimes appear over the same material.

This problem is averted by the 'Extended Head Principle' (Bresnan et al. 2016: 135-137), which in effect says that a phrase can have a 'displaced head', as long as that head appears within a higher phrase having the same f-structure correspondent. The definition of 'extended head' is:

(59) Given a c-structure containing nodes $\mathcal{N}$ and C and a c- to f-structure correspondence ϕ, $\mathcal{N}$ is an extended head of C if and only if $\mathcal{N}$ is the minimal node in $\phi^{-1}(\phi(C))$ that c-commands C without dominating C. (Bresnan et al. 2016: 136)

and the principle is:

(60) Every (phrasal)[33] lexical category has an extended head

Although we have discussed only I, Welsh also has a functional projection C, containing complementizers as discussed by for example Roberts (2005). Therefore it is a 2-level language like English, but differs from English in that the complement of I is S rather than VP. This is required because in Welsh, the position of the subject is after the auxiliary rather than before it. 'I+S' languages, such as for example Tagalog, also often have the property that S can have PP, AP, and NP as well as VP as the predicate phrase, but this does not appear to be the case for Welsh, since it uses a copula in sentences where these play the semantic role of predicate.[34] Welsh also differs from English (and is similar to most other Germanic languages) in that all verbs can appear in I, rather than only a restricted class of 'auxilaries'.

5.2 More English

As observed above, the fact that English puts the subject in front of I rather than after it indicates that it has VP rather than S as complement of I, an analysis corroborated by the fact that a verb is obligatory in finite clauses (recalling that IP shares category features with its complement, so excludes a non-verbal complement). But nevertheless, as observed by Chung & McCloskey (1987), English does arguably have nonfinite clauses where S expands to NP subject and a predicate phrase, which can be NP, AP, VP or PP. These so-called 'small clauses' are used in English to express a combination of incredulity and often dismay (Akmajian 1984):

(61) a. What?? Him an alcoholic??
  b. What?? Her sick with the flu??
  c. What?? Him run(ning) a company??
  d. What?? That guy in the White House??

Along with nonfiniteness comes accusative case on the subject and no agreement with any verbal element. This is evidence that in English, nominative case on NPs is a marker of finiteness on the clause, somewhat in the manner of the

[33]This qualification is absent from the original, but seems to me to make the principle work properly.

[34]Bresnan et al. (2016: 130) suggest that Welsh has these as possible predicates under S, but no examples are provided, and Welsh appears in fact to use a copula. See Borsley (2019) for a recent discussion of Welsh copular clauses in the framework of HPSG.

proposals of Pesetsky & Torrego (2007) within the Minimalist program, although an LFG analysis would have quite a different implementation, similar to the treatment of 'modal case' in Tangkic languages, as presented in somewhat different forms by Andrews (1996) and Nordlinger (1998). In these languages, nominative forms would have an equation specifying an appropriate tense feature value for whatever they were subject of.

Another plausible case for S would be gerundive nominalizations with accusative (rather than genitive) subjects, as analysed by Schachter (1976):

(62) John giving/*give an invited talk might be a good idea

A potential issue with having NP expanding to S here is that the predicate phrase of the S is restricted to being a VP whose verb is marked with *-ing*. But this can be accomodated by including an appropriate constraining equation in the c-structure rule:

(63) NP $\longrightarrow$ $\begin{array}{c}\text{S}\\ \uparrow=\downarrow\\ (\downarrow \text{VFORM}) =_c \text{ING}\end{array}$

The constraining equation will require the predicate of the S to be a verb, as well as a verb marked with *-ing*. Taking this analysis further would require entering the realm of 'mixed category constructions', beyond the scope of this chapter. But there is clearly a question of what causes S to have a rather limited distribution in English, as opposed to other languages such as Welsh or Tagalog.

5.3 Russian revised

Moving on to Russian, King's analysis, discussed in subsection 3.3, diverges from the previous analyses we have discussed in this section in using a two-level VP, with the top level introducing a subject. This difference can be easily eliminated by replacing the upper level VP with S, but is there any serious motivation for doing that?

The structure-function mapping principles listed in (52) are not entirely clear on this: (52c) says that specifiers of functional categories are the grammaticalized discourse functions, which suggests no, but grammaticalized discourse functions can also be adjoined (52e), which suggests possibly yes. A general point that suggests that the S analysis is correct is that the question of what should appear in the specifier of VP in many languages, such as English, has always been rather controversial. Fukui (1986) proposes that only functional, not lexical, categories, have specifiers, and this appears to be consistent with what we have seen here.

On the other hand, specifier is at least a somewhat plausible place to locate quantity and degree modifiers of nouns and adjectives, but there could be a weaker position that specifiers of lexical categories do not supply arguments to those categories, but serve different functions when they exist at all.

If we accept this reanalysis, a natural concomitant is to place postposed focussed subjects and certain other NPs under S, following the VP:

(64) S $\longrightarrow$ (XP) VP XP*
(↑ SUBJ)=↓ ↑=↓ (↑ GF)=↓
(↑ FOC)=↓

King (1995: 210) put these under V′ in order to treat examples such as:

(65) Russian (King 1995: 209)
Kupila plat'e Inna.
bought dress Inna
'Inna.FOC bought a dress.'

But this actually does run against the structure-function mapping principle (52d). Using S in this way gives us a version of King's analysis that is clearly within the framework of the 2001 Synthesis.

5.4 German again

German, however, presents a problem. Choi's treatment was within the Synthesis, in spite of originating in 1996, but Berman (2003) eroded an important aspect of it, that subjects appear under S. In particular, following Haider (1990, 1995), she concluded that various kinds of presumed subjects were included in VP-preposing, including the rather hard to dismiss unergatives, which cannot be construed as nominative objects, as is possible for some of the other examples:

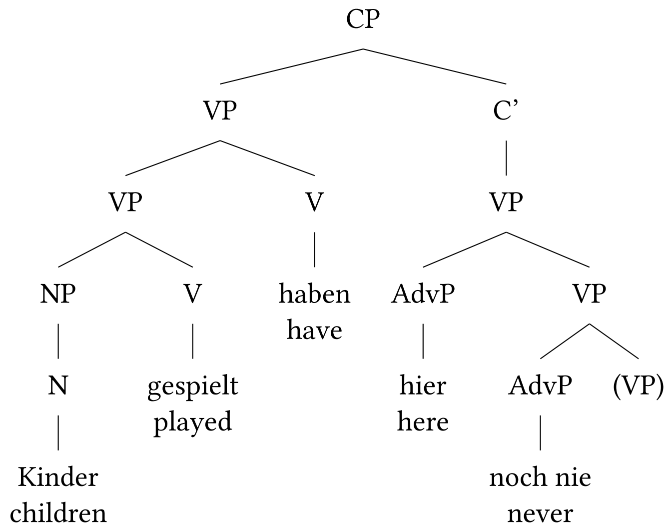
(66) German (Berman 2003: 36)
Kinder gespielt haben hier noch nie.
children played have here still never
Children have never played here.

Her solution was to introduce all arguments, including subjects, under VP in nested VPs, her VP not being clearly distinct from Choi's $\overline{\mathrm{V}}$, with a complex verb at the bottom, as discussed in a considerable amount of literature in different frameworks.[35]

Unfortunately, she did not provide an actual tree for this example ((28b) in the text, nor for the similar (28a)), but I suggest:

[35]For example Wurmbrand (2017) in Minimalism, Zaenen & Kaplan (1995) in LFG.

(67) German



Where the parenthesized (VP) is an informal notation for the fact that the head of the unparenthesized VP above it is supplied by the Extended Head Principle from the Spec of CP position.

This would seem to call for revision of Choi's structures for German such as (44a), but, on the other hand, the VP-internal subjects appear to be restricted; for example, they must be indefinite. So it is not excluded that there is both an S where most subjects reside, as proposed by Choi, and subjects appearing in the VP, as proposed by Haider. Alternate word orders are also produced by 'scrambling', with strong effects on information structure, as extensively discussed by Choi.

5.5 Final remarks

We can sum up the discussion in the previous subsections into a set of rules for when to posit S as opposed to VP as complement of I and in some other environments.

(68) a. If SPEC of IP has a subject position, and subjects appear there, rather than in the complement of I, then the complement of I is VP or possibly other maximal lexical projections.

b. If subjects appear in the complement rather than in the SPEC of IP, then the complement of I is S.

c. This can happen in two ways: either S is nonconfigurational, introducing predicates and arguments in variable order (with the possibility of the predicates being ordered at the end or beginning of S), or S is configurational, dominating NP and a predicate phrase (maximal projection).

In this context, we interpret 'subject' as the core grammatical function (Bresnan et al. 2016: 97) that is also a discourse (topic) grammatical function (Bresnan et al. 2016: 100). We also of course assume S when there is no evidence for I, as in Korean.

The organization of functional projections, on the other hand, is to be ascertained by the arrangements of elements marked by verbal features with respect to other members of the clause, with a general tendency for there to be more intervening items when the projections precede the main verb position than when they follow, leading to a tendency for verb-final languages to tend to appear to lack verbal functional projections.

It is clear that many of the languages we have considered are due for careful reanalysis, due to unresolved issues and discrepancies between earlier and later versions of LFG, and also taking into account the new phrase structure theory of Lowe & Lovestrand (2020). Of particular interest would be the nature of clitic placement in Tagalog, and the issue of flat versus nested structure in nonconfigurational languages such as Tagalog and German, where the order of arguments is free, but the predicate is fixed at one end or the other.

A major study of clause structure that we have not tried to work through is Sells (2001) on Swedish, for the reason that this makes heavy use of Optimality Theory in ways that have not become mainstream in LFG. However, word order and clause structure in Scandinavian languages is an extremely complex and interesting area that deserves further investigation.

I will conclude with a speculation about the nature of nonconfigurationality, which is that nonconfigurational S might be what ensues diachronically when a language makes such extensive use of discourse-conditioned preposing that the syntax learner simply gives up and returns, at the level of c-structure, what is essentially a list of fragments, similarly to the XLE system (Forst & King forthcoming [this volume]) when it can't find a parse. However, in such cases, the f-structure construction system can bring additional resources to bear and produce an integrated result at f-structure and other levels. The idea of nonconfigurational S as a kind of partial failure of c-structure parsing would be consistent with the revision of the theory of Lowe & Lovestrand (2020) proposed in Lovestrand (2022), whereby nonconfigurational S has no category feature value.

Acknowledgements

I would like to acknowledge the help of Jane Simpson and three anonymous reviewers.

References

Akmajian, Adrian. 1984. Sentence types and the form-function fit. *Natural Language & Linguistic Theory* 2(1). 1–23.

Akmajian, Adrian, Susan Steele & Thomas Wasow. 1979. The category AUX in Universal Grammar. *Linguistic Inquiry* 10. 1–64.

Andrews, Avery D. 1996. Semantic case-stacking and inside-out unification. *Australian Journal of Linguistics* 16(1). 1–55. DOI: 10.1080/07268609608599530.

Andrews, Avery D. Forthcoming(a). Clause structure and configurationality. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Andrews, Avery D. Forthcoming(b). Complex predicates. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Asudeh, Ash. Forthcoming. Glue Semantics. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Austin, Peter K. & Joan Bresnan. 1996. Non-configurationality in Australian aboriginal languages. *Natural Language & Linguistic Theory* 14(2). 215–268. DOI: 10.1007/bf00133684.

Banfield, Ann. 1982. *Unspeakable sentences: Narration and representation in the language of fiction*. Boston: Routledge & Kegan Paul.

Belyaev, Oleg. Forthcoming. Core concepts of LFG. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Berman, Judith. 2003. *Clausal syntax of German*. Stanford: CSLI Publications.

Borsley, Robert D. 2019. The complexities of the Welsh copula. In Stefan Müller & Petya Osanova (eds.), *Proceedings of the 26th international conference on Head-Driven Phrase Structure Grammar*, 5–25. Stanford: CSLI Publications.

Bresnan, Joan. 1982a. Control and complementation. *Linguistic Inquiry* 13(3). Reprinted in Bresnan & Kaplan 1982, 343–434.

Bresnan, Joan (ed.). 1982b. *The mental representation of grammatical relations*. Cambridge, MA: The MIT Press.

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-Functional Syntax*. 2nd edn. (Blackwell Textbooks in Linguistics 16). Malden, MA: Wiley-Blackwell.

Bresnan, Joan & Ronald M. Kaplan. 1982. Introduction: Grammars as mental representations of language. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, xvii–lii. Cambridge, MA: The MIT Press.

Choi, Hye-Won. 1999. *Optimizing structure in context: Scrambling and information structure*. Stanford: CSLI Publications.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.

Chomsky, Noam. 1970. Remarks on nominalization. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English transformational grammar*, 184–221. Waltham, MA: Ginn.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris Publications. DOI: 10.1515/9783110884166.

Chung, Sandra & James McCloskey. 1987. Government, barriers and small clauses in Modern Irish. *Natural Language & Linguistic Theory* 18. 173–237.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III & Annie Zaenen (eds.). 1995. *Formal issues in Lexical-Functional Grammar*. Stanford: CSLI Publications.

Devine, A. M. & Laurence D. Stephens. 2000. *Discontinuous syntax: Hyperbaton in Ancient Greek*. Oxford: Oxford University Press.

Falk, Yehuda N. 1984. The English auxiliary system: A lexical-functional analysis. *Language* 60(3). 483–509. DOI: 10.2307/413988.

Foley, William A. & Robert Van Valin Jr. 1984. *Functional syntax and universal grammar*. Cambridge, UK: Cambridge University Press.

Forst, Martin & Tracy Holloway King. Forthcoming. Computational implementations and applications. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Fukui, Naoki. 1986. *A theory of category projection and its applications*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).

Gazdar, Gerald. 1982. Phrase structure grammar. In Pauline Jacobson & Geoffrey K. Pullum (eds.), *The nature of syntactic representation*, 131–186. Dordrecht: Reidel. DOI: 10.1007/978-94-009-7707-5_5.

Gerassimova, Veronica A. & Peter Sells. 2008. Long distance dependencies in Tagalog: The case for raising. In Charles B. Chang & Hannah J. Haynie (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 190–198. Somerville, MA: Cascadilla Press.

Grimshaw, Jane. 2000. Locality and extended projection. In Peter Coopmans, Martin Everaert & Jane Grimshaw (eds.), *Lexical specification and insertion*, 115–133. Amsterdam/Philadelphia: John Benjamins. DOI: 10.1075/cilt.197.07gri.

Haider, Hubert. 1990. PRO-BLEME. In Gisbert Fanslow & Sascha W. Felix (eds.), *Structuren und Merkmale syntaktischer Kategorien*, 121–143. Tübingen: Narr.

Haider, Hubert. 1995. *Studies in phrase structure and economy.*

Hale, Ken. 1981. *On the position of Warlpiri in a typology of the base.* Bloomington, IN: Indiana University Linguistics Club. Distributed by Indiana University Linguistics Club, Bloomington.

Hale, Kenneth. 1973. Person marking in Warlpiri. In Stephen R. Anderson & Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 308–344. New York: Holt, Rein-hart & Winston.

Halpern, Aaron. 1995. *On the placement and morphology of clitics.* Stanford: CSLI Publications.

Ishikawa, Akira. 1985. *Complex predicates and lexical operations in Japanese.* Stanford: Stanford University. (Doctoral dissertation).

Jackendoff, Ray. 1977. *$\bar{X}$ syntax: A study of phrase structure* (Linguistic Inquiry Monographs 2). Cambridge, MA: The MIT Press.

Kaufman, Daniel. 2010. *The morphosyntax of Tagalog clitics: A typologically driven approach.* Ithaca: Cornell University. (Doctoral dissertation).

Kaufman, Daniel & Victoria Chen. 2017. *Clausal constituency and historical Austronesian morphosyntax.* Handout for talk at 24th meeting of the Austronesian Formal Linguistics Association, University of Washington. https://bahasawan.com/wp-content/uploads/2020/01/Clausal-constituency-and-historical-AN-morphosyntax-H.pdf.

Keenan, Edward L. 1976. Towards a universal definition of subject. In Charles N. Li (ed.), *Subject and topic*, 303–333. New York: Academic Press.

King, Tracy Holloway. 1994. Focus in Russian yes-no questions. *Journal of Slavic Linguistics* 2. 92–120.

King, Tracy Holloway. 1995. *Configuring topic and focus in Russian.* Stanford: CSLI Publications.

Kroeger, Paul R. 1993. *Phrase structure and grammatical relations in Tagalog.* Stanford: CSLI Publications.

Kuhn, Jonas. Forthcoming. LFG, Optimality Theory and learnability of languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar.* Berlin: Language Science Press.

Laughren, Mary. 1989. The configurationality parameter and Warlpiri. In László Marácz & Pieter Muysken (eds.), 319–354. Republished in 2019 by De Gruyter Mouton. Dordrecht: Foris Publications.

Laughren, Mary. 2002. Syntactic constraints in a "free word order" language. In Mengistu Amberber & Peter Collins (eds.), *Language universals and variation*, 83–130. New York: Ablex Publishing Corp.

Legate, Julie Anne. 2008. Warlpiri and the theory of second position clitics. *Natural Language & Linguistic Theory* 26. 3–60. DOI: 10.1007/s11049-007-9030-0.

Liu, (Adlay) Kun Long. 2017. *Syntactic interactions with information structure in Squliq Atayal*. Canberra: Australian National University. (Doctoral dissertation). DOI: 10.25911/5d7632016f800.

Louagie, Dana & Jean-Christophe Verstraete. 2016. Noun phrase constituency in Australian languages: A typological study. *Linguistic Typology* 20(1). 25–80. DOI: 10.1515/lingty-2016-0002.

Lovestrand, Joseph. 2022. *Barayin morphosyntax*. Oxford: Oxford University Press.

Lovestrand, Joseph & John J. Lowe. 2017. Minimal c-structure: Rethinking projection in phrase structure. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '17 conference*, 285–305. Stanford: CSLI Publications.

Lowe, John J. & Joseph Lovestrand. 2020. Minimal phrase structure: A new formalized theory of phrase structure. *Journal of Language Modelling* 8(1). 1–51. DOI: 10.15398/jlm.v8i1.247.

Manning, Christopher D. 1996. *Ergativity: Argument structure and grammatical relations* (Dissertations in Linguistics). Stanford: CSLI Publications.

Matsumoto, Yo. 1996. *Complex predicates in Japanese: A syntactic and semantic study of the notion 'word'*. Stanford: CSLI Publications.

Mohanan, K. P. 1982. Grammatical relations and clause structure in Malayalam. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 504–589. Cambridge, MA: The MIT Press.

Nash, David G. 1986. *Topics in Warlpiri grammar*. Published version of 1980 MIT Dissertation. New York: Garland.

Nordlinger, Rachel. 1998. *Constructive case: Evidence from Australian languages*. Stanford: CSLI Publications.

Nordlinger, Rachel. Forthcoming. LFG and Australian languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Pesetsky, David & Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In Simin Karimi, Vida Samiian & Wendy K. Wilkins (eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation*. Amsterdam: John Benjamins. DOI: 10.1075/la.101.14pes.

Roberts, Ian G. 2005. *Principles and parameters in a VSO language: A case study in Welsh*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780195168211.001.0001.

Rudin, Catherine. 1985. *Aspects of Bulgarian syntax: Complementizers and WH constructions*. Columbus, OH: Slavica.

Schachter, Paul. 1976. The subject in Philippine languages: Actor, topic, actor-topic, or none of the above? In Charles N. Li (ed.), *Subject and topic*, 491–518. New York: Academic Press.

Schachter, Paul. 1977. Reference-related and role-related properties of subjects. In Peter Cole & Jerrold M. Sadock (eds.), *Grammatical relations* (Syntax and Semantics 8), 279–306. New York: Academic Press.

Schachter, Paul & Fe T. Otanes. 1972. *Tagalog reference grammar*. University of California Press. DOI: 10.1525/9780520321205.

Schultze-Berndt, Eva. 2022. When subjects frame the clause: Discontinuous noun phrases as an iconic strategy for marking thetic constructions. *Linguistics* 60(3). 865–898. DOI: 10.1515/ling-2020-0138.

Schultze-Berndt, Eva & Candide Simard. 2012. Constraints on noun phrase discontinuity in an Australian language: The role of prosody and information structure. *Linguistics* 50. 1015–1058. DOI: 10.1515/ling-2012-0032.

Sells, Peter. 1994. Sub-phrasal syntax in Korean. *Language Research* 30(2). 351–386.

Sells, Peter (ed.). 2001. *Formal and empirical issues in Optimality Theoretic syntax*. Stanford: CSLI Publications.

Simpson, Jane. 1983. *Aspects of Warlpiri morphology and syntax*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation). https://dspace.mit.edu/handle/1721.1/15468. Published in revised form as Simpson (1991).

Simpson, Jane. 1991. *Warlpiri morpho-syntax: A lexicalist approach*. Dordrecht: Kluwer Academic Publishers.

Simpson, Jane & Joan Bresnan. 1983. Control and obviation in Warlpiri. *Natural Language & Linguistic Theory* 1(1). 49–64. DOI: 10.1007/bf00210375.

Sityar, Emily M. 1989. *Pronominal clitics in Tagalog*. Los Angeles: University of California, Los Angeles. (MA thesis).

Stucky, Susan. 1983. Verb phrase constituency and linear order in Makua. In Gerald Gazdar, Ewan Klein & Geoffrey K. Pullum (eds.), *Order, concord and constituency*. Dordrecht: Foris Publications. DOI: 10.1515/9783110876734-005.

Toivonen, Ida. 2001. *The syntax of non-projecting words*. Stanford: Stanford University. (Doctoral dissertation).

Wurmbrand, Susi. 2017. Verb clusters, verb raising, and restructuring. In Martin Everaert & Henk van Riemsdijk (eds.), *The Wiley Blackwell companion to syntax*, 2nd edn. Hoboken: John Wiley & Sons. DOI: 10.1002/9781118358733.wbsyncom103.

Zaenen, Annie. Forthcoming. Information structure. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Zaenen, Annie & Ronald M. Kaplan. 1995. Formal devices for linguistic generalizations: West Germanic word order in LFG. In Jennifer S. Cole, Georgia M. Green & Jerry L. Morgan (eds.), *Linguistics and computation*, 3–27. Stanford: CSLI Publications. Reprinted in Dalrymple, Kaplan, Maxwell & Zaenen (1995: 215–240).

Chapter 2

LFG and historical linguistics

Hannah Booth
Ghent University

Miriam Butt
University of Konstanz

This chapter looks at the opportunities and perspectives that LFG offers for the study of language change, surveying existing LFG approaches within historical linguistics and providing examples of sample phenomena. We discuss how reanalysis, a major driver of language change, can be accounted for elegantly within LFG's parallel architecture thanks to its crucial separation of form from function and, moreover, how different types of reanalysis can be understood, whether they involve rebracketing, recategorization, or changes at the lexical level commonly discussed in terms of grammaticalization. As we also discuss, LFG's fundamental design principles and resulting flexibility of c-structure allow for complex, nuanced accounts of word order change. Furthermore, we survey the opportunities that LFG offers for exploring the complex relationship between variation and change, and in particular frequency effects and gradual change which proceeds via competition. Finally, we signpost future possibilities for work in this relatively underexplored but promising area.

1 Introduction

This discussion of historical linguistic work in LFG builds on two previous meta-discussions.[1] One is Vincent's (2001) wide-ranging and satisfyingly deep, comparative look at the possibilities which LFG's particular projection architecture

[1] An additional discussion of previous diachronic LFG accounts relating specifically to the history of English is Allen (2012).

Hannah Booth & Miriam Butt. Forthcoming. LFG and historical linguistics. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

and its combination of both functionalist and generative perspectives offer for an analysis of various different types of language change. The other is a recent handbook article by Börjars & Vincent (2017), which offers more detail on c-structural analyses of language change in several Germanic languages that have emerged since the seminal *Time Over Matter* book. The *Time Over Matter* book (Butt & King 2001) represents the first collection of historical linguistic work within LFG, in which Vincent's (2001) contribution represents more of a position paper than a mere introduction to a collected volume.

In this chapter, we provide a discussion of architectural issues and perspectives on language change in Section 2. However, our intention is not to reproduce the in-depth discussions already found in Vincent (2001) and Börjars & Vincent (2017), so we keep this section comparatively brief and proceed on to discussing examples of lexical and functional change in Section 3. This includes phenomena generally dealt with under the rubric of "grammaticalization", but also an understanding of complex predication, passives and case. Section 4 provides a discussion of language change at c-structure, which includes "growing" functional categories, understanding changes in word order and the syntactic configuration of a language and the development of mixed categories. Finally, we address the possibilities for modelling the complex relationship between variation and change within LFG in Section 5.

2 The LFG Architecture and mechanisms of language change

As Vincent (2001) and Börjars & Vincent (2017) point out, LFG is like most theories in the generative tradition in that it was not specifically designed with diachrony in mind. There is no paper tackling language change in the landmark Bresnan (1982) volume and historical work within LFG in the main began in the 1990s, an early exception being Allen (1986), which we discuss in Section 3.2. However, as demonstrated by Vincent (2001), LFG's fundamental design principles and its parallel projections are particularly well-suited to modelling diachronic change (see also the discussion of paradigms in Börjars et al. 1997).

In their textbook on language change, Harris & Campbell (1995) articulate a position whereby reanalysis, along with extension and borrowing, is seen as a key mechanism of language change. Reanalysis in their terms covers quite a broad range of phenomena, involving morphophonological and morphosyntactic changes.[2] A relevant touchstone here is Langacker's (1977: 59) classic definition

[2] While phonological change has been overall a central topic in historical linguistics, the focus in

of reanalysis, which Vincent (2001: 11) notes is most useful in an LFG setting, namely that reanalysis concerns "[a] change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation". Because LFG does not conflate syntactic position and syntactic functions and, thus, by extension, is able to cleanly separate out surface appearance (at c-structure) from functional and semantic import (at f-structure, a(rgument)-structure and s(emantic)-structure), it is particularly well-suited to help understand different types of reanalysis, whether they concern simple syntactic rebracketing, morphosyntactic change of the type where a dative argument or object is reanalyzed as a subject (Allen 1995; Schätzle 2018), the rise of a recipient passive (Allen 1995, 2001) or the development of complex predication (Börjars & Vincent 2017). Other changes may involve the reanalysis of one syntactic category as another (Börjars & Burridge 2011), also leading to the existence of mixed category phenomena (Nikitina 2008), for example, again with attendant functional changes. van Gelderen (2011) discusses such diachronic developments in terms of "Linguistic Cycles" and works with changes in feature specifications that are attached to lexical items and categories. The analyses are couched within Minimalism and work with a very restricted set of features — we would argue that LFG is much better poised to account for changes in feature systems in relation to phrase structure (see Section 4.1).

Cases of classic reanalysis at a lexical level, many of which have been prominently discussed as instances of grammaticalization (Hopper & Traugott 2003; Narrog & Heine 2017, but also see Roberts & Roussou 2003 within Minimalism) are also easily modelled and predicted by an architecture which separates out surface syntactic form (c-structure) from function (f-structure). As we discuss in Section 3.1, a verb can retain its surface form, but begin functioning as a perception raising verb, an auxiliary or a light verb (Barron 2001; Butt & Lahiri 2013). Over time, these functional changes may also result in a change in the surface form of the relevant item – typically some kind of morphophonological reduction, but also changes in the paradigmatic behaviour. The design of the LFG architecture allows for this associated process to be captured distinctly from the actual strict process of reanalysis as per Langacker's definition. In fact, it can also predict which types of functional and semantic elements are more prone to change than others and in what way. For instance, in terms of lexical or semantic "bleaching" one would predict that a verb loses its predicational power (the

this chapter is on morphosyntactic change, reflecting the centrality of syntax in the LFG architecture and the fact that diachronic work within LFG has focused on morphosyntax. A natural framework to work within from an LFG perspective with respect to sound change would be Lexical Phonology (Kiparsky 1982a,b; Mohanan 1986), for example as in Lahiri (2000a).

PRED feature) at f-structure, but retains certain functional information. Or, with respect to originally spatial terms being drawn into the case-marking system of a language, one can imagine working with underspecification and/or the loss of particular features characterizing the spatial terms, e.g., the PLACE and/or PATH specifications (Ahmed Khan 2009).

On the other hand, some instances of language change concern changes at the c-structural level, without necessarily resulting in attendant functional change. Examples are changes with respect to word order and constituent structure such as those found in Germanic and Romance languages, where a previously freer distribution and "discourse-configurational" organization (É. Kiss 1995) yields to a system where grammatical relations are increasingly licensed by position (Kiparsky 1995, 1997; Hinterhölzl & Petrova 2010; Luraghi 2010; Ledgeway 2012; Ponti & Luraghi 2018; Booth et al. 2017; Booth & Schätzle 2019; Booth & Beck 2021). Given that LFG's c-structure represents actual linear order, constituency and hierarchical relations, is not dogmatic about binary branching and allows for endocentric as well exocentric phrasal organization, there are several parameters across which languages would be expected to vary and change and they indeed do. Thus, the common trend for languages to shift over time from a freer word order to a more fixed word order can be captured in terms of the development of an increasingly endocentric c-structure, as we discuss in Section 4.3. In such a scenario, the mappings between c-structure and f-structure will necessarily change, fed by the changing positions licensed at c-structure, as typically manifested in the changing realization of grammatical relations, as we discuss in Section 4.5.

Of course, most instances of language change do not involve just one change within one module of grammar (i.e. c-structure or f-structure), but are more complex. Given the inherently interactional nature of language change, certain changes which initially occur at c-structure may in turn feed changes at f-structure, and vice versa. In this sense, keeping surface form, syntactic categorization and functional information apart as in the LFG architecture allows one to neatly model the step-wise nature of such developments. Vincent (2001) notes that one of the most complex series of changes he has seen analyzed is that presented by Simpson (2001) on the grammaticalization of associated path in Warlpiri (see Section 3). Indeed, as Vincent (2001) also points out, another consequence of the complex interactional architecture of LFG is that an LFG perspective on language change does not expect abrupt, cataclysmic shifts in grammar as proposed in the influential work by Lightfoot (1979, 1991, 1997), for example. Rather, it is expected that a series of small changes, many of them at the lexical level, will combine together and gradually, over time and with attendant variation in usage, will result in a major structural change. This is what is argued for with respect to the introduction of ergativity in Indo-Aryan by Butt & Ahmed Khan (2011), for example (cf. also Traugott & Trousdale (2010) for an overview discussion).

Indeed, variation as an inherent property of language has been closely linked to language change in a strong tradition of work (e.g., Kroch 1989, 2001; Labov 1994; Pintzuk 2003). However, this empirical fact generally presents a challenge for generative frameworks, which did not originally feature gradience or stochastic variation as part of their basic design principles. One proposed architectural solution here has been in terms of Optimality Theory (Kager 1999), in particular stochastic Optimality Theory (Boersma & Hayes 2001). As we discuss in Section 5, this avenue has also been explored within LFG as a way to model gradual syntactic change via competing variants (Clark 2004), using OT-LFG (see Kuhn forthcoming [this volume]).

Having briefly surveyed the explanatory potential of LFG with respect to diachronic change via a set of examples in this section, we delve into the issues and phenomena more deeply in the next sections, also involving other parts of LFG's projection architecture, most prominently a(rgument)-structure.

3 Lexical and functional change

We begin by going through examples of lexical and functional change in Section 3.1, many of which have been analyzed as instances of grammaticalization. We then move on to more complex series of changes which involve a restructuring of the mapping relationship between semantic arguments (a-structure) and grammatical relations (f-structure) in Section 3.2.

3.1 Grammaticalization

The original characterization of the idea behind grammaticalization goes back to Meillet (1912: 131), who defines it as: "l'attribution du caractère grammatical à un mot jadis autonome [the attribution of a grammatical value to a formerly autonomous word]" (Vincent & Börjars 2020: 134). Essentially, this is a process by which an item with lexical content becomes reanalyzed as a functional element. Recent decades have seen a substantial body of work on grammaticalization phenomena, where grammaticalization has been treated both as a grammatical framework (e.g., Lehmann 1985; Hopper & Traugott 2003) and as a pre-theoretical notion which is to be formalized via the tools and concepts available within a particular framework (e.g., Campbell 2000; Newmeyer 2000; Roberts & Roussou

2003; van Gelderen 2011). LFG belongs in the latter class: it sees grammaticalization as a pre-theoretical concept which describes certain observed historical changes that are to be modelled via the formal tool-kit and assumptions available as part of the projection architecture.

The progression from lexical to functional is typically not accomplished in one fell swoop, but consists of the combined effects of a number of individual changes (see, e.g., the various papers in Traugott & Trousdale 2010). The grammaticalization literature proposes that change progresses along a cline, for example as shown in (1) for a crosslinguistically well-established change in which auxiliaries, clitics and finally affixes develop from an originally contentful lexical verb. This change is also generally associated with the concept of "semantic bleaching", by which the item undergoes the gradual loss of semantic content until the formerly lexical content word is reanalyzed as a functional element.

(1) full verb > (vector verb) > auxiliary > clitic > affix
Typical Grammaticalization Cline (Hopper & Traugott 2003: 108)

Note that the cline represents a mixture of surface and functional changes (form and function): the change from a main (full) verb to auxiliary mainly revolves around a change in function, while the change from auxiliary to clitic/affix is very prominently a change in surface form. Given that LFG very clearly differentiates between form (c-structure) and function (f-structure), it seems particularly perspicuous to address issues of grammaticalization from the perspective of LFG, as we aim to illustrate in this section.

The category *vector verb* in the cline in (1) was introduced specifically for instances of light verb formation in Indo-Aryan (Hook 1991) and this has been taken up in discussions within LFG by Butt & Lahiri (2013), who argue that light verbs should not be placed on a grammaticalization cline, but are diachronically stable. Butt and Lahiri contrast the diachronic evidence available for Indo-Aryan light verbs with that of auxiliary formation and show that these two categories exhibit very different diachronic behaviour. While light verbs show no signs of morphophonological surface changes or further functional changes which often follow a categorial reanalysis in instances of grammaticalization, auxiliaries do. This is illustrated in (2) for the Bengali verb 'be' and in (3) for the Urdu verb 'go'.

In Bengali the verb ɑʧʰ 'be' can function as a full verb (2a), but also as a light verb (2b), in which case it is always form-identical to the main verb. On the other hand, the same verb 'be' has given rise to new verbal paradigms whereby the perfect is realized via a cliticized version of a former auxiliary version of ɑʧʰ 'be' (2c) and the progressive shows a fully affixal version (2d).[3] The cliticized and affixal versions of a former auxiliary are as expected/predicted by the grammaticalization cline in (1).

(2) Bengali

a. ɑmi bʰɑlo ɑʧʰi
I.NOM well be.PRES.1
'I am well.' Main Verb

b. amar mone ɑʧʰe
I.GEN mind.LOC be.PRES.3
'I remember.' Light Verb

c. rɑm ʧitʰi pe-(y)e=ʧʰ-ilo
Ram.NOM letter.NOM receive-PERF=be-PAST.3
'Ram had received letters.' Clitic

d. rɑm ʧitʰi pɑ-ʧʧʰ-ilo
Ram.NOM letter.NOM receive-be-PAST.3
'Ram was receiving letters.' Affix

Similarly, the verb *ja* 'go' in Urdu/Hindi has a light verb use (3b) that is always form-identical to its main verb use (3a). When the surface form of the main verb changes due to language change, the light verb version mirrors this change. On the other hand, the auxiliary version of 'go' that furnished the basis for the innovated future morpheme in Urdu/Hindi went through a clitic phase (3c) and is now an affix whose surface form is -g-, as in (3d).

(3) Urdu/Hindi

a. mɛ̃ gɑ-ya
I.NOM go-PERF.M.SG
'I went.' Main Verb

b. bɑʧʧa gɪr gɑ-ya
child.M.NOM fall go-PERF.M.SG
'The child fell (down).' Light Verb

c. kɑh-ũ=hi=ga
say-1.SG-EMPH-FUT.M.SG
'I will say (it), of course.' Clitic
(Kellogg 1893: s399)

[3] For a full analysis of the morphophonological changes that led to the formation of new verbal paradigms in Bengali, see Lahiri (2000b).

d. pʊlɪs ʧor=ko pɑkr̥-e-g-i
police.F.SG.NOM thief.M.SG=ACC catch-3.SG-FUT-F.SG
'The police will catch the thief.' Affix

Butt and Lahiri focus on developing a theory as to why light verbs should be diachronically stable, proposing an underspecified approach to the deployment of lexical semantic information by which light verbs are inextricably linked to their full verb versions via a single underlying entry, see (4). When deployed as a light verb, they require combination with another predicational element, with which they form a complex predicate.

(4)

Underlying Verb
- Main Verb —— Auxiliary
- Light Verb

The diachronic path of change from verb to auxiliary to clitic and potentially an affix is assumed to be based on the main verb version. Along with other work on grammaticalization in LFG, Butt and Lahiri assume that grammaticalization primarily involves a loss or difference in functional information at f-structure, while the surface form is initially held constant. That is, the difference between a main verb use of 'go' and an auxiliary use of 'go' would be expressed in terms of a difference in functional information associated with the respective lexical entries. In the illustrative main verb entry (V) in (5a) vs. the auxiliary version (I) in (5b) one major difference in functional information involves the loss of the predicational power of the verb in terms of its PRED function. This then also instantiates the "semantic bleaching" generally observed in the grammaticalization process.

(5) a. *go* V (↑ PRED) = 'GO⟨(↑ SUBJ)(↑ XCOMP)⟩'
b. *goes* I (↑ TENSE) = FUT

Thus, in the main verb use, the verb 'go' subcategorizes for a SUBJ and an XCOMP. In the auxiliary use that develops over time, this information is absent and is instead replaced by a futurate use of the verb. As such, the auxiliary version then merely provides tense information to an overall predication. That is, grammaticalization primarily involves a change in the functional information associated with an item. This functional change then engenders further changes, such as the reanalysis of the syntactic category of the item (from V to I) and possible ensuing changes in the morphophonological realization of the item due to its new functional status, so that the item eventually develops into an affix (typically via a clitic stage), as illustrated for Bengali and Urdu/Hindi in (2) and (3) above.

Grammaticalization does not tend to occur in one sudden step, but happens gradually over time and tends to involve several intermediate steps. It also does not take place randomly, but is generally motivated by a semantic reinterpretation of a given configuration (e.g., 'goes to go' ⟶ 'will go'). This type of semantically motivated change is also discussed in a recent paper by Börjars & Vincent (2019) with respect to Germanic *will* verbs. Börjars & Vincent (2019) propose an LFG analysis of how an original verb of desire ('want') undergoes change to a verb of intention and further to prediction, giving rise to a new modal in some languages and a futurate auxiliary in others. This semantic change goes hand in hand with a change in functional information (e.g., from a control to a raising verb) and a concomitant reanalysis at c-structure.

Similarly, Camilleri & Sadler (2018, 2020) postulate a total of four separate steps in the formation of a progressive auxiliary from a main verb meaning 'sit' in Arabic. Unlike Butt and Lahiri and Börjars and Vincent, who work with diachronic data, but in keeping with many studies on grammaticalization, the evidence Camillieri and Sadler adduce is mainly from synchronic variation found in dialects of Arabic, which are taken to be indicative of stages of diachronic development.

Camillieri and Sadler associate the origin of the progressive auxiliary in Arabic with constructions in which the posture verb 'sit' is used together with an adjunct clause, as in (6). The verb 'sit' is considered to be a V that projects a VP and this is modified by a VP adjunct. The corresponding (simplified) f-structure analysis is given in (7).[4]

(6) Wādi Ramm Jordanian Arabic
(Almashaqba 2015, cited by Camilleri & Sadler 2020: 24)

lagē-ta-h	gāʕid	ya-smaʕ	al-giṣidah
find.pfv-1sg-3.sg.m.acc	sit.act.ptcp.sg.m	3m-hear.1pfv.sg	def-poem

'I found him sitting down listening to the poem.'

[4]Note that the f-structure in (7) differs from the original one in Camilleri & Sadler (2020: 26) in that we have rendered the xadj as a set containing one element, which is what is described in Camilleri & Sadler (2020: 26), but not represented in their f-structure.

(7)
$$\left[\begin{array}{ll}
\textsc{pred} & \text{‘}\textsc{sit}\langle\textsc{subj}\rangle\text{’} \\
\textsc{subj} & \left[\begin{array}{ll}\textsc{pred} & \text{‘}\textsc{pro}\text{’}\end{array}\right] \\
\textsc{xadj} & \left\{\left[\begin{array}{ll}
\textsc{pred} & \text{‘}\textsc{listen}\langle\textsc{subj obj}\rangle\text{’} \\
\textsc{subj} & \\
\textsc{obj} & \left[\begin{array}{ll}\textsc{pred} & \text{‘}\textsc{poem}\text{’} \\ \textsc{def} & +\end{array}\right]
\end{array}\right]\right\}
\end{array}\right]$$

In a second stage this optional modifying XADJ is reanalyzed as an obligatory clausal complement of the posture verb ‘sit’, as shown in the f-structure in (9) for the attendant example in (8).

(8) Kuwaiti Arabic (Camilleri & Sadler 2020: 28)
layla gāʕd-a ta-dris
Layla sit.ACT.PTCP-SG.F 3F-study.IPFV.SG
‘Layla is (sitting) studying.’

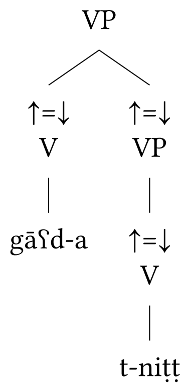
(9)
$$\left[\begin{array}{ll}
\textsc{pred} & \text{‘}\textsc{sit}\langle\textsc{subj xcomp}\rangle\text{’} \\
\textsc{subj} & \left[\begin{array}{ll}\textsc{pred} & \text{‘Layla’}\end{array}\right] \\
\textsc{xcomp} & \left[\begin{array}{ll}
\textsc{pred} & \text{‘}\textsc{study}\langle\textsc{subj}\rangle\text{’} \\
\textsc{subj} &
\end{array}\right]
\end{array}\right]$$

This use of ‘sit’ in combination with a clausal argument is in turn reanalyzed as signaling durational, stative semantics, abstracting away from the original postural, locational meaning. Once the step from a concrete postural meaning to an aspectual meaning dimension is made, the verb is assumed to lose its predicational power in terms of the PRED feature and to only contribute the durational aspectual information to the clause, resulting finally in an innovated progressive marker (cf. Deo (2015) for more discussion and evidence of this type of crosslinguistically attested language change). An example is shown in (10) with a corresponding f-structure analysis in (11). Under Camillieri and Sadler’s analysis the syntactic category of ‘sit’ itself is not reanalyzed; it merely no longer projects a VP of its own, but functions as a co-head with the formerly embedded verb, as shown in (12).

(10) Kuwaiti Arabic (Camilleri & Sadler 2020: 30)
gāʕd-a t-niṭṭ
PROG-SG.F 3F-jump.IMPFV.SG
‘She is jumping.’

(11) $\left[\begin{array}{ll} \textsc{pred} & \text{‘}\textsc{jump}\langle\textsc{subj}\rangle\text{’} \\ \textsc{subj} & \left[\textsc{pred}\ \text{‘}\textsc{pro}\text{’}\right] \\ \textsc{tense} & \textsc{pres} \\ \textsc{asp} & \textsc{prog} \end{array}\right]$

(12)



The change by which an adjunct is reanalyzed over time as a core argument of a verb has also been argued to play a role in Latin in the innovation of raising predicates such as ‘seem’ from verbs of perception such as ‘see’ (Barron 2001) and the grammaticalization of associated path in the Australian languages Warlpiri and Warumungu (Simpson 2001). It also plays a role in the spread of dative subjects in Icelandic, as argued for by Schätzle (2018) and discussed in Butt forthcoming [this volume], as well as in Section 3.2 below.

Barron (2001) provides a theoretically sophisticated account for the development of Latin *videri* ‘seem’ from the perception verb *videre* ‘see’. The general idea is that the epistemic raising verb develops from a passivized version of *videre* in situations where there is a small clause (secondary predication), such as ‘Laelius was seen as an ideal person.’ This was reinterpreted as ‘Laelius seemed an ideal person’ and over time was generally concomitantly structurally reanalyzed as a raising predicate. The analysis is complex and involves changes at the semantic level which translate into functional changes at f-structure.

Another level of complexity is added by Simpson’s (2001) account of associated path in Warlpiri and Warumungu. The puzzle she addresses is how the path expressions (‘thither’ vs. ‘hither’) in (13) came to be grammaticalized as morphemes on a verb, given that the languages generally allow for free word order. She assumes that at some point there must have been a stage in which the path expressions were preferentially placed just after the verbs and that this preferential word order then paved the way for grammaticalization along the cline in (1).

(13) Warumungu (Simpson 2001: 174)

a. **Juku-nturrarni**=angi angkinyi kina ngurraji kina?
carry-THITHER.PAST=2S your to camp to
'Did you take it to your home?'

b. **Juku-ntukarni**=ajjul ngurraji kina.
carry-HITHER.PAST=3S camp to
'They (more than two) brought it here to camp.'

Again, the change is conceived of as a complex chain of reanalyses. The original construction is taken to be one in which a clause like 'I went to camp' is modified by a clausal adjunct, for example: 'after yam digging'. This adjunct was then preferentially realized clause-initially: 'after yam digging, (I) went to camp', thus placing the verbs next to one another in certain situations. This adjacent placement of the verbs is thought to have triggered clause unification, yielding a monoclausal structure in which the former verb of motion is eventually reanalyzed as a morpheme expressing the associated path of the event. We thus have a preferential word order opening the way for a semantic, then syntactic and concomitant functional reanalysis of an original verb into a bound morpheme.

For further discussions and examples of grammaticalization approaches within LFG, also contextualized in terms of comparison of approaches across theories, see Vincent & Börjars (2020). We discuss some aspects of their paper in more detail in Section 4.1, since some of the case studies involve a reanalysis of syntactic categories with attendant "mixed" effects. We return to grammaticalization in Section 4.3 in the context of c-structural change. Before turning to these topics, we discuss instances of language change which primarily involve a change in the linking configuration between semantic arguments and grammatical relations in Section 3.2.

3.2 Arguments and linking

In the previous section on grammaticalization we discussed phenomena of language change that involved a number of different dimensions. In this section, we focus on changes that are primarily concerned with reconfigurations in the linking between semantic arguments and grammatical relations. Work that addresses these kinds of specific changes within LFG is: Allen (1986, 1995); Kibort & Maling (2015); Schätzle (2018) and Beck & Butt (2021).

3.2.1 Experiencer verbs

As mentioned above, a very early application of LFG to diachronic change is Allen (1986), who considers the verb *like* in the history of English. This verb can be analyzed as having an Experiencer (the liker) and a Cause (the cause of liking).[5] Such verbs are interesting diachronically because they show a change in the correspondence between semantic arguments (experiencer, cause) and grammatical relations. In Old English the experiencer argument has the positional and morphological properties of an object, e.g., (14a), but is uncontroversially a subject in Present-day English, e.g., (14b).[6]

(14) a. Old English (Jespersen 1927, as cited in Allen 1986: 376)
Ðam cynge licodon peran
the.DAT king.DAT liked.PL pears
'Pears were pleasing to the king'

b. 'He liked pears.'

Based on detailed investigations of the historical data, Allen (1986) challenges the traditional account for this change (e.g., Jespersen 1927; Lightfoot 1979, 1981), which casts it in terms of a reanalysis of preverbal object experiencers as subjects, as a direct consequence of the loss of case-marking and the fixing of SVO word order. As Allen (1986) points out, there are various problems for this account, including the fact that the OVS order required as a source for the reanalysis is relatively rare with the verb *like*, and because of chronological issues concerning the link with the loss of case-marking. Moreover, the data indicates that the change proceeded gradually, with subject and object experiencers coexisting alongside one another for several centuries over the course of Middle English and Early Modern English, which is not compatible with a "catastrophic" reanalysis account as proposed by Lightfoot, for example.

In light of these observations, Allen (1986) puts forward an alternative account, which involves a gradual change in the mapping correspondences between semantic arguments and grammatical functions, modelled in terms of the introduction and gradual favouring of a new lexical subcategorization frame, employing an early LFG approach to this type of mapping. The new subcategorization frame with a dative-marked subject experiencer, shown here in (15b), is already available in Old English for the verb *lician* 'like', but sits alongside and is less common

[5] Alternatively, this semantic role has been referred to as stimulus or theme, as in, e.g., later work by Allen (1995).

[6] As has been pointed out (Denison 1993: 81), the original example from Jespersen in (14a) is invented and represents a pattern which is in fact rather rare.

than the older frame in (15a) where the dative-marked experiencer maps to object. According to Allen (1986), the gradual favouring of the correspondences in (15b) coincides with changes concerning the assignment of case-marking, specifically a shift from a system where case is lexically assigned to one in which it is structurally assigned on the basis of grammatical relations. Structural case-marking for objects is introduced in the early thirteenth century according to Allen and specification of case-marking for the experiencer subject is optional as of the mid-fourteenth century. Under pressure towards consistent structural case assignment, all lexically determined case-marking is finally lost, subjects are consistently nominative-marked and the frame for allowing object experiencers in (15a) is no longer available.

(15) a. Older frame (adapted from Allen 1986: 388):

			EXP	CAUSE	
lician	V	PRED 'LICIAN <	(OBJ)	(SBJ)	>'
		(↑ OBJ CASE) = DAT			
		(↑ SBJ CASE) = NOM			

b. Newer frame (adapted from Allen 1986: 394):

			EXP	CAUSE	
lician	V	PRED 'LICIAN <	(OBJ)	(SBJ)	>'
		(↑SBJ CASE) = DAT			
		(↑OBJ CASE) = NOM			

Allen (1995) develops this account of experiencer verbs in the history of English further in terms of Lexical Mapping Theory (e.g., Bresnan & Kanerva 1989, cf. Butt forthcoming [this volume] and Findlay & Kibort forthcoming [this volume]), discussing a wider range of data and additional changes including the rise of the recipient passive, which we discuss in Section 4.5. In particular, she demonstrates that lexical semantic factors, rather than loss of case-marking drives the change with respect to experiencer verbs, and that this can be elegantly modelled with LFG's richly articulated lexicon and Lexical Mapping Theory.

3.2.2 Passives and impersonals

Kibort & Maling (2015) address the innovation of a new impersonal construction in Icelandic which they argue has emerged as a syntactically active construction via reanalysis of an impersonal passive with passive morphology. The new construction is thought to have been emerging approximately over the last fifty years (Thráinsson 2007) and has prompted a good deal of debate concerning what the precise analysis should be (Maling & Sigurjónsdóttir 2002; Eythórsson 2008; Jónsson 2009). Maling & Sigurjónsdóttir (2002) show that this is currently a change in progress and that speakers of Icelandic show variation in the interpretation of examples such as (16): some interpret it as an impersonal passive (Reading A), others as an active transitive with a [+human] agentive PRO subject (Reading B). Such variation is expected when there is a change in progress, as we discuss further in Section 5.

(16) Icelandic
Loks var fundið stelpuna eftir mikla leit.
finally was found.N.SG girl.the.F.ACC after great search
Reading A: 'The girl was finally found after a long search.'
Reading B: 'They finally found the girl after a long search.'

Kibort & Maling (2015) argue that this variation and change arises when a former impersonal passive with passive morphology is reanalyzed as an active form. They argue for a series of step-wise changes, beginning with the potential for the linking configurations of an impersonal passive and a regular passive being confused with one another when the OBL_θ agent argument of the passive is left unexpressed, as is often the case in Icelandic language use. This means that a regular passive on the surface often looks very much like an impersonal passive, where there is a covert PRO SUBJ, as illustrated in (17) for transitive verbs (adapted from Kibort & Maling 2015).[7]

(17) a.
$$\begin{array}{lcccc} verb_{passive} & < & \text{arg}_1 & \text{arg}_2 & > \\ & & [-\text{o}] & [-\text{r}] & \\ & & [+\text{r}] & & \\ & & (\text{OBL}_\theta) & \text{SUBJ} & \end{array}$$

b.
$$\begin{array}{lcccc} verb_{impers\_passive} & < & \text{arg}_1 & \text{arg}_2 & > \\ & & [-\text{o}] & [-\text{r}] & \\ & & [+\text{r}] & & \\ & & \text{PRO}_{impers} & \text{SUBJ} & \end{array}$$

Kibort & Maling (2015) assume Kibort's version of Mapping Theory (Kibort 2013, 2014) by which argument slots and types are defined via an overall template

[7]In the version of Mapping Theory which Kibort & Maling (2015) employ, argument positions (arg1, arg2 etc.) are separated out from semantic participants and the thematic roles they instantiate. Here, for ease of exposition, and because this separation is not relevant for the changes discussed, we just represent the argument positions.

allowing for specific types or argument slots, shown above as ‘arg1’, ‘arg2’ in the case of transitives. The linking between arguments and grammatical relations is accomplished via the [±r,o] features of LFG’s standard Mapping Theory. See Findlay & Kibort forthcoming [this volume] for a detailed exposition on Kibort’s Mapping Theory, also cf. Butt forthcoming [this volume]. Maling and Kibort discuss several different paths of change that are predicted as possible within their assumptions as to Mapping Theory and draw on additional examples from Slavic as well as Mayan to illustrate the possible changes.

With respect to Icelandic, they propose that “non-promotional” versions of the passive opened the way for a reanalysis of the impersonal passive as an impersonal active. In these non-promotional passives, the patient/theme argument of transitives is not “promoted” and realized as a SUBJ, as is usually the case under passivization. Instead, it is realized as an OBJ in examples such as (16) (as indicated by the accusative marking on ‘girl’ in (16)). The configuration for the non-promotional passive is shown in (18a). This configuration is in turn very close to that of an active impersonal in which there is a PRO SUBJ and so that is what it is reanalyzed as, see (18b).

(18) a.

verb$_{passive\_obj}$	<	arg$_1$	arg$_2$	>
		[−o]	[−r]	
		[+r]		
		(OBL$_\theta$)	OBJ	

b.

verb$_{impers\_active}$	<	arg$_1$	arg$_2$	>
		[−o][−r]	[−r]	
		PRO$_{impers}$	OBJ	

Thus, a series of reanalyses that initially arose out of ambiguous surface structures are seen to lead to an overall diachronic reanalysis in which an originally passive construction is reinterpreted as a transitive, syntactically active impersonal with a PRO SUBJ and an OBJ, as illustrated in (18). This reanalysis was enabled by the variation in interpretation that arose from the ambiguous surface structures. Note that under this scenario, the verb does not change, nor does the passive morphology. The surface realization remains the same. What changes is the mapping or linking between arguments of the verb and the grammatical relations.

3.2.3 Dative subjects

Another example of a change involving only the linking configuration between semantic arguments and grammatical relations is the rise and spread of dative subjects in Icelandic and Indo-Aryan. Dative subjects were innovated as part of diachronic developments in New Indo-Aryan (NIA) from about 1100 CE onwards. Icelandic dative subjects can be traced back to the earliest documented stages of the language, but these only go back to 1150 CE, about the same time as the new case marking systems of the NIA cousins were developing.

Schätzle (2018) analyzes diachronic data from Icelandic and finds that dative subjects in Icelandic have increased over time. Besides the well-documented process of "dative sickness" or "dative substitution" (Smith 1996; Jónsson 2003; Barðdal 2011), by which accusative experiencer subjects are systematically replaced by datives, Schätzle finds that dative subjects arise via originally middle forms of verbs of searching or perception to give rise to lexicalizations of experiencer predicates which take a dative subject. As in the example of Latin raising verbs (Barron 2001) and the Arabic progressive (Camilleri & Sadler 2020), Schätzle identifies an intermediate stage involving secondary predication as an important step in the series of reanalyses that take place. We do not provide Icelandic examples and details of Schätzle's analysis here; the interested reader is referred to Butt forthcoming [this volume] for a summary and examples.

Schätzle works out a theory of Linking or Mapping that is based on Kibort (2013, 2014), but that crucially integrates an event-based approach. She includes a notion of subevental participants that draws on Ramchand's (2008) tripartite view of events. She further introduces a way of determining relative argument prominence by including a notion of Figure vs. Ground (Talmy 1978), as well as information on Proto-Role properties (Dowty 1991) as suggested by Zaenen (1993) for LFG. The resulting linking system is complex, but it does justice to the complex interface between morphosyntax and lexical and clausal semantics that is involved in the relationship between semantic roles, event semantics and the realization of grammatical relations.

Beck & Butt (2021) refine Schätzle's framework and address dative subjects in both Icelandic and Indo-Aryan. The general linking schema they assume is shown in (19). As can be seen, a maximum of four argument slots are assumed. This number derives from the maximum of four subevental parts identified by Ramchand: 1) the init(iation) subevent, which requires an initiator (or agent) of the event; 2) the proc(ess) or progress of the event, which requires an undergoer or patient of the process; 3) the res(ult) of the event, which requires a resultee argument (often but not necessarily identical to the undergoer of the process).

Finally there is (4) the rheme, which is not strictly speaking a subevent, but which can serve to modify the overall event in some way, i.e. by providing information on where the event took place or the manner in which it took place.

(19) General Linking Schema

		init	proc	res	rh	
Predicate	<	x	x	x	x	>
		FIGURE	GROUND			
		SUBJ	OBJ	OBJ$_{theta}$	OBL	

The linking in (20) provides an example of a typical agentive transitive verb. The verb ‘kill’ has two arguments. One of these (‘Indra’) is associated with the initiation subevent and thus serves as the initiator/agent participant of the event. The other argument (‘serpent’) is the undergoer of the event and thus affected as part of the on-going proc(ess) of the event, with a clear res(ult), namely that it is dead. This argument is thus associated with two subevents.

(20) Indra killed the serpent.

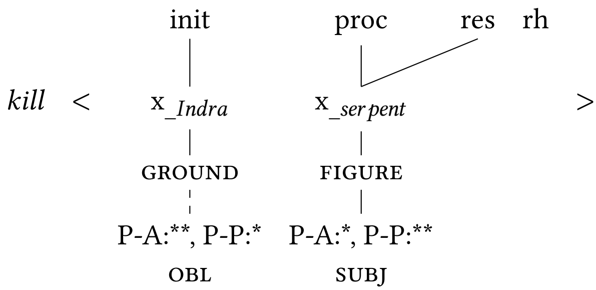
		init	proc	res	rh	
kill	<	x_{Indra}	$x\_{serpent}$			>
		FIGURE	GROUND			
		P-A:***	P-P:***			
		SUBJ	OBJ			

The initiator is naturally also the Figure of the event, and the undergoer serves as the Ground. This basic linking constellation is interpreted with respect to Proto-Role properties in the following way: a) one Proto-Agent (P-A) property each is adduced for: initiation and Figure; b) one Proto-Patient (P-P) property each is adduced for: proc, res, and Ground. In addition, sentient arguments accumulate an additional P-A property. The Proto-Role properties are registered via an '*' in the linking schemas.

Overall, the linking of arguments works as follows. If an argument has more P-A than P-P properties, it is linked to SUBJ. An argument with more P-P than P-A properties is linked to OBJ. When an argument has equally many P-A and P-P properties or no P-A and P-P properties, then other information must be taken into account. In this case the type of subevent the participant is associated with is taken to play a crucial role. That is, if the argument associated with init vs. res have equal amounts of P-A and P-P properties, the init argument will be associated with SUBJ. This is also true for third and potentially fourth arguments of an event – once the SUBJ and OBJ linking has been determined, the subevental semantics play a role in determining the linking to a secondary object (OBJ_θ) or an oblique (OBL). Obliques are likely to correspond to spatial terms and paths (rhemes) or an init-GROUND combination. Secondary objects are likely to be related to undergoer semantics. An init argument that serves as the GROUND rather than the FIGURE is prohibited from being linked to the SUBJ — this is the well-known effect of passivization that has often been described as "demotion" or "inversion" in the literature (e.g., Perlmutter & Postal 1977).

This constellation is illustrated in (21). The association of arguments with subevents remains the same, but the Figure-Ground relationship is flipped. This affects the number and type of Proto-Role properties associated with each argument.

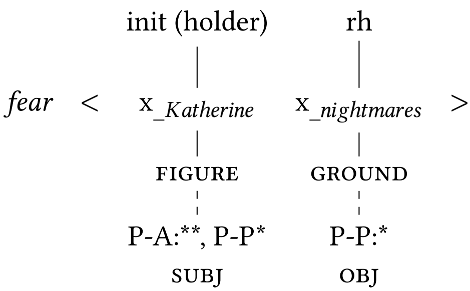
(21) The serpent was killed by Indra.



Since ‘Indra’ is no longer available to be linked to SUBJ, the ‘serpent’ is the SUBJ. Because ‘Indra’ still has more P-A (one for init, one for sentience) than P-P (one for GROUND) properties, it is not associable with an OBJ or OBJ_θ, but is linked to OBL.

An example of a linking configuration for an experiencer predicate is provided in (22). As per Ramchand’s analysis, the holder of the state of experiencing something is associated with the init subevent, and the stimulus is analyzed as a rheme (since it is neither part of the process or the result of the overall event). The experiencer ‘Katherine’ in (22) is also the FIGURE; the stimulus ‘nightmares’ is the GROUND. As a sentient argument who is also a FIGURE, ‘Katherine’ receives two P-A properties. As the holder of a state, this argument receives one P-P property. The ‘nightmares’ accumulate one P-P property from being associated with GROUND. Because rhemes are not properly event participants, they contribute neither P-A nor P-P properties for the calculation. Since ‘Katherine’ has the most P-A properties, it is linked to SUBJ. The ‘nightmares’ argument has only P-P properties and is thus linked to OBJ.

(22) Katherine fears nightmares.



With this basic event-based linking schema in place, Beck & Butt (2021) chart a path of development for dative subjects in New Indo-Aryan. A crucial component is the innovation of ergative transitive active clauses from originally adjectival participles which featured a nominative and an instrumental adjunct (e.g., 'The by Indra killed serpent.'). Beck and Butt posit that the original instrument adjunct was a GROUND which was reinterpreted as a FIGURE in situations where the instrument could be seen as a sentient agent. This then opened the door to further Figure-Ground flips, such as with originally locative structures as in (23) (cf. also Bresnan & Kanerva 1989; Landau 2010 on locative inversion).

(23) Urdu/Hindi
ındra=ko dɑr lɑg-a
Indra.M=Dat fear.M.Nom be.attach-Perf.M.Sg
'Indra was afraid.'

Their proposal is that the original locative predication involved a linking configuration as in (24). The overall predication is stative, so the two arguments involved are a holder of a state and a rheme. The 'Indra' argument is the location of the 'fear', so Indra is associated with a rheme and the GROUND. The 'fear' argument is then interpreted as the holder of a state: as a FIGURE it is located somewhere and receives one P-A property for being a Figure, and one P-A property for being the holder of a state. 'Indra' receives one P-P property for being the GROUND and one P-A property because it is a sentient argument. Both arguments thus have an equal number of P-A and P-P properties, but 'fear' is linked to SUBJ because it is associated with the init subevent.

(24) Fear was attached to Indra.

	init (holder)	rh (loc)	
be.attach <	x_*fear*	x_*Indra*	>
	FIGURE	GROUND	
	P-A:*, P-P:*	P-A:*, P-P:*	
	SUBJ	OBL	
	nominative	'at/to'	

It is not difficult to see that the linking configuration in (24) is unstable. The two arguments have equal numbers of P-A and P-P properties and the sentient argument is associated with GROUND, which is non-canonical (Talmy 1978). Beck and Butt propose that as a consequence, in a series of steps, both the Figure-Ground relation and the association with init and rheme are flipped and the resulting linking configuration is as shown in (25), corresponding to (23). This configuration is clearly more stable as the sentient argument is more prominent and accumulates more P-A properties.

(25) Indra was afraid.

		rh	init (holder)	
be.attach	<	x_*fear*	x_*Indra*	>
		GROUND	FIGURE	
		P-P:*	P-A:**, P-P:*	
		OBJ	SUBJ	
		nominative	dative	

The original spatial adpositions ('at, to') in fact gave rise to the current dative/accusative case markers in Urdu/Hindi (and across New Indo-Aryan), resulting in the innovation of dative subjects from originally spatial terms (cf. Montaut 2003; Butt & Ahmed Khan 2011).

To summarize, the innovation of dative subjects is seen primarily as the reanalysis of an originally stative locative predication as an experiencer verb. The main change involves a reanalysis of various parts of the overall linking configuration. Beck and Butt see this as being common to both Icelandic and Indo-Aryan. In Indo-Aryan, however, there is an additional concomitant but independent change in syntactic category, from a spatial adposition to a case marker. Changes involving reanalysis of an item's syntactic category are taken up in Section 4.1, with Section 4 focusing overall on change with respect to c-structure.

4 Syntactic change

4.1 Recategorization

Studies of language change abound with instances of syntactic recategorization, that is, instances in which an item belonging to one syntactic category is reanalyzed as belonging to a different one. We have already caught a glimpse of such a reanalysis with respect to the grammaticalization cline in (1), whereby a main verb is gradually reanalyzed as an auxiliary, which in turn often becomes a verbal affix. Within LFG, such c-structural recategorization is seen as being preceded by a change in an item's functional import. That is, with respect to the well-studied changes such as the development of new futurate markers (e.g., Fleischman 2009), as we saw in Section 3.1 an originally fully predicating verb such as 'want' or 'go' can be used in situations describing the future attainment of a state or event (Börjars & Vincent 2019). Over time, this usage becomes conventionalized and the verb is seen as routinely fulfilling an additional function, namely the temporal placement of an event in the future. This meaning of the verb ceases to predicate fully and it develops into a functional item. Often the original lexical/content verb continues to exist side-by-side with the new auxiliary; in other cases it ceases to be used as a main verb. In English, for example, the item *will* is no longer used as a modal meaning 'want', but only as a futurate marker. On the other hand, as we saw for the examples taken from Urdu and Bengali in Section 3.1, the verbs 'be' and 'go' continue to exist as main verbs while also serving as auxiliaries and giving rise to new verbal affixes.

Within this same verbal domain, Börjars & Vincent (2017) argue for the historical development of a causative light verb in Romance from the Latin verb *facere* 'make, do', which in turn results in the reanalysis of a formerly biclausal construction as a monoclausal predication. This reanalysis of biclausal predications, where one verb embeds another into monoclausal structures, also generally results from the reanalysis of main verbs as auxiliaries (Butt et al. 2004; Butt 2010).

Vincent & Börjars (2020) go through a number of further examples of syntactic recategorization from an LFG perspective, including the development of adpositions from nouns, infinitival markers from prepositions, complementizers from prepositions (P to C; Vincent 1999) and case-marking functions from prepositions. They also engage in a comparison of analyses across frameworks (Minimalism and HPSG) and ask the question of whether anything in LFG's architecture predicts the mainly *unidirectional* change in categorial reanalysis (meaning that a verb will change into an affix, but an affix does not change into a fully predicating main verb). The answer to this question is "no", unlike the clear predictions made by Roberts & Roussou (2003) within Minimalism, for example, where such reanalysis is seen as an instance of a lexical category raising upwards into the functional domain of a syntactic tree and eventually being reanalyzed as simply originating in that functional position. Upward "mobility" is expected in this framework, while downward movement is prohibited. However, Vincent & Börjars (2020) point out that while this type of grammaticalization along a cline from more lexical to more functional can be accounted for well within Minimalism, instances of "lateral" change whereby the recategorization involves adjacent categories like deictic markers into copular verbs (Börjars & Vincent 2017) are more challenging. In this case, an originally nominal category is reanalyzed as belonging to the verbal domain.

Another example of a recategorization that does not necessarily involve directionality can be observed in Chinese, where Börjars & Payne (2021) argue that nouns which originally denoted a container or a measure and which had the syntactic distribution and modificational properties of standard nouns in the language were reanalyzed over time as measure words and classifiers. They argue that the reanalysis as measure words involves only a syntactic recategorization by which these nouns have a more restricted syntactic distribution and modificational properties in comparison to standard nouns. This syntactic change is not accompanied by a systematic semantic or functional change: the words still measure out units, as in the original usage and appear to retain their full lexical semantics.

van Gelderen (2011) seeks to address issues of recategorization by thinking of language change in terms of cycles (though there seems to be no discussion of actual full cycles of language change) and working heavily with features that are associated with syntactic categories. Changes in the features associated with an item eventually lead to reanalysis at the level of syntactic categories. Interestingly, this take on language change seems to move towards the separation out of functional vs. categorial characteristics of an item that is already in-built into LFG, but with a comparatively impoverished understanding of feature theory.

The difference in how many and what kinds of features and functionality are associated with an item can also lead to debates as to whether syntactic recategorization has taken place at all. Vincent & Börjars (2020: 144) discuss this with respect to prepositions being used to mark the subcategorized for arguments of a predicate, thus acquiring the properties of case markers and note: "Within Minimalism such shifts can be seen as involving a change from P to K, whereas once again, in HPSG and LFG, the change is in the information associated with the argument of P rather than in the category itself." This view of the relationship between adpositions and case clashes with the lexical semantic approach to case taken by Butt & King (1991, 2003, 2004), who use the category K to model the status of case markers in New Indo-Aryan as independent clitics that have a range of functional and lexical information associated with them. Ahmed Khan (2009) shows how spatial adpositions may be associated with feature structures specifying PATH and PLACE and how changes in the specification of these features can result in case markers. Butt & Ahmed Khan (2011) further chart the development of originally spatial adpositions into case markers, analyzed as K, in modern Urdu.

4.2 Mixed categories

Recategorization as described in Section 4.1 also generally does not happen in one fell swoop, but via a number of intermediate stages. One side-effect of these intermediate stages is the emergence of *mixed categories*. Verbal nouns or gerundives, which have the external distribution of nouns, but the internal properties of verbs, are one well-known example. A recent survey and analysis of mixed categories by Nikolaeva & Spencer (2020) shows that there are several different types of mixed categories (see also the discussion in Lowe 2016b).

In her diachronic take on mixed categories, Nikitina (2008) argues for a clear disassociation of the lexical and syntactic components of category mixing, precisely because of the range of mixed properties displayed by syntactic categories. She investigates and analyzes phenomena from Romance and Wan (Mande) and proposes that a clear distinction be made between instances of *function retaining* derivational changes in syntactic category and *structural reanalysis*, including rebracketing. An example of the former function retaining change is the English *-ing* nominalization, whereby the head distributes as a noun, but retains the functional predication of a verb. Over time, this retention of function may be lost, resulting in a straightforward nominal, rather than a syntactic category with mixed properties. Examples of the latter (structural reanalysis) include types of instances discussed in the previous section, i.e. the development of adpositions and case markers from nouns and the development of complementizers from verbs (Lord 1993) or adpositions.

The work by Vincent & Börjars (2010) on the slippery slope between adpositions and adjectives serves as another example. Vincent & Börjars (2010) look at Germanic and Romance languages which are losing their overall case system and investigate paths of change that serve to compensate for this loss. One path of change involves the use of adpositions like Latin *prope* 'near' and English 'near'. These subcategorize for an OBJ and have the spatial meaning of adpositions, but can be used as adjectives and take comparative and superlative morphology. See also Vincent & Börjars (2020) for some further discussion on this issue.

Formal analyses of mixed categories in LFG have often invoked *lexical sharing* of one type or another (Bresnan & Mugane (2006); Lowe (2015, 2016a); see discussion in Section 4.4). Butt et al. (2020) propose an alternative to this approach in their analysis of complement clauses in Tamil. The complement clauses show a mixed set of nominal and verbal properties, which is due to a historical development by which an originally relative clause type structure incorporated a pronoun and thus acquired the external properties of a noun while retaining the internal verbal predication of a finite clause. Butt et al. (2020) propose an analysis of mixed categories in terms of the formal tool of *complex categories* first introduced in the context of the ParGram computational grammar development effort (Butt et al. 1999; Crouch et al. 2008). This approach essentially allows for the parameterization of syntactic categories, avoiding the monolithic assignment of one syntactic category to a given lexical item or phrase.

4.3 The growth of structure

The diachronic phenomena discussed under the term "grammaticalization" in Section 3.1 involved a change whereby an individual lexical item comes to be reanalyzed as a functional element. This could be considered grammaticalization in the narrow sense, in which the focus is on the changing status of a particular item, as in much of the classic work on grammaticalization where changes occurring above the level of individual lexemes (e.g., changes in word order) are typically secondary concerns. For example, in their seminal textbook, Hopper & Traugott (2003: 24, 59) suggest that word order changes, though "deeply interconnected" with grammaticalization, are not to be considered under the term on the basis that they do not exhibit the unidirectionality typical of grammaticalization (see also Sun & Traugott 2011 for a similar view). At the same time, some authors have called for a broader take on grammaticalization, encompassing also cases where a particular fixed word order comes to encode certain functional information. This type of change is argued for by Börjars et al. (2016) as grammaticalization involving a "template" made up of slots and categories, one example being V1 clauses which have grammaticalized to varying degrees so as to encode conditionality across the Germanic languages (see also Hilpert 2010). Börjars et al. (2016) extend the remit of grammaticalization further still, proposing a specific type of grammaticalization, couched within LFG, which involves two concomitant developments: (i) the development of a grammaticalized meaning in a particular item and (ii) the increasing association of that grammaticalized meaning with a particular structural position.

Börjars et al. (2016) propose this special type of grammaticalization on the basis of diachronic data concerning the development of definite markers and noun phrase syntax in North Germanic, specifically from Old Norse to modern Faroese. They provide empirical evidence which shows that Old Norse lacks an obligatory dedicated (in)definiteness marker via paired examples such as (26), where the bold element in (26a) receives a definite interpretation and that in (26b) an indefinite interpretation, despite the fact that neither is overtly marked for (in)definiteness.

(26) Old Norse (Börjars et al. 2016: e10)

a. **Austmaðr** kvezk …
east.man said
'The Norwegian said …'

b. Ok gekk **kona** fyrir útibúrsdyrrin
and went woman in.front.of outhouse.door.DEF
'A woman went in front of the door of the outbuilding'

Moreover the relative order of elements within the noun phrase is relatively free, although the initial position is associated with prominent and contrastive elements, as in the two instances of prenominal possessive pronouns with contrastive emphasis in (27).

(27) Old Norse (Börjars et al. 2016: e14)
at **minn** **faðir** væri eptirbát **þins** **fǫður**
COMP 1.SG.POSS father was after.boat 2.SG.POSS.GEN father.GEN
'that my father trailed in the wake of yours'

In the next diachronic stage which Börjars et al. consider – early Faroese (ca. 1298 CE) – they provide evidence which indicates that overt marking of definiteness is now obligatory, since unmodified nouns must occur with a definiteness marker in order to receive a definite interpretation, e.g., (28a). Indefinite markers are not yet obligatory, however, as evidenced by examples like (28b), which is interpreted as indefinite without any overt indefinite marker.

(28) Early Faroese (Börjars et al. 2016: e18)

a. Bardr Peterson war ritade **brefet**
B.P. was written letter.DEF
'Barður Peterson had written the letter.'

b. Ef **sauþr** gengi j annars haga …
if sheep goes in other's field
'If a sheep goes into another man's field …'

Moreover, unlike in Old Norse, where different definiteness markers can co-occur, by this stage of Faroese they are in complementary distribution. Only later in the history of Faroese — the representatives which Börjars et al. examine are a newspaper from the 1890s and data from Present-day Faroese — does overt marking for indefiniteness become obligatory via *ein*, e.g., (29a), and a prenominal syntactic definiteness marker (*tann* or *hinn*) is generally required when there is premodification, leading to "double definiteness", e.g., (29b)–(29c).

(29) Present-day Faroese (Börjars et al. 2016: e22–e23)

a. ein ungur maður
INDEF young.STR man
'a young man'

b. tann stóra gatan
DEF big.WK mystery.DEF
'the/that big mystery'

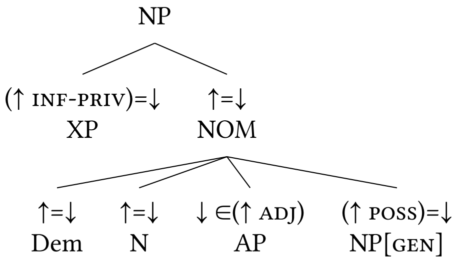
c. hin størsta vindmylluparkin í Europa
DEF biggest windmill.park.DEF in Europe
'the biggest wind farm in Europe'

In sum, the history of Faroese exhibits the grammaticalization of dedicated (in)definiteness markers which only later come to be associated with a particular structural position (the left edge of the noun phrase), in line with an overall increasingly fixed word order within the noun phrase.[8] Börjars et al. (2016) take advantage of the flexible nature of LFG's c-structure to model the observed gradual changes, building on a diachronic account of word order change in the context of grammaticalization concerning Romance prepositions by Vincent (1999). Central to both Vincent's (1999) and Börjars et al.'s (2016) accounts is the assumption that a new category can emerge diachronically without it necessarily needing to project a full endocentric phrase straightaway. Indeed, in the two accounts the full endocentric phrase projected by the new category is "grown" gradually at c-structure over time. This view is in line with the c-structure principle of Economy of Expression (Bresnan et al. 2016), which privileges lexical over phrasal expression and is radically different to the more standard universal application of X-bar theory in certain other generative approaches whereby, as soon as one posits a category, one also needs to posit a full endocentric phrasal projection complete with a specifier and complement position (cf. also Toivonen 2001, 2007 on non-projecting categories within LFG).

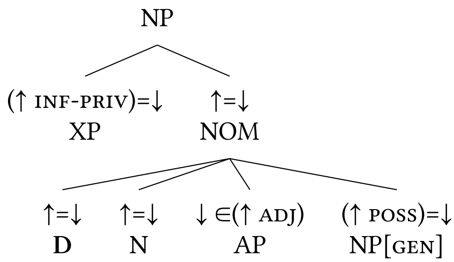
[8]Specifically, the definiteness marker was originally associated with the adjective in Old Norse, and frequently occurred postnominally, to the left of the adjective. The proposal by Börjars et al. is that, as adjectives became increasingly prenominal, the definiteness marker became associated with the left edge of the noun phrase overall.

Applied to the Faroese story, Börjars et al. (2016) propose three c-structures to capture the structure of nominal phrases in the three periods: Old Norse (30a), early Faroese (ca. 1298) (30b) and Present-day Faroese (30c). In the earliest structure in (30a), word order is largely free (captured in the flat structure under NOM(inal)) but there is an initial position associated with information-structurally privileged elements. Crucially, in (30a) there is no category D; this only develops in early Faroese, cf. (30b), but at this point it is not yet associated with a particular structural position. Once definite markers are structurally associated with the left edge of the noun phrase, one can assume a projectional functional category, as captured in the endocentric DP structure for Present-day Faroese in (30c). The proposed growth of c-structure thus captures the grammaticalization of (in)definiteness in the context of a gradual shift from relatively free word order driven by information structure to a much more rigid, syntactically constrained word order as exhibited in modern Faroese.

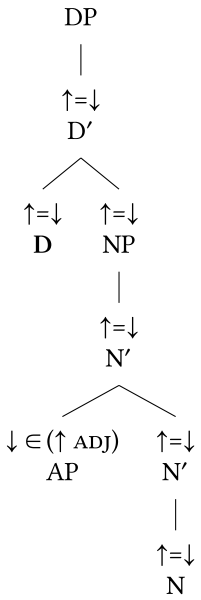
(30) a. Old Norse



b. Early Faroese



c. Present-day Faroese

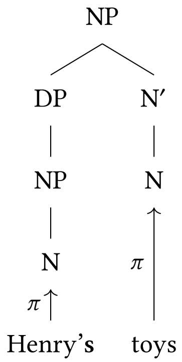


4.4 Degrammaticalization and lexical sharing

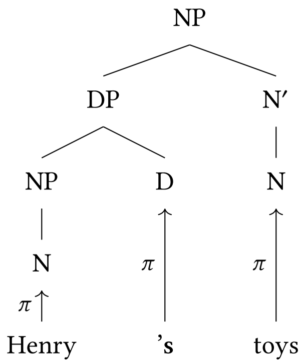
Another way in which gradual degrees of syntactic change have been captured at c-structure is via "Lexical Sharing", as originally proposed within LFG by Wescoat (2002, 2005, 2007, 2009) and further developed by Lowe (2015, 2016a) as "Constrained Lexical Sharing". Lexical Sharing is essentially a mechanism which allows two or more constituents at c-structure to map to a single lexical element. As we discuss in this section, Lowe (2015) employs this in the context of diachrony as a way to model degrees of "*de*grammaticalization" (see e.g., Norde 2009; Willis 2017) with respect to the English possessive marker *'s*, building on a synchronic analysis of Present-day English in Lowe (2016a).

The starting point for Lowe's account is Present-day English, which indicates a mixed picture with respect to whether the possessive marker *'s* has clitic or affixal status. This is reflected in the literature, where some argue for it to be a clitic (e.g. Quirk et al. 1985; Anderson 2008) and others for it to be an affix (e.g., Zwicky 1987; Payne 2009). In this context, Lowe (2016a) claims that synchronically *'s* has dual status, i.e. that Present-day English exhibits both clitic forms and affixal forms and shows how this complex status can be modelled via Constrained Lexical Sharing. As Lowe points out, the lexicalism underpinning LFG leads to a discrete distinction between clitic and affix. An affix is assumed to attach to its host in the lexicon and will thus map to the same c-structure node as its host, e.g., (31a), while a clitic is a distinct lexical element which occupies its own c-structure node, e.g., (31b). The example c-structures here are as in Lowe (2015: 213).[9]

(31) a. Affix:



b. Clitic:

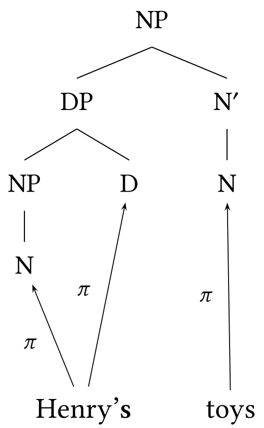


Rather than assume the straightforward affixal analysis in (31a) for affix-like instances of *'s*, however, Lowe (2016a) proposes an account involving (Constrained) Lexical Sharing. This allows one to capture the affixal status of *'s* whilst being

[9]In both Lowe (2015: 213) and Lowe (2016a: 174), different structures are provided for the possessum *toys* depending on whether one assumes the affix analysis or the clitic/lexically shared affix analysis; for the affix analysis, the immediate daughter of N' is N, but for the clitic and lexically shared analysis the immediate daughter of N' is given as an NP. Although Lowe does not provide any explanation for this difference, for the sake of consistency we simply repeat the structures here.

able to maintain a consistent syntactic structure for possessive phrases, irrespective of the affixal/clitic status of *'s*. Wescoat's original formulation of Lexical Sharing (Wescoat 2002, 2005, 2007, 2009) assumes an additional dimension, l(exical)-structure, which consists of a linearly ordered set of words. The idea behind Lexical Sharing is that it is possible for two adjacent c-structure elements to map to a single element at l-structure, i.e. "sharing" the same lexical exponent. Within the Constrained Lexical Sharing of Lowe (2015, 2016a), Wescoat's l(exical)-structure is identified with the syntactic string (s-string) of Kaplan (1989), and thus Lexical Sharing refers to instances where a single element at the s-string is associated via the relation π with two adjacent c-structure nodes. In Lowe's account, the affix-like *'s* is a lexically shared affix, e.g., (32), i.e. constitutes a single lexical element with its host but maps to a separate node from the host at c-structure, resulting in an overall structural configuration parallel to that for the clitic analysis in (31b).

(32) Lexically shared affix:



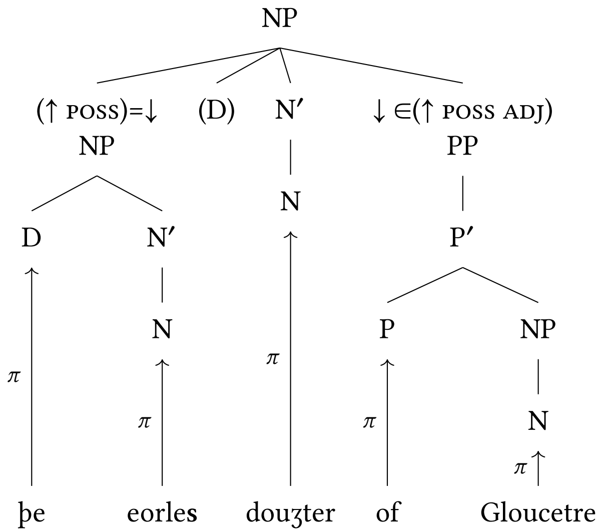
Applied specifically to diachrony, Lowe (2015) shows how this approach can be used to represent degrees of "*de*grammaticalization" of the English possessive marker, which "*de*grammaticalizes" over time from an unambiguous affix to a clitic (cf. the typical grammaticalization cline in (1) above). At the earliest attested stage, the Old English ancestor of Present-day English *'s*, *-es*, is one of a number of genitive case allomorphs and is uncontroversially an affix which is fully integrated with its stem in the s-string and maps to the same c-structure node as its host, cf. (31a). Crucially, as Lowe (2015) points out, drawing on data discussed by Allen (1997, 2003, 2008), the emergence of the clitic over the subsequent centuries is gradual and involves degrees of degrammaticalization and small-step changes from affix to clitic.

Specifically, in the period ca. 1100–1400 CE several changes are underway which affect the affixal status of the possessive marker: the various genitive case forms are largely lost and *-(e)s* becomes the possessive marker for most nouns, while possession is increasingly marked on just the head of the possessor, rather than on every element of the possessor. Lowe (2015) cites two construction types in particular which are attested during this period and indicate the beginning of a change in the morphosyntactic status of the possessive marker: (i) possessors which involve coordination where possession is marked only on the rightmost head, e.g., (33a) and (ii) possessor phrases with split postmodification flanking the possessum, where possession is marked on the head of the possessor, e.g., (33b).

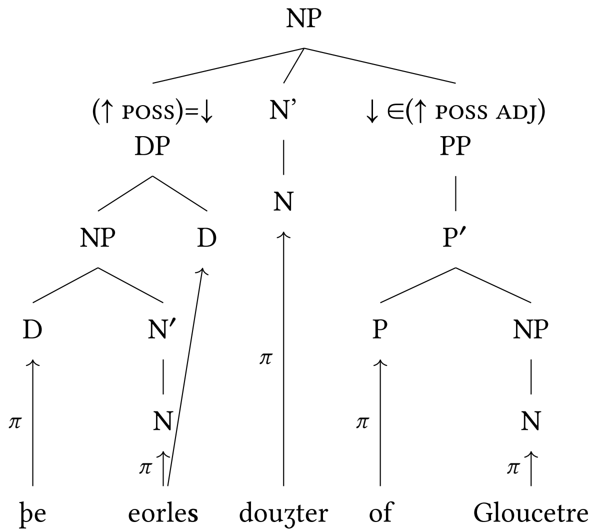
(33) Middle English
  a. wif & weres gederunge
     wife and man.GEN union
     'The union of man and wife.'
     (*Hali Meidenhad*, c. 1225 CE)
  b. þe eorles douȝter of Gloucetre
     the earl.GEN daughter of Gloucester
     'The Earl of Gloucester's daughter'
     (*Polychronicon* VIII, ca. 1380)
     (Lowe 2015: 217–218)

The two constructions in (33) show a strong positional constraint, whereby the possessive marker on the head of the possessor phrase must immediately precede the possessum. Lowe interprets this as evidence that the possessive marker is no longer fully affixal, since it is now constrained by the syntactic context in which it appears, rather than just being dependent on the position of the word to which it attaches. In the Lexical Sharing approach, this change can be captured by assuming that the original affix (modelled in (34a)) is reanalyzed as a lexically shared affix as shown in (34b) (structures from Lowe 2015: 222). In (34a), an optional D node is assumed, which is later incorporated into the head of the possessor phrase (see (34b)), once possessor phrases come to supply the definiteness of the possessum, in line with the broader grammaticalization of the definite article which is underway in the period.

(34) a.



b.



However, the Middle English lexically shared affix in (34b) is not yet equivalent to the status of the possessive marker in Present-day English but rather has a subtle difference. As Lowe (2015) points out, in the former, the noun component must supply the head of the possessor phrase. This is captured in the lexical entry in (35a), where the N component is associated with an f-description which requires that the f-structure of the noun serve as the value (or a member of a set of values) of the feature POSS in a wider f-structure; the D does not need to have any f-descriptions associated with it. By contrast, for Present-day English – where the possessive marker is closer to a clitic – Lowe assumes the lexical entry in (35b), "Partitioned Lexical Sharing", which involves two c-structure nodes each with their own set of f-descriptions (structures from Lowe 2015: 215, 223).

(35) a. Unified Lexical Sharing:

eorles:	N	D
	(↑ PRED) = 'earl' (POSS (ϵ) ↑) ((POSS ↑) DEF) = +	

b. Partitioned Lexical Sharing:

species':	N	D
	(↑ PRED) = 'species'	(POSS ↑)

According to Lowe, only from the end of the 14th century are "phrasal possessives proper" attested, i.e. phrasal possessives with postmodified possessors with possessive marking on the right edge of the postmodifier, rather than on the head of the possessor as in early examples, cf. (33b). The example Lowe provides is from Chaucer (ca. 1400 CE), shown here in (36). These more clitic-like examples coexist alongside the more affix-like split examples as in (33b) throughout the Middle English and Early Modern English periods, with an increasing preference for the more clitic-like type in (36). This is modelled in terms of gradually shifting preferences in favour of the Partitioned Lexical Sharing analysis, cf. (35b) over the Unified Lexical Sharing analysis, cf. (35a).

(36) Middle English (Lowe 2015: 219)
The grete god of Loves name
'The great God of Love's name.'

Taking advantage of Lexical Sharing thus allows Lowe (2015) to model the nuanced steps involved in syntactic change via degrammaticalization and, applied to Present-day English, also to capture the coexistence of older and newer variants at a synchronic level (Lowe 2016a).

A very different approach to the mix of clitic and affixal properties is exemplified by Bögel & Butt (2012), who work out an analysis of various different Urdu possessive constructions, including case clitics and the originally Persian *ezafe* construction. Their analysis factors in prosodic features typical of clitics (rather than leaving them out) and avoids the introduction of lexical sharing or other complex formal machinery.

4.5 Grammatical relations and licensing

As Vincent (2001) points out, the fundamental design of the LFG architecture, in which position and function are captured separately, means that it is well suited to modelling changes concerning grammatical relations. A consequence of assuming f-structure as an independent level of representation for abstract functional information is that grammatical functions such as SUBJ and OBJ are viewed as basic building blocks of the theory. As such, unlike in some other generative approaches, SUBJ and OBJ need not be defined in terms of structural position. This allows one to neatly capture the full cross-linguistic spectrum with respect to how languages encode grammatical relations, from those where structural position plays a strong role, e.g., modern English, to those where morphological marking is the dominant encoding means, e.g., Latin, but also languages which use a mixture of means, e.g., Chicheŵa (Bresnan & Mchombo 1987) and Icelandic (Zaenen et al. 1985). Previous work in this area has tended to focus on the cross-linguistic possibilities from a synchronic perspective; see Nordlinger (1998) and Snijders (2015) for relevant typologies. But this approach to grammatical relations also has much to offer for diachronic studies, since change concerning how languages encode grammatical relations is well-attested across languages (e.g., Kiparsky 1997; Hewson & Bubenik 2006; Ponti & Luraghi 2018) and individual historical stages will naturally exhibit intermediate stages along a change trajectory, with a particular balance between structural and morphological encoding strategies.

4.5.1 Word order and recipient passives

A complex change in this area which has been investigated in detail by Allen (1995, 2001) is the rise of the recipient passive in English, i.e. constructions like (37), where the recipient rather than the theme is treated as the subject.

(37) He was given a book.

According to Allen (2001), recipient passives are unattested in Old English; the earliest known example of a recipient passive is from 1375 CE, alongside other scarce examples from the late fourteenth and early fifteenth centuries. The earliest example according to Allen, from 1375, is shown here in (38).

(38) Middle English (Award Blount, p. 207, Allen 2001: 51)
Item as for the Parke, she is alowyd Every yere a dere and xx Coupull of Conyes and all fewell Wode to her necessarye…

'Item: as for the park, she is allowed a deer every year and 20 pairs of rabbits and all firewood necessary to her...'

As with the change concerning experiencer verbs discussed in Section 3.2, Allen again challenges classic accounts of this development (e.g., Jespersen 1927; van der Gaaf 1929; Campbell 1998), which assume that recipient passives emerged via reanalysis of an ambiguously case-marked fronted indirect object as the subject, as in (39).

(39) Middle English (Ric.Couer de L. 1307, Auchinleck MS, Allen 2001: 49)
The Duke Mylon was geven hys liff, and fleygh out of land with his wife.
'Duke Mylon was given his life, and fled out of the country with his wife.'

Allen (2001) points out that the chronology does not stack up to support the classic account, for a variety of reasons. This includes the observation that the loss of the morphological distinction between nominative and dative which results in the prerequisite ambiguity for the classic reanalysis account occurred long before the first recipient passive examples are attested, with a gap of 175 years.

Allen (2001) argues instead for a change involving reanalysis of the indirect object (theme) of active sentences as the direct object, which in turn has consequences for the status of the recipient argument and ultimately facilitates its promotion to subject under passivization. Rather than being driven by ambiguous case-marking, as assumed in the classic accounts, Allen argues that her reanalysis story was triggered by the fixing of the relative word order of two objects. This is based on the observation that the first attestation of recipient passives coincides with the disappearance of examples like (40), in which a (non-pronominal) NP which is the Theme precedes a (non-pronominal) NP which is the Recipient. According to Allen, such orderings with two NPs are unattested as of the last quarter of the fourteenth century.

(40) I gave [a gift]$_{\text{theme}}$ [the king]$_{\text{recipient}}$

Allen suggests that once nominal recipients became fixed in the immediately postverbal position, the simplest analysis from the perspective of the language learner was to analyze the recipient as an OBJ, due to the fact that the learner could now calculate grammatical relations directly on the basis of position, and in turn the semantic relations too. Specifically, a new processing strategy arose which stated that the first non-pronominal NP after the verb would be the OBJ, provided no pronoun preceded it – the strategy for pronouns would be rather different according to Allen, presumably owing to the special positional distribution of pronouns in Early English. As a result, the semantic role could now be determined on the basis of position: if the OBJ was followed by another NP, the OBJ could only be the Recipient and the second NP could only be OBJ_θ, with the only possibility in terms of semantic role as the Theme. In this way, the fixing of the order eases the hearer's processing concerning the assignment of grammatical relations and thematic roles. Moreover, since the Recipient as an OBJ is now −r, it can map to SUBJ under passivization in accordance with the natural classes which fall out from the features [±r,o]. Thus recipient passives are now possible.

Allen's account thus investigates the connection between word order and the assignment of grammatical relations and in particular a change whereby grammatical relations become increasingly encoded via position. Next, we discuss other work which has considered this type of change within LFG.

4.5.2 Positional licensing and information structure

Change whereby structural position becomes an increasingly dominant licensing strategy for grammatical relations over time is well attested cross-linguistically; cf. the rise of (argument) configurationality (e.g., Hewson & Bubenik 2006; Luraghi 2010). In the linking theory developed by Kiparsky (1987, 1988, 1997, 2001), where case, agreement and position are viewed as interacting licensing strategies for grammatical relations, this type of change has been formalized as the rise of "positional licensing" (Kiparsky 1997). Focusing on the history of English, Kiparsky (1997) argues that as English lost its morphological case system, position became the dominant licensing strategy for grammatical relations. Beyond Allen's analysis of the recipient passive, this idea has been explored more recently within LFG by Booth et al. (2017), who present a positional licensing account in LFG terms for the diachrony of subjects in Icelandic. As both Kiparsky (1997) and Booth et al. (2017) point out, Icelandic offers an interesting point of comparison with Kiparsky's original account since, unlike English, Icelandic has maintained rich morphological case up to the present day.

Booth et al. (2017) observe that Icelandic subjects are increasingly realized in the clause-initial prefinite position and capture this change in terms of the rise of positional licensing, also bringing information structure into the account. Two other concomitant changes are observed and feed into their analysis: (i) a decrease in V1 declaratives as in (41) and (ii) the emergence of the expletive *það* which is positionally restricted to the clause-initial prefinite position; cf. the contrast in (42), when the expletive is ruled out in contexts where this position is otherwise occupied.

(41) Middle Icelandic (Georgius, 1525, Booth et al. 2017: 111)

Sýndi drottinn mikla miskunn vin sínum
show.PST.3SG lord.NOM.DEF great.ACC mercy.ACC friend.DAT his-own.DAT
sankti Georgíum
saint.DAT George.DAT
'The Lord showed great mercy to his friend St. George.'

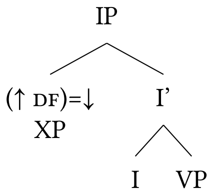
(42) Modern Icelandic (Booth et al. 2017: 111–112)

a. **Það** var ekki minnst á önnur dýr.
EXPL be.PST.3SG NEG mention.PTCP on other.ACC animals.ACC

b. Ekki var minnst á önnur dýr.
NEG be.PST.3SG mention.PTCP on other.ACC animals.ACC
'There was no mention of other animals.'

Following a proposal by Hinterhölzl & Petrova (2010) for the history of West Germanic, Booth et al. (2017) put forward an information-structural account for the rise of the expletive *það* and in turn the decrease in V1 declaratives, assuming that the finite verb serves as an information-structural boundary separating topic and comment. The change is captured in terms of the growth of structure, whereby a flat structure lacking functional categories yields to a more articulated structure making use of functional categories and projections, similar to the account of North Germanic noun phrases by Börjars et al. (2016) (see Section 4.3). In Booth et al.'s account at the clausal level, the relevant functional projection which emerges from an earlier flat structure is IP, headed by the finite verb in I; cf. the LFG accounts of modern Scandinavian clause structure by Sells (2001, 2005) and Börjars et al. (2003).

Once the IP structure is established, various changes occur concerning the nature of the clause-initial prefinite position, i.e. SpecIP. The information-structural role of the finite verb as a boundary between topic and comment leads to SpecIP becoming increasingly associated with a discourse function (DF) capturing given or topical information, cf. (43).

(43)


In their account, this increasing association of SpecIP with topicality can in turn explain the observed increasing realization of subjects in this position. Agentive and sentient entities tend to make for better topics than non-sentient entities, and, since subjects typically represent the more agentive, sentient semantic participants, subjects will accordingly often occur in this new topic position (see also Givón (1990), who discusses subjects as "grammaticalized topics").

However, as Booth et al. show, SpecIP does not straightforwardly develop into a designated subject position, since subjects can still occur postfinitely in modern Icelandic. In particular, in clauses which lack a topic altogether, i.e. impersonal and presentational constructions, the expletive occurs in the SpecIP topic position as a signaller that the clause lacks a topic. As such, they propose the functional annotations in (44) for SpecIP in the modern stage: it can be occupied by any topical GF, or alternatively the expletive, provided the clause lacks a topic.

(44)

IP

$\left\{ \begin{array}{l} (\uparrow \text{TOPIC}) = \downarrow \\ (\uparrow \{\text{COMP}|\text{XCOMP}\}^{*}\ \text{GF}) = \downarrow \end{array} \right.$

$\left| \begin{array}{l} (\uparrow \text{EXPLETIVE}) =_c + \\ \neg(\uparrow \text{TOPIC}) \end{array} \right\}$

XP

I'

I …

Thus the reorganization of information structure and word order in Icelandic, and in particular the changing status of SpecIP, is seen as the underlying shift which results in a decrease in V1 in favour of V2 sentences with a clause-initial topic or an expletive in topicless contexts. Booth et al.'s study shows that information-structural properties are an important consideration in the context of change with respect to word order and the licensing of grammatical relations. For a similar account which presents this change in terms of shifting correspondences between c-structure, f-structure and LFG's i(nformation)-structure (Butt & King 1996; King 1997), see Booth & Schätzle (2019) and Booth & Beck (2021). In a similar vein, Booth (2021) shows how assuming that languages can gradually change their status with respect to argument configurationality and discourse configurationality can account for subtle changes in word order between Old and Modern Icelandic, which have otherwise prompted heated debate within approaches which assume configurationality to be a binary parameter (e.g. Faarlund 1990; Rögnvaldsson 1995).

5 Variation and change

In this last section, we turn to the question of the role of variation. It is well-known that language change is gradual and goes hand-in-hand with variation (e.g., Weinreich et al. 1968; Kroch 1989; Labov 1994, 2001; Pintzuk 2003; Chambers & Schilling 2013). However, formal grammars are discrete in nature, so a natural question which arises is how to combine the inherent variability and gradualness associated with language change into formal models of grammar.

One very popular way forward has been the combination of *Optimality Theory* (Boersma et al. 2000; Kager 1999) with stochastic methods (Boersma 2000; Boersma & Hayes 2001), which has been argued to account for patterns of variation and language acquisition. Optimality Theory (OT) was adapted into LFG very early on to yield a version of LFG dubbed "OT-LFG" (Kuhn forthcoming [this volume], Bresnan 1996, 1998) and combined with stochastic methods to yield explanations for gradience in judgements (Bresnan et al. 2007; Bresnan & Nikitina 2010) and variation across dialects (Bresnan 2007; Bresnan & Hay 2008; Sharma et al. 2008; Bresnan et al. 2001).

In terms of historical linguistics, Clark (2004) is the first to lay out a formal model of diachronic variation that has led to gradual change.[10] He works with two case studies from the history of English: 1) word order change from primarily OV to VO; 2) the preferred association of subjects with the clause-initial position. Clark models the observed changes within stochastic OT-LFG and shows how the model parallels the observed stages in the historical development and variation in the corpora. Change is essentially effected via competing variants, as in much influential work on syntactic change in recent decades (e.g., Kroch 1989; Pintzuk 2003). In Clark's OT-LFG account, these competing variants are taken to be the result of constraints that are liable to be re-ranked with respect to one another due to inherent "noise" in the communication process between humans and the asymmetry of goals between perception (more information is generally useful for decoding) vs. production (producing less information is generally less burdensome). The stochastic OT-LFG approach is able to capture the steady quantitative rise in the use of an innovated structure by associating it with gradual changes in the relative strength of the relevant constraints. A constraint re-ranking that may be due to "noise" variation may become statistically preferred and from there finally lead to a categorical change.

[10] Note that 'gradual' in this context refers not to incremental steps along, e.g., a grammaticalization cline as in (1) but rather the gradual diffusion of a particular change through a population of speakers, or even possibly the gradual establishment of an innovation in the grammar of an individual.

In a comprehensive look at auxiliary contraction in English, Bresnan (2021) proposes a new hybrid model of LFG and a usage-based mental lexicon to explain the synchronic distribution and diachronic development of auxiliary contractions/clitics. The usage-based mental lexicon is conceived of as a combination of ideas coming from Pierrehumbert's examplar-based model of the mental lexicon (Pierrehumbert 2001, 2002) and Bybee's usage-based approach (Bybee 2006). The frequent co-occurrence of certain combinations of words, e.g., *you+are* or *we+will*, are predicted to undergo contraction. Frequencies are calculated using a measure termed *Informativeness*, which is the logarithm of the inverse of the conditional probability of one word following the other. The corpus study, the frequency calculations and the development of the hybrid LFG plus usage-based mental lexicon make for a rich and complex paper which combines the strengths of formal grammar modelling with by now well-established effects of frequency and variation in usage.

The auxiliary contractions themselves are modelled via the formal concept of lexical sharing within LFG (cf. Section 4.4) and this is where Bresnan (2021) draws a concrete connection to diachrony: "lexical sharing as a formal construct can be viewed as a grammaticalization of high-probability syntactic distributions in usage…". However, as argued for by Bögel (2015), clitics do not necessarily need to be modelled via lexical sharing. She instead proposes a more articulated architecture of the prosody-syntax interface for an analysis of clitics and provides a means for integrating effects of frequency and variation (Bögel & Turk 2019). Frequency effects can also be modelled via preferences applied directly to rules or lexical items, as practiced with respect to computational grammar development in LFG (Frank et al. 2001; Dost & King 2009).

Overall, the area of variation and change provides an interesting area of research for LFG, with initial architecturally complex and sophisticated proposals having recently been formulated, pointing the path towards innovative and exciting research in this area.

6 Summary

This chapter has endeavored to provide an overview of the types of work done within LFG on language change. Like most formal theories of grammar, LFG did not address language change from the outset and the first serious work on language change only began appearing in the 1990s. However, as LFG is fundamentally designed to separate out form from function, the complex interaction between the function of an item and its overt realization can be modelled very well. Indeed, given LFG's complex formal design, in which the different components of grammar are represented with component-appropriate rules and representations, and interface with one another via the projection architecture, one might argue that there are too many moving parts to permit a clearly delineated theory of language change.

In one sense this is correct, but in another sense, one could argue that, as with synchronic description, what LFG provides is a broad formal framework, which must be specified by linguistic theorizing. By its very nature LFG pursues an *inductive* approach — the framework provides a broad perspective on the data (e.g., form and function are assumed to be separate, c-structure is assumed to model linear order, constituency and hierarchical relations, grammatical relations and argument structure are core objects over which generalizations can be stated, etc.) but the linguistic explanations and generalizations themselves emerge from the data and can be stated independently of the theory, thus allowing for potential cross-theoretic validity. Another aspect of the inductive approach is that the overall framework can be adjusted in the face of strong empirical evidence. For instance, when there is strong empirical evidence that information about frequency of items plays a role, the theory is adjusted and opens up interesting new ways of modelling language change, as we saw in Section 5.

Furthermore, the fact that LFG is very functionally oriented allows for an open channel of communication with the functional-descriptive and grammaticalization literature, leading to natural and insightful accounts of lexical and functional change, as we discussed in Section 3. Change in terms of syntactic categories and clausal organization is seen as being motivated by changes in function in the first place, with syntactic recategorization and reorganization following to reflect the change in underlying function (Section 4).

In this chapter, we have followed the very broad notion of reanalysis adopted by Harris & Campbell (1995) and hope to have shown how LFG can naturally account for reanalysis at various different levels: lexical, functional, categorical. In fact, one could see LFG as providing a firm formal basis for understanding the possible moving parts involved in reanalysis as part of language change, while also providing a basis for understanding how the well-attested gradualness, variation and frequency-based effects of language change can be modelled formally. That said, and despite the length of the chapter, it should be obvious to the reader that the existing body of historical linguistic work within LFG is so far not large and there is thus much room for investigation into language change. There is also room for experimentation and innovation with respect to how to represent and understand language change. This might include a new model of the lexicon, a new version of LFG or the introduction of new methods of probabilistic modelling, as we have seen above, or working with new ways of accessing the diachronic data, for example by means of a platform developed together with experts from visual analytics (e.g., Schätzle et al. 2019; Beck & Butt 2020; Beck et al. 2020).

References

Ahmed Khan, Tafseer. 2009. *Spatial expressions and case in South Asian languages*. Konstanz: University of Konstanz. (Doctoral dissertation).

Allen, Cynthia L. 1986. Reconsidering the history of "like". *Journal of Linguistics* 22(2). 375–409. DOI: 10.1017/s0022226700010847.

Allen, Cynthia L. 1995. *Case marking and reanalysis: Grammatical relations from Old to Early Modern English*. Oxford: Oxford University Press.

Allen, Cynthia L. 1997. The origins of the 'Group Genitive' in English. *Transactions of the Philological Society* 95(1). 111–131. DOI: 10.1111/1467-968x.00015.

Allen, Cynthia L. 2001. The development of a new passive in English. In Miriam Butt & Tracy Holloway King (eds.), *Time over matter: Diachronic perspectives on morphosyntax*, 43–72. Stanford: CSLI Publications.

Allen, Cynthia L. 2003. Deflexion and the development of the genitive in English. *English Language & Linguistics* 7(1). 1–28.

Allen, Cynthia L. 2008. *Genitives in early English: Typology and evidence*. Oxford: Oxford University Press.

Allen, Cynthia L. 2012. New perspectives, theories and methods: Lexical Functional Grammar. In Alexander Bergs & Laurel Brinton (eds.), *English historical linguistics: An international handbook*, vol. 2, 1646–1658. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110251609.1646.

Almashaqba, Bassil Mohammad. 2015. *The phonology and morphology of Wadi Ramm Arabic*. Salford: University of Salford. (Doctoral dissertation).

Anderson, Stephen R. 2008. The English "group genitive" is a special clitic. *English Linguistics* 25(1). 1–20.

Barðdal, Jóhanna. 2011. The rise of dative substitution in the history of Icelandic: A diachronic Construction Grammar account. *Lingua* 121(1). 60–79. DOI: 10.1016/j.lingua.2010.07.007.

Barron, Julia. 2001. Perception and raising verbs: Synchronic and diachronic relationships. In Miriam Butt & Tracy Holloway King (eds.), *Time over matter: Diachronic perspectives on morphosyntax*, 73–103. Stanford: CSLI Publications.

Beck, Christin, Hannah Booth, Mennatallah El-Assady & Miriam Butt. 2020. Representation problems in linguistic annotations: Ambiguity, variation, uncertainty, error and bias. In *Proceedings of the 14th Linguistic Annotation Workshop*, 60–73. Barcelona: Association for Computational Linguistics. https://www.aclweb.org/anthology/2020.law-1.6.

Beck, Christin & Miriam Butt. 2020. Visual analytics for historical linguistics: Opportunities and challenges. *Journal of Data Mining & Digital Humanities: Special Issue on Visualisations in Historical Linguistics*. 1–23. DOI: 10.46298/jdmdh.6707.

Beck, Christin & Miriam Butt. 2021. The rise of dative subjects: Relative prominence in event structure. Manuscript under review.

Boersma, Paul. 2000. Learning a grammar in functional phonology. In Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer (eds.), *Optimality Theory: Phonology, syntax, and acquisition*, 465–523. Oxford: Oxford University Press.

Boersma, Paul, Joost Dekkers & Jeroen van de Weijer. 2000. Introduction — Optimality Theory: Phonology, syntax, and acquisition. In Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer (eds.), *Optimality Theory: Phonology, syntax, and acquisition*. Oxford: Clarendon Press.

Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32(1). 45–86. DOI: 10.1162/002438901554586.

Bögel, Tina. 2015. *The syntax-prosody interface in Lexical Functional Grammar*. Konstanz: University of Konstanz. (Doctoral dissertation).

Bögel, Tina & Miriam Butt. 2012. Possessive clitics and *ezafe* in Urdu. In Kersti Börjars, David Denison & Alan Scott (eds.), *Morphosyntactic categories and the expression of possession* (Linguistik Aktuell/Linguistics Today 199), 291–322. Amsterdam: John Benjamins. DOI: 10.1075/la.199.11bog.

Bögel, Tina & Alice Turk. 2019. Frequency effects and prosodic boundary strength. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (eds.), *Proceedings of the International Congress of Phonetic Sciences (ICPhS)*, 1014–1018.

Booth, Hannah. 2021. Revisiting the configurationality issue in Old Icelandic. *Glossa: a journal of general linguistics* 6(1). DOI: 10.16995/glossa.5804.

Booth, Hannah & Christin Beck. 2021. Verb-second and verb-first in the history of Icelandic. *Journal of Historical Syntax* 5(28). 1–53.

Booth, Hannah & Miriam Butt. Forthcoming. LFG and historical linguistics. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Booth, Hannah & Christin Schätzle. 2019. The syntactic encoding of information structure in the history of Icelandic. In Miriam Butt, Tracy Holloway King & Ida Toivonen (eds.), *Proceedings of the LFG '19 conference*, 69–89. Stanford: CSLI Publications.

Booth, Hannah, Christin Schätzle, Kersti Börjars & Miriam Butt. 2017. Dative subjects and the rise of positional licensing in Icelandic. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '17 conference*, 104–112. Stanford: CSLI Publications.

Börjars, Kersti & Kate Burridge. 2011. From preposition to purposive to infinitival marker: the Pennsylvania German *fer... zu* construction. In Michael T. Putnam (ed.), *Studies on German-language islands*, 385–411. Amsterdam: John Benjamins.

Börjars, Kersti, Elisabet Engdahl & Maia Andréasson. 2003. Subject and object positions in Swedish. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '03 conference*, 43–58. Stanford: CSLI Publications.

Börjars, Kersti, Pauline Harries & Nigel Vincent. 2016. Growing syntax: The development of a DP in Northern Germanic. *Language* 92. e1–37. DOI: 10.1353/lan.2016.0002.

Börjars, Kersti & John R. Payne. 2021. Decategorialization and Chinese nouns. In I Wayan Arka, Ash Asudeh & Tracy Holloway King (eds.), *Modular design of grammar: Linguistics on the edge*. Oxford: Oxford University Press. DOI: 10.1093/oso/9780192844842.003.0014.

Börjars, Kersti & Nigel Vincent. 2017. Lexical-Functional Grammar. In Adam Ledgeway & Ian Roberts (eds.), *The Cambridge handbook of historical syntax*, 642–663. Cambridge, UK: Cambridge University Press.

Börjars, Kersti & Nigel Vincent. 2019. Modelling step change: The history of WILL-verbs in Germanic. In Nuria Yañez-Bouza, Willem Hollmann, Emma Moore & Linda van Bergen (eds.), *Categories, constructions and change in english syntax*, 283–314. Cambridge, UK: Cambridge University Press. DOI: 10.1017/9781108303576.013.

Börjars, Kersti, Nigel Vincent & Carol Chapman. 1997. Paradigms, periphrases and pronominal inflection: A feature-based account. In Gert Booij & Jaap van Marle (eds.), *Yearbook of morphology*, 155–180. Dordrecht: Kluwer Academic Publishers. DOI: 10.1007/978-94-017-3718-0_10.

Bresnan, Joan (ed.). 1982. *The mental representation of grammatical relations*. Cambridge, MA: The MIT Press.

Bresnan, Joan. 1996. LFG in an OT setting: Modelling competition and economy. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '96 conference*. Stanford: CSLI Publications. Handout from keynote address.

Bresnan, Joan. 1998. Morphology competes with syntax: Explaining typological variation in weak crossover effects. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky (eds.), *Is the best good enough? Optimality and competition in syntax*, 59–92. Cambridge, MA: The MIT Press & MITWPL.

Bresnan, Joan. 2007. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In Sam Featherston & Wolfgang Sternefeld (eds.), *Roots: Linguistics in search of its evidential base* (Studies in Generative Grammar), 77–96. Berlin: Mouton de Gruyter.

Bresnan, Joan. 2021. Formal grammar, usage probabilities, and auxiliary contraction. *Language* 97(1). 108–150. DOI: 10.1353/lan.0.0249.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-Functional Syntax*. 2nd edn. (Blackwell Textbooks in Linguistics 16). Malden, MA: Wiley-Blackwell.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & R. Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Bouma, Irene Krämer & Joost Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. KNAW.

Bresnan, Joan, Shipra Dingare & Christopher Manning. 2001. Soft constraints mirror hard constraints: Voice and person in English and Lummi. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '01 conference*, 13–32. Stanford: CSLI Publications.

Bresnan, Joan & Jennifer Hay. 2008. Gradient grammar: An effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua* 118. 245–259. DOI: 10.1016/j.lingua.2007.02.007.

Bresnan, Joan & Jonni M. Kanerva. 1989. Locative inversion in Chicheŵa: A case study of factorization in grammar. *Linguistic Inquiry* 20(1). 1–50.

Bresnan, Joan & Sam A. Mchombo. 1987. Topic, pronoun, and agreement in Chicheŵa. *Language* 63(4). 741–782. DOI: 10.2307/415717.

Bresnan, Joan & John Mugane. 2006. Agentive nominalizations in Gĩkũyũ and the theory of mixed categories. In Miriam Butt, Mary Dalrymple & Tracy Holloway King (eds.), *Intelligent linguistic architectures: Variations on themes by Ronald M. Kaplan*, 201–234. Stanford: CSLI Publications.

Bresnan, Joan & Tatiana Nikitina. 2010. The gradience of the dative alternation. In Linda Ann Uyechi & Lian-Hee Wee (eds.), *Reality exploration and discovery: Pattern interaction in language and life* (CSLI Lecture Notes 197), 161–184. Stanford: CSLI Publications.

Butt, Miriam. 2010. The light verb jungle: Still hacking away. In Mengistu Amberber, Brett Baker & Mark Harvey (eds.), *Complex predicates: Cross-linguistic* *perspectives on event structure*, 48–78. Cambridge, UK: Cambridge University Press.

Butt, Miriam. Forthcoming. Case. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Butt, Miriam & Tafseer Ahmed Khan. 2011. The redevelopment of Indo-Aryan case systems from a lexical semantic perspective. *Morphology* 21(3). 545–572.

Butt, Miriam & Tracy Holloway King. 1991. Semantic case in Urdu. In Lise M. Dobrin, Lynn Nichols & Rosa M. Rodriguez (eds.), *Papers from the 27th regional meeting of the Chicago Linguistic Society*, 31–45.

Butt, Miriam & Tracy Holloway King. 1996. Structural topic and focus without movement. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '96 conference*. Stanford: CSLI Publications.

Butt, Miriam & Tracy Holloway King (eds.). 2001. *Time over matter: Diachronic perspectives on morphosyntax*. Stanford: CSLI Publications.

Butt, Miriam & Tracy Holloway King. 2003. Case systems: Beyond structural distinctions. In Ellen Brandner & Heike Zinsmeister (eds.), *New perspectives on case theory*, 53–87. Stanford: CSLI Publications.

Butt, Miriam & Tracy Holloway King. 2004. The status of case. In Veneeta Dayal & Anoop Mahajan (eds.), *Clause structure in South Asian languages*. Berlin: Springer Verlag. DOI: 10.1007/978-1-4020-2719-2.

Butt, Miriam, Tracy Holloway King, María-Eugenia Niño & Frédérique Segond. 1999. *A grammar writer's cookbook*. Stanford: CSLI Publications.

Butt, Miriam & Aditi Lahiri. 2013. Diachronic pertinacity of light verbs. *Lingua* 135. SI: Complex Predicates, 7–29. DOI: 10.1016/j.lingua.2012.11.006.

Butt, Miriam, María-Eugenia Niño & Frédérique Segond. 2004. Multilingual processing of auxiliaries within LFG. In Louisa Sadler & Andrew Spencer (eds.), *Projecting morphology*. Stanford: CSLI Publications. Previously published as Butt, Niño & Segond (1996).

Butt, Miriam, María-Eugenia Niño & Frederique Segond. 1996. Multilingual processing of auxiliaries in LFG. In D. Gibbon (ed.), *Natural language processing and speech technology: Results of the 3rd KONVENS conference*, 111–122. Berlin: Mouton de Gruyter.

Butt, Miriam, S. Rajamathangi & Kengatharaiyer Sarveswaran. 2020. Mixed categories in Tamil via complex categories. In Miriam Butt & Ida Toivonen (eds.), *Proceedings of the LFG '20 conference*, 68–88. Stanford: CSLI Publications.

Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language* 82(4). 711–733. DOI: 10.1353/lan.2006.0186.

Camilleri, Maris & Louisa Sadler. 2018. Schematising (morpho)syntactic change in LFG: Insights from grammaticalisation in Arabic. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '18 conference*, 129–149. Stanford: CSLI Publications.

Camilleri, Maris & Louisa Sadler. 2020. The grammaticalisation of an auxiliary and a copula: The Arabic 'sit' participle. *Journal of Historical Syntax* 4(6). 1–60.

Campbell, Lyle. 1998. *Historical linguistics*. Edinburgh: Edinburgh University Press.

Campbell, Lyle. 2000. What's wrong with grammaticalization? *Language Sciences* 23(2). 113–161. DOI: 10.1016/S0388-0001(00)00019-X.

Chambers, J. K. & Natalie Schilling (eds.). 2013. *The handbook of language variation and change*. 2nd edn. Oxford: Wiley-Blackwell. DOI: 10.1002/9781118335598.

Clark, Brady Z. 2004. *A stochastic Optimality Theory approach to syntactic change*. Stanford: Stanford University. (Doctoral dissertation).

Crouch, Richard, Mary Dalrymple, Ronald M. Kaplan, Tracy Holloway King, John T. Maxwell III & Paula Newman. 2008. *XLE Documentation*. Xerox Palo Alto Research Center. Palo Alto, CA. https://ling.sprachwiss.uni-konstanz.de/pages/xle/doc/xle_toc.html.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III & Annie Zaenen (eds.). 1995. *Formal issues in Lexical-Functional Grammar*. Stanford: CSLI Publications.

Denison, David. 1993. *English historical syntax*. London: Longman. DOI: 10.4324/9781315842042.

Deo, Ashwini. 2015. The semantic and pragmatic underpinnings of grammaticalization paths: The progressive to imperfective shift. *Semantics and Pragmatics* 8. 1–52. DOI: 10.3765/sp.8.14.

Dost, Ascander & Tracy Holloway King. 2009. Using large-scale parser output to guide grammar development. In Tracy Holloway King & Marianne Santaholma (eds.), *Proceedings of the 2009 workshop on Grammar Engineering Across Frameworks (GEAF 2009)*, 63–70. Association for Computational Linguistics. DOI: 10.3115/1690359.1690367.

Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67. 547–619. DOI: 10.2307/415037.

Eythórsson, Thórhallur. 2008. The new passive in Icelandic really is a passive. In Thórhallur Eythórsson (ed.), *Grammatical change and linguistic theory: The Rosendal papers*, 173–219. Amsterdam: John Benjamins. DOI: 10.1075/la.113.07eyt.

Faarlund, Jan Terje. 1990. *Syntactic change: Toward a theory of historical syntax*. Berlin: De Gruyter. DOI: 10.1515/9783110854947.

Findlay, Jamie Y. & Anna Kibort. Forthcoming. Argument structure and mapping theory. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Fleischman, Suzanne. 2009. *The future in thought and language: Diachronic evidence from Romance*. Cambridge, UK: Cambridge University Press.

Frank, Anette, Tracy Holloway King, Jonas Kuhn & John T. Maxwell III. 2001. Optimality Theory style constraint ranking in large-scale LFG grammars. In Peter Sells (ed.), *Formal and empirical issues in Optimality Theoretic syntax*, 367–397. Stanford: CSLI Publications.

Givón, Talmy. 1990. *Syntax: A functional-typological introduction*. Amsterdam: John Benjamins.

Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge, UK: Cambridge University Press.

Hewson, John & Vit Bubenik. 2006. *From case to adposition: The development of configurational syntax in Indo-European languages*. Amsterdam: John Benjamins. DOI: 10.1075/cilt.280.

Hilpert, Martin. 2010. What can synchronic gradience tell us about reanalysis? Verb-first conditionals in written German and Swedish. In Elizabeth Closs Traugott & Graham Trousdale (eds.), *Gradience, gradualness and grammaticalization*, 181–201. Amsterdam: John Benjamins. DOI: 10.1075/tsl.90.10hil.

Hinterhölzl, Roland & Svetlana Petrova. 2010. From V1 to V2 in West Germanic. *Lingua* 120(2). 315–328. DOI: 10.1016/j.lingua.2008.10.007.

Hook, Peter Edwin. 1991. The emergence of perfective aspect in Indo-Aryan languages. In Elizabeth Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*, vol. II: Types of grammatical markers, 59–89. Amsterdam: John Benjamins. DOI: 10.1075/tsl.19.2.05hoo.

Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization*. 2nd edn. (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9781139165525.

Jespersen, Otto. 1927. *A modern English grammar on historical principles*. Vol. 3. London: Allen & Unwin.

Jónsson, Jóhannes Gísli. 2003. Not so quirky: On subject case in Icelandic. In E. Brandner & H. Zinsmeister (eds.), *New perspectives on case and case theory*, 129–164. Stanford: CSLI Publications.

Jónsson, Jóhannes Gísli. 2009. The new impersonal as a true passive. In Artemis Alexiadou, Jorge Hankamer, Thomas McFadden, Justin Nuger & Florian

Schäfer (eds.), *Advances in Comparative Germanic Syntax*, 281–306. Amsterdam: John Benjamins. DOI: 10.1075/la.141.12the.

Kager, René. 1999. *Optimality Theory*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/cbo9780511812408.

Kaplan, Ronald M. 1989. The formal architecture of Lexical-Functional Grammar. *Journal of Information Science and Engineering* 5. 305–322. Revised version published in Dalrymple, Kaplan, Maxwell & Zaenen (1995: 7–27).

Kellogg, Samuel Henry. 1893. *Grammar of the Hindi language*. Delhi: Munshiram Manoharlal Publishers Pvt. Ltd. Second Edition, reprinted 1990.

Kibort, Anna. 2013. Objects and lexical mapping theory. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '13 conference*. Stanford: CSLI Publications.

Kibort, Anna. 2014. Mapping out a construction inventory with (Lexical) Mapping Theory. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '14 conference*, 262–282. Stanford: CSLI Publications.

Kibort, Anna & Joan Maling. 2015. Modelling the syntactic ambiguity of the active vs. passive impersonal in LFG. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '15 conference*, 145–165. Stanford: CSLI Publications.

King, Tracy Holloway. 1997. Focus domains and information structure. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '97 conference*, 2–13. Stanford: CSLI Publications.

Kiparsky, Paul. 1982a. From cyclic phonology to lexical phonology. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations*, vol. 1, 131–175. Dordrecht: Foris Publications.

Kiparsky, Paul. 1982b. Lexical morphology and phonology. In I.-S. Yang (ed.), *Linguistics in the morning calm*, 3–91. Seoul: Hanshin.

Kiparsky, Paul. 1987. Morphology and grammatical relations. Unpublished manuscript, Stanford University.

Kiparsky, Paul. 1988. Agreement and linking theory. Unpublished manuscript, Stanford University.

Kiparsky, Paul. 1995. Indo-European origins of Germanic syntax. In Ian G. Roberts & Adrian Battye (eds.), *Clause structure and language change*, 140–167. Oxford: Oxford University Press.

Kiparsky, Paul. 1997. The rise of positional licensing. In Ans van Kemenade & Nigel Vincent (eds.), *Parameters of morphosyntactic change*, 460–494. Cambridge, UK: Cambridge University Press.

Kiparsky, Paul. 2001. Structural case in Finnish. *Lingua* 111(4). 315–376. DOI: 10.1016/S0024-3841(00)00035-8.

É. Kiss, Katalin. 1995. *Discourse configurational languages*. Oxford: Oxford University Press.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1(3). 199–244. DOI: 10.1017/s0954394500000168.

Kroch, Anthony. 2001. Syntactic change. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*, 699–729. Oxford: Blackwell.

Kuhn, Jonas. Forthcoming. LFG, Optimality Theory and learnability of languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Labov, William. 1994. *Principles of linguistic change: Internal factors*. Oxford: Blackwell.

Labov, William. 2001. *Principles of linguistic change: Social factors*. Oxford: Blackwell.

Lahiri, Aditi (ed.). 2000a. *Analogy and markedness: Principles of change in phonology and morphology*. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110899917.

Lahiri, Aditi. 2000b. Hierarchical restructuring in the creation of verbal morphology in Bengali and Germanic: Evidence from phonology. In Aditi Lahiri (ed.), *Analogy, levelling, markedness: Principles of change in phonology and morphology*, 71–123. Berlin: Mouton de Gruyter.

Landau, Idan. 2010. *The locative syntax of experiencers*. Cambridge, MA: The MIT Press.

Langacker, Ronald W. 1977. Syntactic reanalysis. In Charles N. Li (ed.), *Mechanisms of syntactic change*, 59–139. Austin: University of Texas Press.

Ledgeway, Adam. 2012. From Latin to Romance: Configurationality, functional categories and head-marking. *Transactions of the Philological Society* 110(3). 422–442. DOI: 10.1111/j.1467-968x.2012.01310.x.

Lehmann, Christian. 1985. Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile* 20. 303–318.

Lightfoot, David. 1979. *Principles of diachronic syntax*. Cambridge, UK: Cambridge University Press.

Lightfoot, David. 1981. The history of NP movement. In C. L. Baker & John McCarthy (eds.), *The logical problem of language acquisition*, 86–119. Cambridge, MA: The MIT Press.

Lightfoot, David. 1991. *How to set parameters: Arguments from language change*. Cambridge, MA: The MIT Press.

Lightfoot, David. 1997. Catastrophic change and learning theory. *Lingua* 100(1–4). 171–192. DOI: 10.1016/s0024-3841(93)00030-c.

Lord, Carol. 1993. *Historical change in serial verb constructions*. Amsterdam: John Benjamins.

Lowe, John J. 2015. Degrees of degrammaticalization: A lexical sharing approach to the English possessive. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '15 conference*, 208–228. Stanford: CSLI Publications.

Lowe, John J. 2016a. English possessive *'s*: Clitic *and* affix. *Natural Language & Linguistic Theory* 34. 157–195.

Lowe, John J. 2016b. Participles, gerunds and syntactic categories. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-Driven Phrase Structure Grammar and Lexical Functional Grammar*, 401–421. Stanford: CSLI Publications.

Luraghi, Silvia. 2010. The rise (and possible downfall) of configurationality. In Silvia Luraghi & Vit Bubenik (eds.), *Continuum companion to historical linguistics*, 212–229. London: Continuum.

Maling, Joan & Sigrídur Sigurjónsdóttir. 2002. The new impersonal construction in Icelandic. *Journal of Comparative Germanic Linguistics* 5(1–3). 97–142.

Maling, Joan & Annie Zaenen (eds.). 1990. *Modern Icelandic syntax* (Syntax and Semantics 24). San Diego, CA: Academic Press. DOI: 10.1163/9789004373235.

Meillet, Antoine. 1912. L'évolution des formes grammaticales. *Scientia* (Rivista di Scienza) 12(6). Reprinted in Antoine Meillet, 1958, *Linguistique historique et linguistique générale*, Paris, Champion.

Mohanan, K. P. 1986. *The theory of lexical phonology*. Dordrecht: Reidel. DOI: 10.1007/978-94-009-3719-2.

Montaut, Annie. 2003. Oblique main arguments in Hindi/Urdu as localizing predications: Questioning the category of subject. In Peri Bhaskarao & K.V. Subbarao (eds.), *Non-nominative subjects*, 33–56. Amsterdam: John Benjamins.

Narrog, Heiko & Bernd Heine. 2017. Grammaticalization. In Adam Ledgeway & Ian Roberts (eds.), *The Cambridge handbook of historical syntax* (Cambridge Handbooks in Language and Linguistics), 7–27. Cambridge, UK: Cambridge University Press. DOI: 10.1017/9781107279070.002.

Newmeyer, Frederick J. 2000. Deconstructing grammaticalization. *Language Sciences* 23(2). 187–229. DOI: 10.1016/S0388-0001(00)00021-8.

Nikitina, Tatiana. 2008. *The mixing of syntactic properties and language change*. Stanford: Stanford University. (Doctoral dissertation).

Nikolaeva, Irina & Andrew Spencer. 2020. *Mixed categories: The morphosyntax of noun modification*. Cambridge, UK: Cambridge University Press.

Norde, Muriel. 2009. *Degrammaticalization*. Oxford: Oxford University Press.

Nordlinger, Rachel. 1998. *Constructive case: Evidence from Australian languages*. Stanford: CSLI Publications.

Payne, John R. 2009. The English genitive and double case. *Transactions of the Philological Society* 107(3). 322–357. DOI: 10.1111/j.1467-968x.2009.01226.x.

Perlmutter, David M. & Paul M. Postal. 1977. Toward a universal characterization of passivization. In *Proceedings of the 3rd annual meeting of the Berkeley Linguistics Society*, 394–417. Berkeley: Berkeley Linguistics Society. DOI: 10.3765/bls.v3i0.2266.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee & Paul J. Hopper (eds.), *Frequency and the emergence of linguistic structure*, 137–157. Amsterdam: John Benjamins. DOI: 10.1075/tsl.45.08pie.

Pierrehumbert, Janet B. 2002. Word-specific phonetics. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory phonology*, vol. 7, 101–140. Berlin: De Gruyter. DOI: 10.1515/9783110197105.1.101.

Pintzuk, Susan. 2003. Variationist approaches to syntactic change. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 509–528. Oxford: Blackwell. DOI: 10.1002/9781405166201.ch15.

Ponti, Edoardo Maria & Silvia Luraghi. 2018. Non-configurationality in diachrony: Correlations in local and global networks of Ancient Greek and Latin. *Diachronica* 35(3). 367–392. DOI: 10.1075/bct.113.03pon.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. Harlow: Longman.

Ramchand, Gillian. 2008. *Verb meaning and the lexicon: A first phase syntax*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/cbo9780511486319.

Roberts, Ian & Anna Roussou. 2003. *Syntactic change: A minimalist approach to grammaticalisation*. Cambridge, UK: Cambridge University Press.

Rögnvaldsson, Eiríkur. 1995. Old Icelandic: A non-configurational language? *NOWELE: North-Western European Language Evolution* 26(1). 3–29. DOI: 10.1075/nowele.26.01rog.

Schätzle, Christin. 2018. *Dative subjects: Historical change visualized*. Konstanz: University of Konstanz. (Doctoral dissertation).

Schätzle, Christin, Frederik L. Dennig, Michael Blumenschein, Daniel A. Keim & Miriam Butt. 2019. Visualizing linguistic change as dimension interactions. In *Proceedings of the 1st international workshop on computational approaches to historical language change*, 272–278. Florence: Association for Computational Linguistics. DOI: 10.18653/v1/W19-4734.

Sells, Peter. 2001. *Structure, alignment and optimality in Swedish*. Stanford: CSLI Publications.

Sells, Peter. 2005. The peripherality of the Icelandic expletive. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '05 conference*, 408–428. Stanford: CSLI Publications.

Sharma, Devyani, Joan Bresnan & Ashwini Deo. 2008. Variation and change in the individual: Evidence from the survey of English dialects. In Robin Cooper & Ruth Kempson (eds.), *Language in flux: Dialogue coordination, language variation, change and evolution*, 265–321. London: College Publications.

Simpson, Jane. 2001. Preferred word order and the grammaticalization of associated path. In Miriam Butt & Tracy Holloway King (eds.), *Time over matter: Diachronic perspectives on morphosyntax*, 173–208. Stanford: CSLI Publications.

Smith, Henry. 1996. *Restrictiveness in case theory*. Cambridge, UK: Cambridge University Press.

Snijders, Liselotte. 2015. *The nature of configurationality in LFG*. Oxford: University of Oxford. (D.Phil. Thesis).

Sun, Chaofen & Elizabeth Closs Traugott. 2011. Grammaticalization and word order change. In Heiko Narrog & Bernd Heine (eds.), *The Oxford handbook of grammaticalization*, 378–388. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199586783.013.0030.

Talmy, Leonard. 1978. Figure and ground in complex sentences. In Joseph H. Greenberg (ed.), *Universals of human language*, vol. 4, 625–649. Stanford: Stanford University Press.

Thráinsson, Höskuldur. 2007. *The syntax of Icelandic*. Cambridge, UK: Cambridge University Press.

Toivonen, Ida. 2001. *The syntax of non-projecting words*. Stanford: Stanford University. (Doctoral dissertation).

Toivonen, Ida. 2007. Verbal agreement in Inari Saami. In Ida Toivonen & Diane Nelson (eds.), *Saami linguistics* (Current Issues in Linguistic Theory 288), 227–258. Amsterdam: John Benjamins. DOI: 10.1075/cilt.288.09toi.

Traugott, Elizabeth Closs & Graeme Trousdale. 2010. *Gradience, gradualness and grammaticalization*. Amsterdam: John Benjamins.

van der Gaaf, Willem. 1929. The conversion of the indirect personal object into the subject of a passive construction. *English Studies* 11. 58–67. DOI: 10.1080/00138382908596544.

van Gelderen, Elly. 2011. *The linguistic cycle: Language change and the language faculty*. Oxford: Oxford University Press.

Vincent, Nigel. 1999. The evolution of c-structure: Prepositions and PPs from Indo-European to Romance. *Linguistics* 37(6). 1111–1153. DOI: 10.1515/ling.37.6.1111.

Vincent, Nigel. 2001. LFG as a model of syntactic change. In Miriam Butt & Tracy Holloway King (eds.), *Time over matter: Diachronic perspectives on morphosyntax*, 1–42. Stanford: CSLI Publications.

Vincent, Nigel & Kersti Börjars. 2010. Complements of adjectives: A diachronic approach. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '10 conference*, 458–478. Stanford: CSLI Publications.

Vincent, Nigel & Kersti Börjars. 2020. Heads and history. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences I: Syntax inside the grammar*, 133–157. Berlin: Language Science Press.

Weinreich, Uriel, William Labov & Marvin Herzog. 1968. Empirical foundations for a theory of language change. In Winfried P. Lehmann & Yakov Malkiel (eds.), *Directions for historical linguistics: A symposium*, 95–188. Austin: University of Texas Press.

Wescoat, Michael T. 2002. *On lexical sharing*. Stanford: Stanford University. (Doctoral dissertation).

Wescoat, Michael T. 2005. English nonsyllabic auxiliary contractions: An analysis in LFG with lexical sharing. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '05 conference*, 468–486. Stanford: CSLI Publications.

Wescoat, Michael T. 2007. Preposition-determiner contractions: An analysis in Optimality-Theoretic Lexical-Functional Grammar with lexical sharing. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '07 conference*, 439–459. Stanford: CSLI Publications.

Wescoat, Michael T. 2009. Udi person markers and lexical integrity. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '09 conference*, 604–622. Stanford: CSLI Publications.

Willis, David. 2017. Degrammaticalization. In Adam Ledgeway & Ian Roberts (eds.), *The Cambridge handbook of historical syntax* (Cambridge Handbooks in Language and Linguistics), 28–48. Cambridge, UK: Cambridge University Press. DOI: 10.1017/9781107279070.003.

Zaenen, Annie. 1993. Unaccusativity in Dutch: Integrating syntax and lexical semantics. In James Pustejovsky (ed.), *Semantics and the lexicon*, 129–161. Dordrecht: Kluwer Academic Publishers. DOI: 10.1007/978-94-011-1972-6_9.

Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language & Linguistic Theory* 3(4). 441–483. DOI: 10.1007/bf00133285. Reprinted in Maling & Zaenen (1990: 95–136).

Zwicky, Arnold M. 1987. Suppressing the Zs. *Journal of Linguistics* 23. 133–148. DOI: 10.1017/s0022226700011063.

Chapter 3

LFG, Optimality Theory and learnability of languages

Jonas Kuhn

University of Stuttgart

Optimality-Theoretic accounts of grammar specification do not rely (exclusively) on rules and descriptions that every expression in the implied formal language has to satisfy strictly. Instead, a universal set of violable constraints is posited, and most grammatical expressions are not perfectly harmonic in all respects, but violate some of the constraints. A particular language is characterized by a certain priority ranking over the universal constraint set, which results in cross-linguistic variation with respect to the optimal way of expressing an underlying meaning: the realization that incurs the least serious constraint violations. The framework comes with a straightforward algorithmic formulation of the language learning problem (as error-driven constraint reranking).

Optimality Theory (OT) was combined with many linguistic description frameworks in the 1990s. LFG turned out to be a very appropriate base formalism for specifying the candidate representations in OT's competition-based definition of grammaticality. The novel way of characterizing formal languages prompted a range of debates regarding central assumptions in linguistic theorizing; various extensions of the competition-based setup were proposed; and the concept of violable constraints was taken over as an effective modeling device for managing ambiguities in broad-coverage computational grammar development. This chapter provides an introduction of the core concepts of OT as fleshed out on the basis of LFG, it illustrates the most influential extensions, and it reviews important conceptual debates triggered by the approach.

1 Introduction

The term Optimality Theory (OT) refers to a family of grammatical frameworks developed in various subfields of linguistics following the original proposal for

Jonas Kuhn. Forthcoming. LFG, Optimality Theory and learnability of languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

phonology by Prince & Smolensky (1993). Bresnan (1996), in her keynote address at the first LFG conference,[1] showed that LFG constitutes a natural base formalism for an OT account of syntax, laying the foundation for substantial research activities in the late 1990s and early 2000s.

The key idea behind OT is that grammatical knowledge is not captured in distinct rule systems for each language, but theorists assume a universal set of constraints responsible for determining the harmony or markedness of potential realizations of some underlying input representation. For example, both the clause structure *you see who?* and *who do you see?* are candidate realizations for expressing a *wh*-question that asks for the object of a transitive verb. In some languages, the equivalent of the former is grammatical, whereas in English, the latter is. Neither of the two variants is perfectly unmarked; each violates certain constraints (in essence: an interrogative phrase is best realized at the beginning of the clause vs. the object of a verb is best realized in the canonical VP-internal complement position). What differs across languages is the relative prominence among the universal constraints, such that there are different winners in the competition for the most harmonic/least marked candidate, which is defined to be the language-particular output. Learning a language thus amounts to determining the correct prominence ranking over a known set of constraints to replicate the behavior observed in adult speakers. A more technical illustration of an OT system building on the LFG formalism will follow in Section 1.1.

The notion of violable grammar constraints inspired linguists to explore the explanatory potential of a competition-based definition of grammaticality for a whole range of linguistic subfields. Moreover as Section 1.2 will discuss, the notion turned out fruitful for computational work aiming at linguistically grounded broad-coverage grammars: expanding "classical" constraint-based grammar formalisms with a novel type of violable constraints enables a competition-based filtering of ambiguity sources and thus greatly facilitates the treatment of relatively rare lexical variants and grammatical constructions (Frank et al. 2001). An outline of the remaining sections of this chapter will be provided in Section 1.3.

1.1 Example of an OT-LFG analysis

The tableau in Figure 1 demonstrates more technically how an OT-LFG account predicts *who will she see?* to be the grammatical structure for realizing a matrix question in English. It is adapted from (Bresnan 2001a), a contribution that recasts Grimshaw's (1997) OT analysis of English verb inversion in an LFG framework

[1] Details of the analysis presented in the keynote address are discussed in Bresnan 2001a and Bresnan 2002.

assuming Extended Head Theory (as detailed in Bresnan 2001b: ch. 7 and Bresnan et al. 2016: ch. 7).

Input: $\begin{bmatrix} \textsc{pred} & \text{`see(x, y)'} \\ \textsc{gf}_1 & [\text{“you”}] \\ \textsc{gf}_2 & [\text{“who”}] \\ \textsc{q-foc} & \text{——} \\ \textsc{tns} & \textsc{fut} \end{bmatrix}$	Op-Spec	Ob-Hd	Stay
a. ⟨ [$_{\text{IP}}$ you will [$_{\text{VP}}$ see who]], ...⟩	*!		
b. ⟨ [$_{\text{CP}}$ who *e* [$_{\text{IP}}$ you will [$_{\text{VP}}$ see *e*]]], ...⟩		*!	*
☞ c. ⟨ [$_{\text{CP}}$ who will [$_{\text{IP}}$ you *e* [$_{\text{VP}}$ see *e*]]], ...⟩			**

Figure 1: Tableau for the competition that leads to *who will you see?* being the most harmonic candidate under the constraint ranking for English. The candidates' c-structure trees are shown as bracketing expressions; their full f-structures is not shown here. Only a small subset of the candidate set is listed, omitting other realization variants. These are excluded for English by additional constraints, but could be grammatical in other languages and are thus included in the candidate set.

The tableau shows the comparison of three candidate realizations for the underlying content, the input – here the proposition *see*(x, y), where x is a pronominal for the second person singular, y is the focus of a question, and the clausal realization expresses future tense. Particular syntactic choices, for instance the realization of the semantic arguments by grammatical relations such as SUBJECT and OBJECT are underspecified in the input. In the OT-LFG framework, the range of realization options across languages can be captured very well by assuming that the input is a partially specified f-structure, and the candidate set consists of all fully specified LFG analyses whose f-structure subsumes the input.

The excerpt from the full candidate set shown in Figure 1 focuses on c-structural realization alternatives for clause structure. In Extended Head Theory, functional categories (such as I and C) and their c-structure projections are mapped to the same f-structure as their corresponding lexical categories (V). The functional projections occur optionally, but they offer the possibility to realize additional head elements such as tense auxiliaries and they come with one c-structural specifier position each, which can for instance be used for *wh* operators. Also, it is a possibility across languages for verbal head elements to be realized in a higher, c-commanding, position (e.g., V in I or I in C, as in candidate c.). By definition, elements in such a functional position are still *extended heads* of the lower projections. This nonderivational account of head mobility, which Bresnan (2001b:

ch. 7) also formalizes in a non-OT setting of LFG, thus opens up a considerable spectrum of alternative candidate realizations for the basic clause structure.

Let us now consider what effect the three violable constraints have that Figure 1 shows in the order of the ranking for English (the constraint formalization will be discussed in more detail in Section 2.2). The harmony/markedness of a candidate is determined by its constraint violation profile. In the tableau notation, a violation is signalled by a star in the row for the corresponding constraint. Since OT assumes a *ranking* of the constraints (and not a more general weighing as for instance in Harmonic Grammar, Legendre et al. 1990), one can simply proceed from left to right through the constraints to determine the winning candidate(s): at each constraint, only the candidates that have the fewest violations for this constraint are kept. The other candidates have no more chance of winning the competition, even when they do not violate any of the lower-ranking constraints. It is common to mark the decisive constraint violations with an exclamation mark following the star; the cells for the lower-ranking constraints are often grayed out in tableaux. In our example competition, the Op-Spec constraint demands that syntactic operators such as *wh* phrases be realized in c-structural specifier positions (highlighting the operator's prominence). Candidate a. violates this constraint, since the *who* is realized inside the VP. Op-Spec is the highest-ranking constraints and the other candidates satisfy this constraint, hence candidate a. is excluded from further comparison. The Ob-Hd constraint (for obligatory head) says that every projected category (X′, X″) should have a lexically filled extended head. Candidate b., which includes a CP projection without a lexically filled head, incurs a violation of this constraint.[2] The third constraint, Stay,[3] states that categories should *dominate* their extended heads, thus punishing c-structural realizations in higher, c-commanding positions: In candidate c., the *wh* phrase in the CP-Spec position incurs one Stay violation (due to a lexically empty DP inside the VP), the *will* that is not realized within IP another one. Since Stay is low in the constraint ranking for English, candidate c. is optimal despite the two violations: there are no surviving competitors with fewer violations.

Under different rankings of the same three constraints, different clause structures come out as optimal. When Stay outranks Op-Spec, for instance, candidate

[2] The notion of *extended* heads leads to the situation that the empty I head in candidate c. does not incur an Ob-Hd violation: The C head, which contains *will*, is in a c-commanding position and thus acts as an extended head for the I′ and IP.

[3] The name Stay is carried over from Grimshaw's (1997) account, which assumed a derivational base formalism with upward movement in trees.

a. is the most harmonic. This predicts non-echo matrix question like in (1) to be grammatical, as is the case in a *wh*-in-situ language like Chinese.

(1) Mandarin Chinese (Huang 1982: 253)
ni kanjian-le shei?
you see-ASP who
'Who did you see?'

A language learner starts out with an uninformed initial constraint ranking. Whenever they hear an utterance by an adult speaker of the language, they check whether the observed output (for a contextually inferred input) is the same that the learner's current ranking would have predicted. If not, the constraints responsible for the error are demoted in the ranking (Tesar & Smolensky 1998). Details of the OT approach to characterizing grammaticality and the OT-LFG formalization will be discussed Section 2.

1.2 Ranked soft constraints in computational grammar development

As mentioned, Joan Bresnan gave a keynote at the first LFG conference in 1996, presenting her proposal for integrating the competition-based grammatical framework of Optimality Theory with LFG's declarative specification of parallel (candidate) representations that are in imperfect correspondence (Bresnan 1996). The concept of violable constraints immediately prompted interest in the computational LFG community, which was working towards implementations of LFG grammars that robustly cover real-life text data, most notable in the Parallel Grammar Development project ParGram (see also Forst & King forthcoming [this volume]).

As Frank et al. (2001) discuss, violable constraints can alleviate a considerable practical problem with broad-coverage grammars that aim for linguistic precision: the *ambiguity management* problem. To reach an acceptable coverage of real corpus data with a linguistically precise grammar, the rules have to cater for essentially all conceivable realizations of variable phenomena – even for those that are infrequent and tend to underlie special contextual restrictions. In a grammar with such a comprehensive rule set however interactions among multiple variable phenomena can lead to an enormous degree of ambiguity in parsing. Although ambiguity packing techniques keep the computational parsing task itself tractable (Maxwell & Kaplan 1989) (see also Cahill & Way forthcoming [this volume]), it can become hard for the grammar writers to keep track of whether the contextually justified readings are among the readings predicted by the grammar (King et al. 2004).

The following examples may illustrate the problem. In English, temporal adverbials can have the shape of a plain NP headed by a time expression such as *year, day, moment*, etc., as for instance in (2a,b). A carefully constrained grammar will rely on a lexical marking of the relevant nouns to avoid overgeneration. Without such a lexical restriction, any intransitive verb used with an NP in object position, such as (2c), would be incorrectly predicted to form a grammatical sentence.

(2) a. The car had problems all weekend.[4]
    b. What are your future plans the next half year?[5]
    c. * Lou yawned two big bones.

The rule in (3) uses a constraining equation to ensure that the adverbial NP comes with the marking [NTYPE TIME +].

(3)
$$\begin{array}{llllllll} \mathrm{VP} \rightarrow & \mathrm{V} & (\mathrm{NP}) & \{ & \mathrm{PP} & | & \mathrm{NP} & \}^* \\ & \uparrow=\downarrow & (\uparrow\mathrm{OBJ})=\downarrow & & \downarrow\in(\uparrow\mathrm{ADJ}) & & \downarrow\in(\uparrow\mathrm{ADJ}) & \\ & & & & & & (\downarrow\mathrm{NTYPE\ TIME}) =_c + & \end{array}$$

Examples (4a,b) show however that adverbial uses of plain NPs can also occur with nouns that do not refer to time expressions.

(4) a. Being prepared will save your soles from aching the entire trip.[6]
    b. And WoW had problems every expansion for years [...][7]

By a semantic process of type coercion (see e.g. Pustejovsky 1995), NPs referring to segments of a process (e.g. during spatial movement, as in (4a), or in repetitive processes, as in (4b)) can receive a temporal reading. A grammar that strictly ties adverbial NPs to a lexical marking of time expressions will fail on these examples. And since such coercion processes are relatively productive, grammar writers face a difficult tradeoff between maximizing coverage and risking a proliferation of ambiguity.

The issue is amplified when several variation phenomena interact, as the following considerations demonstrate. Completely independent of the type coercion process we just saw, many transitive verbs in English can occur with an

[4] https://www.flickr.com/photos/speedcenter/19225841474
[5] https://www.reddit.com/r/UnderCards/comments/kdjbb1/ama_is_the_way_to_get_a_free_vieeeews/
[6] https://vivaglammagazine.com/how-to-effortlessly-look-good-when-traveling/
[7] https://forums.newworld.com/t/fix-your-product-this-isnt-an-indie-company/247609/59

understood object in appropriate contexts – i.e. without a syntactically realized direct object: *enjoy* and *announce* are both strictly transitive, but (5a,b) show that occasionally, an understood object can be filled in contextually.

(5) a. [Text message with a video posting:]
Here it is. Lean back and enjoy for the next 7 minutes.[8]

b. [News headline:]
Troy Williams will announce tomorrow[9]

To capture such uses, one could have transitive verbs optionally fill in a pronominal PRED value for their object (this can for instance be done by putting an optional equation in the lexical template for transitive verbs: { (↑ OBJ PRED) = 'pro' }). The pronominal PRED value indicates that its referent can be contextually inferred.

Both the temporal type coercion discussed previously and understood objects occur rather infrequently overall; yet, a broad-coverage grammar should cater for these constructions. Unfortunately, it is impossible to fully capture the contextual constraints on the use of such constructions within a knowledge engineering approach – individual uses may depend on inferences involving situational knowledge and general world knowledge.

By employing violable "soft" constraints inspired by Optimality Theory in a broad-coverage LFG grammar, the infrequent constructions can be captured in the grammar without causing a proliferation of ambiguity for the core constructions. Sentence (6a) is an example of a plain transitive use of the verb *announce*. A conspiracy of the two infrequent constructions we just discussed would however also lead to an understood object reading of *announced* (corresponding to the only reading available for (5b)) when *every expansion* would be analyzed as a coerced temporal adjunct (like in (2e)). In effect, every single transitive VP (as in (6b)) would receive duplicate readings. As (6c) shows, even an NP including a time expression will rather fill the position of the direct object than acting as an adverbial NP alongside an understood object (although in this particular case, the two readings are semantically hard to distinguish).

(6) a. The company announced every expansion.

b. The dog chases every bird.

[8] https://mobile.twitter.com/automobilirimac/status/765505874425180160

[9] https://www.on3.com/teams/kentucky-wildcats/news/troy-williams-will-announce-tomorrow/amp/

c. [From a comment page for a music band:]
Super talented, professional. We enjoyed every minute.[10]

The mechanism of ranked violable constraints discussed by Frank et al. (2001), which is implemented in the XLE system (Crouch et al. 2008), provides a highly convenient way of dealing with this situation. In (7), template definitions are seen that optionally[11] introduce the infrequent construction of a type-coerced common noun (such as *expansion, trip* and even *bird*) and an understood object with a transitive verb.

(7) a. COMMON-NOUN(_P) ≡ (↑ PRED) = '_P'
{ (↑ NTYPE TIME) = +
TypeCoercionToTimeNoun ∈ o(*) }

b. TRANSITIVE(_P) ≡ (↑ PRED) = '_P ⟨ (↑ SUBJ) (↑ OBJ)⟩'
{ (↑ OBJ PRED) = 'PRO'
UnderstoodObj ∈ o(*) }

Note that the optional functional descriptions go along with statements of the form 'XYZ ∈ o(*)'. These have the effect of introducing so-called optimality marks (XYZ etc.) to a data structure projected from c-structure to a novel o-projection (for Optimality projection).[12] Mathematically, the data structure is a multiset (i.e. a generalization of sets that can include two or more identical 'copies' as distinct elements). The purpose of this multiset is simply to implement a counting mechanism that keeps track of how many times a certain functional description was used in the construction of a particular LFG analysis. In c-structure rules $A \rightarrow B_1 \ldots B_n$, the multisets from all daughter categories $B_1 \ldots B_n$ are identified with the multiset projected from the mother category A, such that the root node of a c-structure tree will always provide access to the aggregate counts of marks from the full tree.

The multiset of optimality marks provides the basis for filtering a set of candidate analysis, very much like in the OT tableaux illustrated in Section 1.1. The grammar writer can define the relative ranking of the violable constraints (the OPTIMALITYORDER) that guides the filtering:

[10] https://www.esteemlivemusic.com/we-enjoyed-every-moment/

[11] Enclosing functional descriptions in curly brackets has the effect of creating one option with the descriptions and one without.

[12] Like the ϕ-projection, the o-projection starts out from c-structural entities (i.e., lexical or phrasal c-structure nodes). This makes it possible to introduce distinct marks from several c-structure nodes that map to the same f-structure. This can for instance be relevant for economy constraints that favor structures with fewer nodes.

(8) OPTIMALITYORDER UnderstoodObj TypeCoercionToTimeNoun.

Starting from the highest-ranking type of optimality mark (= OT constraint), the set of readings is successively reduced; for each type of OT mark only the candidates that have the fewest marks is kept. The effect of this is that the infrequent constructions will only be seen in the final output of the grammar when there is no alternative, more canonical way of analyzing the same string. In our example scenario, plain transitive sentences like (as (6b)) will hence receive only the canonical subject-object reading. At the same time however, the grammar *does* cover uses of type coercion or understood objects. For instance when analyzing sentences that contain type-coerced adverbial NPs like (4a,b), *all* candidate analyses include at least one mark TypeCoercionToTimeNoun, hence they will survive this step of the filtering.

With such a filtering mechanism, the grammar writers can transparently control many of the interaction effects of the linguistic phenomena they are dealing with – without having to come up with explicit descriptions of contextual conditions for the rare special uses. The approach has hence been widely applied in the broad-coverage grammars of the ParGram project. As we will see in Section 2.6, some modifications of the plain ranking mechanism open up further functionality for grammar development.

To conclude this section we note however that the purpose of applying a competition-based filter on a set of candidate analyses is quite different in OT-LFG (as discussed in Section 1.1) vs. "OT-style" soft constraints in grammar development: in the former case, the competition serves to define grammaticality in a given language, whereas in the latter case, all candidates are assumed to be grammatical in principle, and the competition serves to filter out implausible readings.

1.3 Chapter outline

The remainder of the chapter is organized into four sections and a conclusion: Section 2 discusses the OT framework for defining grammaticality and its combination with the LFG formalism in more detail. Section 3 provides some demonstrations of how the competition-based notion of grammaticality can systematically derive variation patterns within a language and predict typological patterns across languages. Section 4 goes into important extensions of the standard OT framework and reviews discussions of the status of key components of an OT system such as the violable constraints. Section 5 reviews the developments after the mid-2000s, with an emphasis on the relationship between OT and other competition-based architectures and on learnability. Section 6 concludes with a summary of important epistemological considerations.

2 OT: Formalism and computational considerations

Section 2.1 discusses the key components of an OT system in general and their motivation for the Theory of Grammar. On this basis, Section 2.2 addresses the application of OT systems in syntax and expands on the OT-LFG approach outlined in the introduction. Section 2.3 contrasts the OT-LFG approach with OT syntax approaches using a derivational base formalism.

Section 2.4 discusses the most important computational considerations regarding OT-LFG. An excursion in Section 2.5 addresses potential concerns regarding plausibility of OT-LFG that might arise if one assumes that the declarative specification of the candidate set has any direct psychological reality during human sentence processing. Section 2.6 provides some more details on the use of soft constraints in broad-coverage grammar development illustrated in Section 1.2.

2.1 Key components and assumptions of OT

The initial example shown in Figure 1 already provided an illustration of the key components of Optimality Theory in general and OT-LFG in particular. Within an OT system, the set of grammatical expressions for a particular natural language L is defined using (i) a language-independent function *Gen*, which maps any given input representation to a set of candidates, (ii) a universal set of violable constraints *Con*, and (iii) a harmony evaluation function *Eval*, applying a language-specific prominence ranking $\gg_L$ that is defined over the constraint set. We define an expression to be grammatical in L if it is (one of) the most harmonic candidate(s) in the set of competing realizations for some input according to $\gg_L$. A candidate C_1 is more harmonic than a candidate C_2 iff it contains fewer violations for the highest-ranked constraint in which the marking for C_1 and C_2 differs.

An OT system thus does not only predict a single language, but essentially an entire constellation of possible languages: the so-called factorial typology opened up by the different relative rankings that are possible for the assumed constraints. The task for a theorist working on some family of phenomena is to find a set of constraints that plausibly predicts the spectrum of cross-linguistic variation as it is empirically attested in the languages of the world.

A key methodological assumption of linguistic work employing OT systems is that the input includes no language-specific information or restrictions. This is called the "richness of the base" (Smolensky 1996). A construction that is only available in some languages should not be explained by differences in the input, but must be derived as an effect of constraint ranking. In other words, the role of candidate generation, *Gen*, is limited to providing the space of possibilities; the substantive linguistic regularities are all dealt with in the constraint set.

In phonology, the input can be thought of as the target sound sequence a speaker intends to realize. In the case of loan words, this target sequence may include sound clusters that cannot be expressed in the language. For instance, the Dutch loan word *plan* is realized as *päläna* in Sinhalese:

(9) Sinhalese (Indo-Aryan; Boersma et al. 2000: 5, citing Sannasgala 1976)
päläna
'plan'

The OT explanation is that there is a constraint against complex syllable onsets and that in the constraint ranking for Sinhalese, this constraint is ranked higher than a constraint against the addition of epenthetical vowels. As a consequence, a surface realization that deviates from the target sound sequence arises as optimal. To capture the full spectrum of phonological systems, *Gen* must make far-reaching deviations from the input available.

Constraints that evaluate whether a property from the input is preserved in the output are called *faithfulness constraints*; there are three types, punishing (a) the insertion and (b) the deletion of a segment and (c) the alteration of some feature value (e.g. devoicing of a voiced consonant) (Prince & Smolensky 1993; McCarthy & Prince 1995). Which part of the space of more or less faithful candidates a language uses for a certain phenomenon is determined by the relative ranking of faithfulness constraints and a second class of constraints, the so-called *markedness constraints*, which assess the output shape of a candidate, irrespective of the input it is supposed to convey. By and large, OT predicts typological differences to arise from different strategies of reconciling conflicts among the two major types of constraints.

2.2 Optimality-Theoretic Syntax and its formalization

The predictive potential of OT systems regarding language typology as demonstrated in phonology soon attracted researchers working on different subfields of grammar, including syntax (Grimshaw 1997; Samek-Lodovici 1996; Pesetsky

1998; Legendre et al. 1998).[13] Technically, it is relatively straightforward to assume that some modular part of a conventional theory of grammar is replaced by an OT system: instead of the conventional output representation, a set of alternative candidates is assumed whose harmony is then assessed based on ranked violable constraints.[14] (It should be noted however that depending on which part of a conventional approach is being replaced by the output of an OT competition, there can be great differences in the predictive potential of the model component; we come back to this in Section 2.3.)

The fact that the explanatory power of an OT system comes largely from constraint interaction had an interesting effect in terms of sociology of science: Optimality-theoretic extensions of base formalisms from different schools of thought could be subjected to a meaningful comparison, even where diverging assumptions in the conception of the base formalisms themselves had before made comparisons very difficult. This led to an opening up of channels for exchange among theoretical frameworks (see for instance the collection of contributions in Legendre et al. 2000).[15]

Early work on OT syntax sometimes relied on an intuitive and informal conception of which syntactic structures should compete with each other in the same candidate set. To preserve the idea from OT phonology that all conceivable alternatives for saying the same thing are included in the candidate set, the natural assumption for the input in OT syntax is an abstract representation of the syntactically relevant *semantic content* of an utterance. As Bresnan (1996) points out, this intuition can be cashed out straightforwardly if candidate generation is based on a monotonic unification formalism like LFG: the input can be defined as a partial representation including the relevant semantic information, and the set of candidates are all the complete representations expanding this input representation – in the case of LFG pairs of c-structures and corresponding f-structures that are subsumed by the input presentation, as is illustrated in Figure 2 from Bresnan 2001a. Cross-linguistic variation at the level of constituent structure is captured by including all typologically different options for expressing some core semantic information in the candidate set.

[13] The account of unaccusativity in French by Legendre et al. (1990) in the framework of Harmonic Grammar predates much of the work in OT phonology. Nevertheless, many syntacticians presumably only took notice of competition-based approaches in the aftermath of the success of OT phonology and reacting to Grimshaw's (1997) paper.

[14] The approach of "Harmonic Serialism" assumes an architecture that is broken down into a sequence of steps with local optimization. Müller (2020) for instance explores an exploration to morphology.

[15] This unifying potential is discussed for instance by Newmeyer (2002: 44).

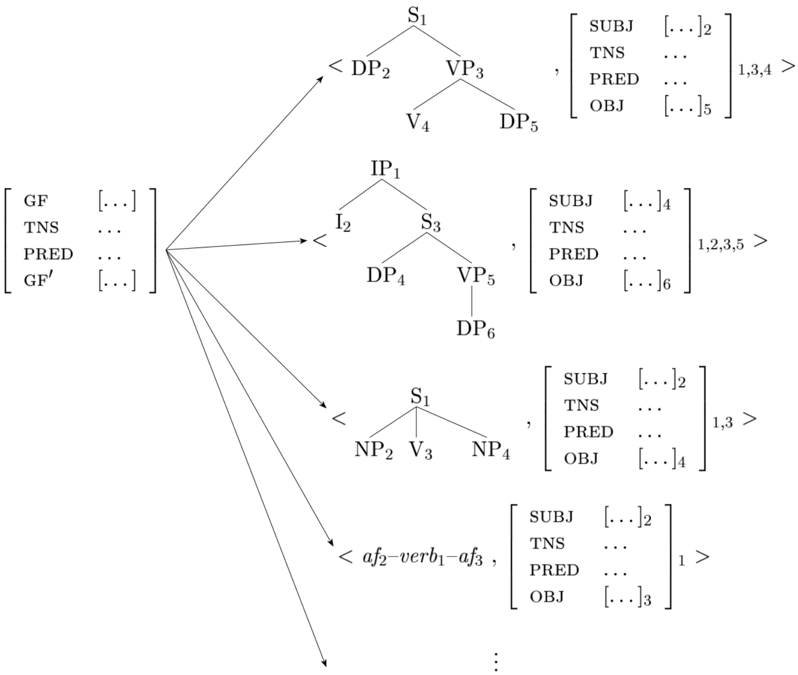


Figure 2: Illustration of candidate generation in OT-LFG (Bresnan 2001a): an underspecified functional structure (on the left) is the basis for generating full LFG analyses (c-structure/f-structure pairs) whose f-structure is subsumed by the input.

A formal definition for the candidate generation function *Gen* in OT-LFG is shown in (10) (Kuhn 2003: 74).[16] *Gen* depends on an LFG-type base grammar G_{inviol}, which specifies the space of structurally valid candidate tuples from the full typological spectrum by means of the 'inviolable principles', and is applied to a partially specified f-structure Φ_{in}.

(10) Restricted definition of *Gen*

$$Gen_{G_{inviol}}(\Phi_{in}) = \{ \langle T, \Phi' \rangle \in L(G_{inviol}) \mid \Phi_{in} \sqsubseteq \Phi', \text{ where } \Phi' \text{ contains no more semantic information than } \Phi_{in} \}$$

The candidate set generated by *Gen* includes all c-structure/f-structure pairs from the base grammar whose f-structure Φ' (i) is subsumed by the input f-structure

[16]The language $L(G)$ generated by an LFG grammar G is here defined as the set of tuples $\langle T, \Phi \rangle$ such that T is a c-structure generated by the context-free skeleton in G and Φ is a valid f-structure for T according to the functional descriptions in G.

Φ_{in} and (ii) does not add any semantic information.[17]

The violable constraints in OT-LFG operate on the candidate representations. It is in principle conceivable to include the detection of constraint violations already in the specification of the base grammar G_{inviol}, using disjunctive f-annotations of the kind shown in Section 1.2 on soft constraints in computational grammar development and introducing marks to a special *o*-projection for counting. However, the specification would presumably become unmanageable fast, since the representational patterns addressed in the various constraints are interdependent; moreover, many constraints address not only patterns in f-structures, but make reference to portions of c-structure. The formalization of OT constraints in (Kuhn 2003: ch. 4) therefore proposes to leave constraint marking out of G_{inviol} and assume a conceptually separate step for detecting the constraint violations for each c-structure/f-structure pair in the candidate set. This step can employ straightforward descriptive constraint *schemata* which use a special metavariable $\star$ that is successively being instantiated to every structural element in a candidate analysis, i.e. to each c-structure node, and to each (sub) f-structure.

In (11), the schema-based formal capturing of Bresnan's (2000) constraints Op-Spec and Ob-Hd, which were discussed in Section 1.1, is seen (for details, see Kuhn 2003: 90ff):[18]

(11) a. Op-Spec ('An operator must be the value of a df in the f-structure.')

$(f\text{-}str(\star) \wedge \exists v.[(\star\ \textsc{op}) = v])$
$\rightarrow \exists f.[(f\ \textsc{df}) = \star]$

"If $\star$ is an f-structure bearing a feature op (with some value v), then there is some f-structure f such that $\star$ is embedded in f under a df function."

b. Ob-Hd ('Every projected category has a lexically filled extended head.')

$(cat(\star) \wedge (bar\text{-}level(\star, 1) \vee bar\text{-}level(\star, 2))$
$\rightarrow \exists n.[ext\text{-}hd(n, \star)]$

"If $\star$ is an X-bar or X-max category, then there is some node n which is the extended head of $\star$."

The evaluation function *Eval* is operationalized as a trial of all schemata in the constraint set C on every structural element of a candidate analysis (c-structure node or partial f-structure). For the trial, the metavariable $\star$ is instantiated to the

[17] The restriction excluding surplus semantic information will be addressed in Section 2.4.

[18] The schemata assumes appropriately defined auxiliary predicates *cat* (for category) etc.

element under consideration; the count for the relevant constraint is increased in case the proposition is satisfied. The language-specific ranking $\gg_L$ over the constraints then controls the filtering of the most harmonic candidate(s), as was illustrated in Figure 1 in Section 1.2.

All candidates that are optimal (under the constraint ranking for a given language L) for some underlying partial f-structure Φ_{in} then form the language (= set of c-structure/f-structure pairs) generated by an OT-LFG system for language L (Kuhn 2003: 117):

(12) Definition of the language generated by an OT-LFG system
$O = \langle G_{inviol}, \langle C, \gg_L \rangle \rangle$ for language L:

$$L(O) = \{ \langle T_j, \Phi_j \rangle \in L(G_{inviol}) \mid \exists \Phi_{in} : \langle T_j, \Phi_j \rangle \in Eval_{\langle C, \gg_L \rangle}(Gen_{G_{inviol}}(\Phi_{in})) \}$$

The OT-LFG formalization thus provides a declarative, fully operationalized framework for specifying a theory of the typological space of options for a range of grammatical phenomena, predicting specific realization patterns for languages associated with certain constraint rankings. Computational considerations, in particular regarding the complexity of *Gen*, will be addressed in Section 2.4.

Combining the OT concept of competition-based specification of grammatical knowledge with the LFG formalism for operationalizing the components of such an account was beneficial for both sides: For theorists interested in exploring the expressiveness of the Optimality-Theoretic approach when applied to phenomena from the broad field of morphology/syntax/semantics, the OT-LFG framework provided a basis whose formal and computational properties were well understood, such that (i) concrete accounts for phenomena of interest could be worked out and tested against attested linguistic data, and (ii) formal and computational implications of the novel conceptualization of modeling grammatical knowledge could be pinpointed. In the reverse direction, the LFG community benefited from the extension of the classical formalism, since OT-LFG provided a systematically constrained framework for predicting patterns in the variation observed across languages (and within a language), grounded in a learning algorithm that realistically captures the situation of learner reacting to the observed language behavior of competent adult speakers. By couching an LFG account for a phenomenon in the OT-LFG framework, its generality becomes testable against empirical data from language typology. And even when trying to capture variation patterns within a single language, an OT-LFG approach readily supports a reasoning that reduces the range of available realization options to a small set of independently justifiable directives.

2.3 OT-LFG's conceptual advantages over a derivational base formalism

A large proportion of work in OT syntax is not couched in the OT-LFG framework, but assumes a derivational base formalism such as Principle-and-Parameter Theory or Minimalism (see Grimshaw's (1997) original account and much subsequent work, e.g. Pesetsky 1998). With candidate analyses that are inherently derivational, an implementation of the original OT idea – that the input corresponds to the semantic content of the expressions under consideration (for which the OT system differentiates between distinct acquired language-specific realization strategies) – inevitably leads to a more complicated architecture. In the study of syntax, the expressions under consideration are full sentences; so, to implement the original OT idea, all alternative surface sentences expressing the same semantic content should be included in a candidate set. The complication for an OT architecture now arises from the fact that in a Chomskyan derivational approach, semantic interpretation is located at the level of logical form (LF), which is by definition one of the end points of a derivational process that starts out from a D-structure (as in the T-model underlying Principle-and-Parameter Theory, shown in Figure 3) or a set of lexical items, the "numeration".

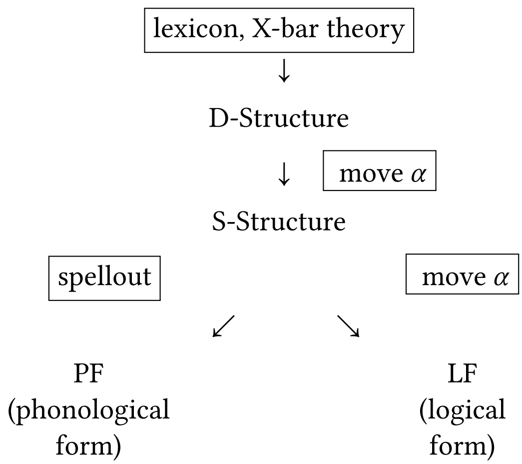


Figure 3: The T-model (or Y-model) of the derivation processes underlying a single analysis couched in Principle-and-Parameter Theory.

In a non-OT framework, the derivational processes that lead to a phonological form and a logical form are controlled by language-particular factors. When the derivational processes are taken to be the candidates of an OT system that adheres to the richness of the base, the derivational mechanisms have to be opened up in such a way that all language-particular restrictions on the derivations are lifted – generating all conceivable variants as alternative candidate derivations and leaving the calculation of language-particular effects to *Eval*, which selects the optimal derivation based on the constraint ranking. The input (for the definition of the OT candidate set) has to be (some relevant part of) an LF representation, and the set of candidates that *Gen* assigns to such an "input" LF must be all derivations that *end in* this LF via unconstrained derivation – from any possible starting point (i.e. any numeration of lexical items that could arrive at this LF through unconstrained transformational derivations). Leaving aside concerns regarding the computational tractability of such a system, it is challenging to conceptualize the workings of the OT constraints if the candidate-internal derivation is taken literally as a (potentially destructive) structure-transforming process: if the evaluation of candidate derivations does not take place until a particular LF has been reached, how can the constraints apply to derivation steps happening early on in the process?

The constraint evaluation challenge can be resolved by viewing the derivational process inherent to the candidates as some abstract process that produces a *representation* including a record of all relevant steps (such as traces of a movement). With this representational strategy, a definition of the candidate set via a shared LF becomes possible. Let us call this an LF-as-input OT system, which preserves the original, meaning-related concept of the OT input (but enforces an abstract view of the derivations, with a representational record).

For a research approach starting out with a derivational framework, it may seem more natural however to resolve the constraint evaluation challenge in a different way: giving up the fully meaning-related notion of the OT input, one can turn to a different way of characterizing the set of competing candidates: since the derivations have their own technical starting point – D-structure or a numeration of lexical items – why not adopt a conceptualization in which this part of the candidate derivations constitutes the input for the OT system and have evaluation compare the various possible derivations with identical D-structures/ numerations? Important typological considerations regarding syntax and morphology can be addressed under this view. In fact, when a study focuses on a small set of specific phenomena – as is common practice both for typological and for language-particular studies in linguistics – the two approaches are often indistinguishable, since the (supposedly) *relevant* variation across candidates on which the study is centered originates from the same derivational subprocess across all candidates. This is presumably the explanation for the fact that the simpler system is often tacitly assumed. However, in situations where attention is not focused on a narrowly delineated range of phenomena, the two ways of conceptualizing the input *do* make an enormous difference. Only a semantically based input will leave the choice among realizations that differ in the lexical material in non-trivial ways to *Con* and *Eval*.[19]

A numeration-as-input approach generally imposes restrictions on the candidate space that are tied to particular languages (since semantically equivalent paraphrases of the same content which use different lexical material do not compete in the same candidate set). It is actually a consequence of the predominance of the numeration-as-input approach (and similar input conceptualizations) that *language-particular ineffability* was considered a central issue for OT accounts of syntax. The issue can be characterized as follows: Since there will always be a most harmonic candidate in any given candidate set, the OT approach appears to systematically exclude that in some languages there is no grammatical realization at all for a certain linguistic construction. Fanselow & Féry (2019) provide the example in (13). While (13a) is grammatical in English, there is no grammatical way of saying (the equivalent of) (13b). They argue that a standard system of OT syntax as Grimshaw's (1997) will include a candidate set of alternative clause realizations for (13b), and one of them will inevitably be the most harmonic, incorrectly predicting that there is some grammatical realization used this set of lexical items.

(13) a. Who did the president think that the foreign minister met in Afghanistan?
  b. * Who did the president resign although the foreign minister met in Afghanistan?

Note however that with a sufficiently abstract semantic input, an OT system *could* be devised that makes reasonable predictions for all languages: Where there is no compact realization of a thought in a single clause, a more verbose, multi-clause paraphrase can be used (for (13b) maybe *Who was it that the foreign minister met in Afghanistan when the president resigned nevertheless?*). This appears like a plausible analogy to a very unfaithful realization of loan words or foreign names that fall outside the phonological patterns of a language (as exemplified by (9) above). Thus, the language-particular ineffability problem arising in

[19] Of course, not all accounts incorporating a competition-based subprocess are necessarily following the idea that all cross-lingual variation should be reduced to a global, fully meaning-based optimization. It is also conceivable to construe distinct derivational steps as separate, self-contained optimizations (compare e.g. Heck & Müller 2000; Müller 2003). Many approaches explicitly couched in a derivational setting indeed assume fairly restricted structural or derivational *domains* of local optimization (Müller 2012: sec. 4).

numeration-as-input approaches is not a problem under a global meaning-based conceptualization of the input.

The considerations in this subsection have shown that while a competition-based construal of dedicated derivational subprocesses may be fruitful in order to systematically derive typological patters for a certain submodule of a broader derivational theory of grammar, the original OT idea of deriving all cross-lingual differences – and thus learnability – as an effect of constraint reranking presupposes a comprehensive global competition among all alternative candidate expressions for a given underlying meaning. For capturing constraint evaluation in global competition with a derivational base formalism, there does not seem to be a good alternative to using representational traces of the (abstract) derivational process inherent to each candidate representation.

Bresnan's 1996 recasting of Grimshaw's (1997) account of extended projections – employing Extended Head Theory as sketched in Section 1 – can be viewed as a blueprint of a strategy that translates some relevant key aspects of a derivational approach into a representational approach. So, one can effectively view OT-LFG not only as an OT extension of the LFG framework; it also provides a feasible implementation for LF-as-input approaches, in particular a whole range of work that follows the general spirit of Grimshaw's (1997) proposal.

To sum up the previous and this section, OT-LFG as proposed by Bresnan 1996, 2000, 2001a, 2002 spells out a conceptualization of OT syntax that allows for a clean and comprehensive separation of language-independent candidate generation and violable constraints capturing the spectrum of typological variation. LFG's representational framework makes this separation conceptually simple. Global competition, for instance between morphological vs. syntactic means of expressing for realizing an underlying feature is naturally accounted for since LFG assumes parallel correspondence among all representational structures.

2.4 OT-LFG: Computational considerations

Conceptually straightforward as it is, the idea of relying on feature structure subsumption to define the set of candidate analyses (i.e. tuples of c-structure, fully specified f-structure and potentially more LFG projections) in *Gen* as specified in Section 2.2 raised computational concerns. Johnson (1998) points to the issue that the parsing problem for an OT-LFG grammar is undecidable in the general case. To solve the parsing problem given a string *s*, all optimal analyses have to be found that have *s* as the yield of their c-structure tree.[20]

[20] Johnson (1998) assumes stronger conditions for the optimal candidate: the input representation determined in the first step needs to be included in the optimal analysis of the string. This

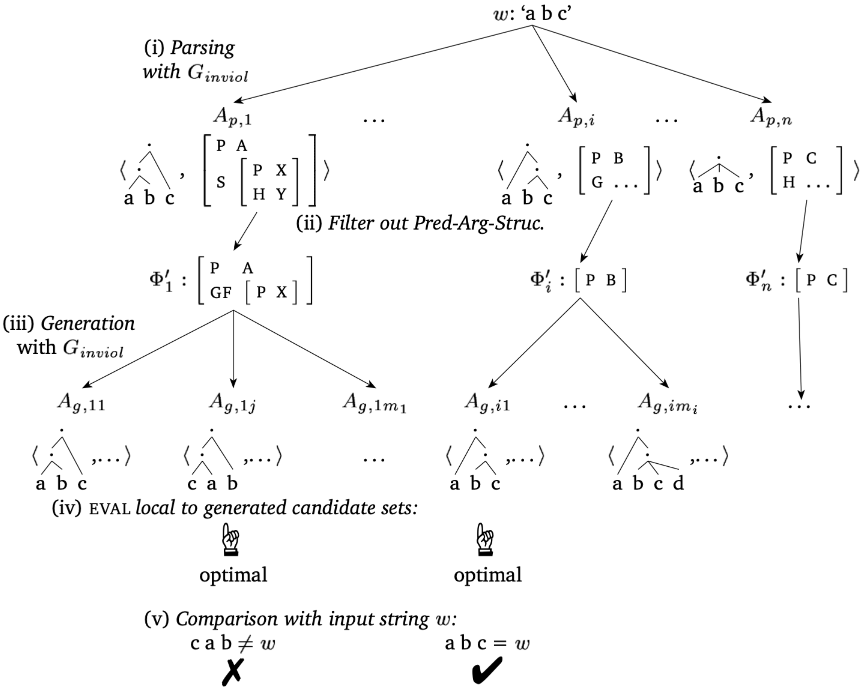


Figure 4: Illustration from Kuhn 2003: 173 for the parsing procedure for a standard OT system, with grammaticality based on production-based optimality.

Figure 4 from Kuhn 2003: 173 illustrates the procedure with a semi-abstract example. In a first step (i) all potentially underlying input representations have to be found. This can be achieved by standard LFG parsing (which is decidable), using the base grammar that defines the set of all universally available candidate structures. The input information is by definition part of the f-structure in each candidate. This predicate-argument structure representation has to be filtered out (step (ii)). In a next step (iii), for each potential underlying input representation, the set of all candidates has to be generated (including the original as well as all alternative full c-structure/f-structure tuples); the constraints are applied to each candidate (iv) and the most harmonic candidate can be determined, given the relevant constraint ranking. This candidate is only included in the set of valid analyses for the original string *s* if *s* is indeed the yield of the optimal candidate (v).

corresponds to (strong) bidirectional optimality, as discussed in Section 4.4. For standard OT, only the production-based competition is relevant, as assumed in the remainder of this section.

We note that step (iii) involves generation with an LFG-type grammar from a partially specified f-structure. Wedekind 1999 shows that the general problem of generation from partial f-structures, given some LFG grammar, is undecidable. This implies that there could be cases in which the candidate set for a given input f-structure cannot be computed, so it would also be impossible to determine the effect of an OT system. However, Wedekind also points out that the decidability problem for generation with plain LFGs occurs only with certain technical feature representations that are not used to represent the semantics of natural-language sentences. How does this translate to the application of an LFG-style grammar in OT candidate generation? Is it possible to formulate restrictions on admissible partial f-structures and thus guarantee decidability of the parsing problem?

Potentially problematic cases are candidates that include violations of faithfulness constraints, as they are for instance assumed to derive *do*-insertion in English questions (Bresnan 2000: (44)). As we will see in the following, restrictions on the formalism can be devised that guarantee decidability but nevertheless permit the use of faithfulness constraints to derive syntactic variation of this kind. Taking advantage of the explanatory potential of OT, Bresnan's 2000 analysis does not stipulate a special, PRED-less lexicon entry in *do* insertion (like in standard LFG), but derives insertion of an additional verb as a consequence of the ranking of violable constraints. To achieve this effect, the *Gen* function underlying in this system has to be able to add "unfaithful material" quite freely. Does this mean that we are confronted with the decidability problem?

As Kuhn (2003: ch. 4) discusses, the option of adding unlimited amounts of material in a candidate is not only undesirable from a computational point of view. It also goes against the key idea of treating all candidates as potential verbalizations of some identifiable semantic content. The definition of *Gen* should therefore be restricted in such a way that the candidates' f-structures (i.e. the interpretable part of the representation) are not only subsumed by the input f-structure, they also may not contain any additional semantic information. (It is allowed for candidates to contain additional non-semantic f-structure information and material at the level of c-structure and other projections; this is unproblematic for decidability, since the amount of information that can be added is bounded by the size of the grammar.) With this restriction, it can be guaranteed that the set of candidates stays computationally tractable (Kuhn 2002, Kuhn 2003: 199ff). The definition (10) above already incorporated the necessary restriction.

Figure 5 illustrates a formalization of faithfulness violations that is compatible with this restriction and can be used to derive *do* insertion in English. The key idea from Bresnan (2000: (44)) is that the lexical contribution of elements that are inserted "unfaithfully" (as a way of satisfying some other constraint that is higher-ranking than faithfulness) does not make it into the candidate's f-structure. In Figure 5, we see the technical formalization from Kuhn 2003: 112, which uses a special λ-projection to keep track of all lexical contributions from the words, which may or may not re-appear in f-structure. In case they do *not* re-appear, a violation of a faithfulness constraint (Dep-IO) will be recorded by *Eval*.

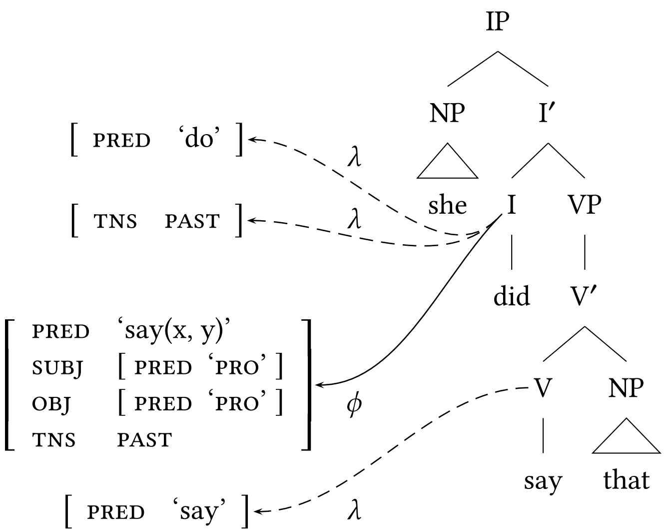


Figure 5: Illustration of a candidate incurring a faithfulness violation, following the technical formalization of Kuhn 2003: 112; c-structure nodes project not only to f-structure (via function ϕ), but also to a (very local) lexical structure (via function λ).

The lexical entries underlying the analysis in Figure 5 are specified as shown in (14) (adapted from Kuhn 2003: (140)):

(14) *did*   I $g_1 = \lambda(M*)$
$(g_1\ \textsc{pred}) =$ 'DO'
$\{ g_1 = \uparrow \}$
$g_2 = \lambda(\mathcal{M}*)$
$(g_2\ \textsc{tns}) = \textsc{past}$
$\{ g_2 = \uparrow \}$

say	V	$g_1 = \lambda(\mathcal{M}*)$
		$(g_1\ \textsc{pred}) =$ ‘SAY’
		$\{ g_1 = \uparrow \}$

In these entries, the functional annotations make use of local metavariables (such as g_1 and g_2) in the following way: The equation $g_1 = \lambda(M*)$ defines g_1 as the variable name for an attribute-value structure that is λ-projected from the current element's c-structural mother (= the node I for *did*). The equation $(g_1\ \textsc{pred}) =$ ‘do’ introduces a PRED feature and value into the attribute-value structure. The third equation $\{ g_1 = \uparrow \}$ identifies the local l-structure with the c-structural elements normal f-structure (i.e. its ϕ-projection; recall that $\uparrow$ is defined as $\mathcal{M}*$). What is crucial is that the third equation is enclosed in curly brackets, which means that it is applied optionally. Hence, the lexical specification will either be included in the f-structure or not. (Since the rest of the LFG structure leads to an identification of the f-structure projected from *did* and from *say*, maximally one of the two entries can introduce their PRED value to f-structure.)

The enormous variation opened up by this optionality is controlled by the faithfulness constraint Dep-IO, which in this setup can be formalized as in (15):

(15) Dep-IO (referred to as Fill in early OT work)
General OT formulation (Kager 1999: 68): ‘Output segments must have input correspondents.’
$(\textit{atomic-f-str}(\star) \rightarrow$
$\forall n, P.[(\textit{cat}(n) \wedge \textit{featrure-path}(P) \wedge (\lambda(n)P) = \star) \rightarrow (\phi(n)P) = \star])$
“For all categories n and feature paths P, if $\star$ is an atomic value under P in the λ-projection from n, then $\star$ is also the value under P in the ϕ-projection from n.”

Recall that in the application of the *Eval* function, every structural element in a candidate representation – i.e. every c-structure node, every local f-structure, and also every local l-structure – will be tested for all constraints. For the top-most l-structure seen in Figure 5, $\star$ will be instantiated to the PRED-value ‘do’, which *is* an atomic f-structure and which is also in the λ-projection of a node n (namely I) under a path p (namely PRED). So to satisfy the Dep-IO faithfulness constraint, PRED ‘do’ should also be in the f-structure for I, which it is not. This leads to the desired effect of capturing an insertion in c-structure which has no correspondence in f-structure as a Dep-IO violation.

To sum up, the insertions that are required to implement a generalized OT account are compatible with the restricted definition of *Gen*. By definition, the semantic information in a candidate's f-structure is always the same as the semantic information in the underspecified input f-structure Φ_{in}. Additional material can only occur in locally projected l-structures. While it is not forbidden to have infinitely many distinct candidate structures for an input f-structure, the restriction keeps the candidate set computationally tractable (Kuhn 2003: ch. 6). (Exploiting results from Kaplan & Wedekind 2000, recursive loops that lead to infinite candidate sets can receive an aggregate treatment.)

2.5 Excursion: The cognitive status of "directional" candidate set specification

As established in the previous subsection, it is possible to provide a declarative formal characterization of the language generated by an OT system. Thanks to the non-derivational character of (tuples of) LFG structures, it is possible to use the sharing of the semantic part of the structures as the defining element for candidate sets, independent of the question of a procedure could be implemented which applies the OT system in recognition or parsing. Nevertheless, to many LFG practitioners, the computational breakdown of the parsing task for an (LFG-based) OT system sketched in Figure 4 does not seem intuitively appealing. While – as probably everyone admits – LFG's purely description-based specification of representation structures makes the characterization of the OT candidate set very perspicuous, the highly intuitive algorithmic breakdown of parsing with plain LFG does not carry over to parsing in the OT-LFG framework, which necessarily has to incorporate a production-based directionality, while parsing models comprehension. Parsing with a plain LFG grammar follows the logic of the structure specification step by step: given a string, matching lexicon entries and c-structure rewrite rules are used to construct a set of c-structure trees spanning the string. To narrow the set down to the trees which correspond to a valid f-structure, functional descriptions from the lexical annotations and the rule annotations are then taken into account in a process of model construction (in a feature logic), and wherever we find a valid f-structure, we have one possible analysis for the input string. The intuitive simplicity of the algorithmic breakdown has presumably played a role in the attraction of the LFG formalism as a psychologically realistic framework – in particular as it contrasted with the model of Chomskyan derivational approaches, which work with a notion of an underlying deep structure (D-structure or numeration). The theoretically motivated transformation of an underlying deep structure into a surface structure does capture intuitions regarding the highly systematic relationship between expressions like *the cat drinks milk* and *what does the cat drink?* or between *she saw* *the cat* and *the cat was seen*. But in comprehension, a listener is confronted with the linear string from a surface structure (say, *the cat was seen*), and there is no cognitively intuitive algorithmic process that leads back to conceivable underlying deep structures. When spelling out a parsing procedure with OT-LFG, we now find ourselves in a similar situation: if we translate the bidirectional characterization of the set of valid structures for a given string into an algorithmic procedure, we do not arrive at a plausible rendition of what could be going on in comprehension.

One might suspect that these considerations challenge the cognitive plausibility of Optimality Theory. It should be noted however that there is an oversimplification in the reasoning that assumes a direct conceptual mapping of a declarative specification of some function (such as the function from a string of words to a set of c-structures associated with valid f-structures) to the seemingly straightforward algorithmic breakdown of this function. It is misleading in the general case to assume that an intuitively appealing translation of a composite function into some procedure is the only option for realizing the theoretically motivated function in a cognitive system. As a matter of fact, there would be no computationally tractable parsers for standard LFG if one relied on the simple-most translation of the conceptual steps underlying grammatical specification in LFG into an algorithm. As the discussion in Maxwell & Kaplan 1996 illustrates, LFG parsing performs a highly sophisticated interleaving of the various sources of grammatical and lexical knowledge. Vice versa, Edward Stabler's work on the formalization of Chomskyan derivational grammar models (e.g., Stabler 2011, 2013) shows that algorithmic solutions do not have to stick to the seemingly counterintuitive procedures in the underlying theoretical characterization.

The somewhat counterintuitive direct algorithmic implementation of parsing with OT systems is in fact a consequence of incorporating an important observation that most linguists will presumably share into the design of the definition of grammaticality (rather than leaving it as a background assumption that is not enforced by the formal system): for a speaker of language X to know the grammatical way of expressing something in X amounts to knowing which other *potential* ways of expressing the same thing are not available in X. During the acquisition of X, the speaker will have learned from exposure to adult speakers' language behavior which realizational variants can be completely excluded. So, the adult speaker's grammatical knowledge may well be thought of computationally as a "hard-wired" input-output mapping that freezes the patterns which have stabilized in the competition system (superseding a dynamic acquisition phase, during which the relevant constraint rerankings were triggered for language learner).

As a side note, the "pre-compilation" of a cascade of optimality constraints with their step-by-step filtering effect into a single input-output function was subject to research in the context of OT systems that can be fully formalized with finite-state systems. Karttunen 1998 proposed a special finite-state operator (so-called lenient composition) that has the effect of turning a sequence of individual constraints formulated as transducers into a single transducer with the same effect as an OT competition.[21] To achieve a similar computational effect for a syntactic OT-LFG system, the internal data structures built up during parsing (following a chart-based or dynamic programming approach) would have to be re-designed to simultaneously incorporate a production-based and a comprehension-based directionality – which could then be instantiated in a single bottom-up algorithmic pass, avoiding direct bidirectionality processing (compare the "interleaved" bidirectional processing approach proposed in Kuhn 2000b). Since work in computational linguistics focusing on learning without any language-specific prior knowledge had already reached a relatively advanced state using different approaches, there was no substantial practical interest in putting the full OT-LFG account to use on a larger scale.

2.6 OT-style constraint ranking in broad-coverage grammars

Section 1.2 provided an illustration of the constraint ranking mechanism implemented in the XLE system, which is widely used in grammar development. The examples showed how violable constraints allow grammar writers to include rule variants for infrequently occurring constructions without causing a proliferation of implausible readings for canonical constructions. (Computationally, we can note in the light of Section 2.4 that XLE's OT-style constraint ranking is not normally used to modify the notion of grammaticality defined by the base grammar; hence, the approach avoids the additional complexity of a two-way application of the grammar in parsing and generation mode.)

The XLE implementation of violable constraints via a special *o*-projection provides some further mechanisms that are of high practical value in broad-coverage grammar development. The specification of the ranking of the optimality marks in (16) illustrates some of these mechanisms. We will shortly go through number of details, but first of all we note that the ranking is specified in the configuration section of the grammar code. This means that the relative ranking of the marks (and hence of the soft constraints) can be adjusted for different application scenarios of the grammar without changing the grammar code itself. Thus the

[21] A similar approach is discussed in Frank & Satta 1998, a.o.

ranking specification can be used to flexibly adjust the grammar to peculiarities of certain language registers or text domains.

(16) OPTIMALITYORDER +PPasOBL (UnderstoodObj TypeCoercionToTimeNoun) NOGOOD Missing3SgMarking STOPPOINT Fragment.

In (16), the highest-ranking mark PPasOBL is preceded by a plus sign. This indicates that in the filtering of analyses, the mark is not considered to be negative, but positive. When the available readings differ in the count of PPasOBL, the ones with the *maximal* number of marks survive. This provides grammar writers with a way of giving preference to a certain variant rather than having to "punish" a different variant, which, depending on the feature representation adopted, may may be impractical. Rule (17a) exemplifies the introduction of PPasOBL as a preference mark for the oblique object analysis of PPs (like in *wait for someone*). The oblique reading will be preferred over the alternative of analyzing the PP as an adjunct whenever both are available. This has the effect of reducing the number of parses for a considerable number of input sentences. In real-world uses of grammars, the available information makes it often hard to make an informed decision.[22]

(17)
$$\text{VP} \longrightarrow \underset{\uparrow=\downarrow}{\text{V}} \quad \left(\underset{(\uparrow \text{OBJ})=\downarrow}{\text{NP}}\right) \quad \underset{\substack{\{\ \downarrow \in (\uparrow \text{ADJ}) \\ |\ (\uparrow \text{OBL}_{\theta})=\downarrow \\ \text{PPasOBL} \in o(*)\ \}}}{\text{PP*}}$$

The ranking in (16) also illustrates the workings of parentheses and of some specially defined keywords. The marks UnderstoodObj and TypeCoercionToTimeNoun, which were introduced in Section 1.2, are now jointly enclosed in parentheses. The effect of this is that they are treated as equally ranked. In grammar writing practice, such constraint ties are frequently used for phenomena that

[22]Contrary to the situation with very infrequent constructions that were discussed in Section 1.2, the filtering may here quite often have the effect that a contextually inappropriate analysis is chosen over the more appropriate analysis (for instance in *Sue was waiting for hours*). But in a range of applications of the grammatical analysis, this may not be too problematic, while a reduction of the sources of ambiguity can be extremely helpful during the process of extending the grammar or fixing a certain problem. When the grammar is used in application scenarios in which it can be harmful to occasionally choose the contextually inappropriate variant within a systematic pattern of ambiguity, the preference mark can simply be taken out of the ranking specification, so the parser will output both variants and leave the decision to downstream processing components.

are independent from each other, since there is no grammar-internal justification for giving preference to one of them. We will come back to the disambiguation decisions that are left undecided by the constraint tie strategy at the end of this section.

To the right of the two we see the mark NOGOOD. This is a predefined keyword that has the effect that all marks that follow receive a special interpretation. Consider Missing3SgMarking: This mark is introduced in the definition of the template SubjNon3SgAgr in (18). This template is used in present tense verb forms of English like *I laugh* or *they laugh*. There is a similar template for third-person singular forms like *she laughs*. The third disjunct in (18) covers the use of the form *laugh* with a third-person singular NP like in **she laugh*, which is ungrammatical in standard English. By providing this option, the grammar will robustly cover agreement mistakes or it can be used for varieties of English that include this variant.

(18) SubjNon3SgAgr ≡ { (↑ subj num) = pl
| (↑ subj num) = sg (↑ subj pers) ~= 3
| (↑ subj num) = sg (↑ subj pers) = 3
Missing3SgMarking ∈ *o*(*) }

With the ranking specified as in (16), *she laugh* will receive an analysis; however due to the use of the NOGOOD mark, it will be labeled as ungrammatical by the parsing system.

A last point illustrated by (16) is the predefined mark STOPPOINT. This mark offers a way for grammar writers to control the computational behavior of the parser. Putting an optimality mark to the right of STOPPOINT (like Fragment in the example) has the effect that the functional descriptions that it marks will not be used at all in the first pass of running the parser. However, if parsing leads to an empty set of valid analyses, the parser will be reset for a second pass, this time including marks like Fragment. This mechanism can be used to make the grammar more robust (by adding analysis options that will fire when everything else fails) without compromising the runtimes of the parser for "well-behaved" input sentences. A typical example for using STOPPOINT in the ParGram grammars is a special S rule that combines c-structure fragments such as NPs, PPs and certain incomplete verb projections, to ensure that problematic input sentences (e.g. with misspellings or rare constructions) will still receive an analysis for the parts that are covered correctly. Each fragment will introduce one Fragment mark, so the filtering mechanism will output the option with the fewest (i.e. on average the largest) fragments. It is even possible to use several instances of STOPPOINT in a ranking to potentially trigger several resets.

As the various mechanisms we briefly discussed show, a flexible ranking for specially marked parts of the grammar conveniently puts grammar writers in a position of exerting control over the set of valid structures that an LFG grammar will assign to a given input string in parsing (and similarly for a given input f-structure in generation).[23] Of course, there are other ways of controlling the space of ambiguities, most notably probabilistic techniques, which have been widely applied in the context of structural analysis of realistically occurring language (Collins 1997; Riezler et al. 2000, 2002). The standard strategy in the probabilistic approach to the ambiguity problem is to rely on supervised training of probabilistic models that predict the distribution of alternative linguistic representation structures, dependent on a large set of contextual factors. If a sufficiently large training corpus is available that was manually disambiguated by competent speakers of the language, a so-called treebank, the complex interaction of many distinct knowledge sources on the contextually appropriate choice of readings can be captured quite reliably thanks to statistical patterns. When the texts used in application are similar enough to the training corpus, the performance that can be achieved is higher than in a knowledge engineering approach of classical grammar writing, since statistically relevant patterns of all kinds (e.g. word order preferences, lexical-semantic argument selection preferences, but even statistical effects unrelated to grammatical knowledge) are learned "in passing".

The XLE system offers both the optimality ranking approach discussed in this section and a probabilistic filtering approach that relies on supervised treebank training. In the practice of broad-coverage grammar development with a highly expressive formalism such as LFG, both mechansisms have their place and a combination is arguably the most effective way to go: a probabilistic approach exploits the empirical distribution of interacting factors, such that a sufficiently expressive probabilistic formalism (or machine learning model) will induce implicit statistical knowledge even about patterns that could not (yet) be captured in symbolic terms (see e.g. Cahill et al. 2007; Forst 2007). On the downside however, a plain probabilistic approach leaves little leeway for grammar writers to inject specific symbolic knowledge about certain constructions. By using symbolic optimality ranking as a pre-filter for the set of candidate analyses going into treebank training, the grammar writers can easily experiment with alternative strategies (King et al. 2000).

[23] One way of looking at the addition of rankable constraints to an LFG grammar writer's means of expression is to include ideas from grammatical frameworks that never factorized the task of disambiguation out, most notably Constraint Grammar (Karlsson 1990).

As a final remark it has to be noted however that in this combined setup, the grammar writer is not in a very informed position to determine the relative ranking among the optimality constraints for multiple different linguistic constructions – for instance UnderstoodObj and TypeCoercionToTimeNoun from Section 1.2. This is something that an empirically informed training procedure can do better (reaching the statistically most effective ranking for a given treebank). The utility of symbolic soft constraints for linguistically informed ambiguity management lies more in the flexibility of experimenting with preference, dispreference and delayed exection (via one or more STOPPOINT mark) of constraints. By leaving the ranking of OT marks within a section (before NOGOOD, between NOGOOD and STOPPOINT, etc.) very flat through the use of parentheses, the ranking decision is postponed to the subsequent probabilistic filtering module.

There are a number of publications that report on the use of ranked constraints in various contexts of grammar specification, for instance Zaenen & Crouch 2009; Bögel et al. 2009; Dione 2014.

3 Linguistic applications of the competition-based concept of grammaticality

Section 2.6 provided details on the OT-style ranking mechanism that offers very effective functionality for ambiguity management in broad-coverage grammar development. We now go back to the original, theoretically motivated OT syntax model that employs competition among candidates to determine the grammatical way of expressing an underlying meaning. The model has been broadly employed to (a) systematically derive variation patterns within a language and (b) predict typological patterns across languages.

A broad range of syntactic phenomena have been addressed with OT syntax approaches. We will take a closer look at a few accounts in this and the subsequent sections – mostly to illustrate some specific properties of OT systems, in particular in the guise of OT-LFG and extensions that have been proposed. A full overview of all important phenomena addressed in the literature is beyond the scope of this chapter.[24]

This section discusses two important predictive schemes that the OT approach offers at the interface between syntax and morphology. Section 3.1 shows how

[24]Important areas excluded for space reasons are for instance positional alignment accounts of phrase structure such as Sells (1999, 2001). Sten Vikner and collaborators have worked out a detailed account of object shift in OT (Vikner 2001; Engels & Vikner 2014).

morphological blocking phenomena can be derived in a very general way. Section 3.2 reviews the harmonic alignment account of the typological spectrum of differences in argument linking.

3.1 Generalizing blocking accounts to incorporate morphology-syntax competition

LFG's system of corresponding parallel representations can be straightforwardly integrated in the competition-based grammaticality account of OT. This opens up a path for generalizing theories of morphological blocking that were developed before a generic mechanism of comparison was included in the overall formal framework. Prominent examples are the accounts by Andrews (1982, 1990), building on top of the Elsewhere Principle from phonology (Anderson 1969). The idea of morphological blocking offers an explanation of how within a morphological paradigm such as the inflection of English verbs in present tense, an unmarked form (like *laugh*) can fill all cells in which no explicitly marked form (like the third person singular form *laughs*) is available. When alternative forms are available that express different degrees of specificity with respect to certain morphosyntactic features, the existence of a marked form for some specific feature combination (or position in a morphological paradigm) blocks the use of an unmarked form for this particular combination: *laughs*, which is marked for PERSON and NUMBER (to the values 3 and SINGULAR, respectively), blocks the unmarked *laugh*, which is underspecified for PERSON and NUMBER. Technically, Andrews (1982) proposes a blocking condition which states that a less specifically marked form A cannot be used in a position X if there is a form B that comes with a more specific marking, subsumed by A's specification.

With a competition-based definition of grammaticality, blocking effects can be construed as a consequence of general constraint interaction (Bresnan 2002, 2001a): The unmarked form is assumed to incur faithfulness violations, since it does not explicitly realize the underlying feature information in the input. For each inflectional category, a faithfulness constraint (e.g. FAITH^{NUM} for number) checks whether a surface form accurately marks the underlying feature. On the other hand, a markedness constraint is assumed for each specific feature value, punishing the explicit marking of this value (e.g. *PL “avoid marking the plural explicitly”, *SG “avoid marking the singular”). The markedness constraints implement the tendency in natural language to keep expressions as concise as possible. On the basis of these two antagonist constraints, learning their relative ranking for a given language[25] has the effect of learning in which paradigm cells to use a

[25]The learner acquires the constraint ranking through exposure to output produced by adult

marked form vs. the unmarked form:[26] If *PL outranks FAITH$^{\text{NUM}}$ in a language, the plural (of the word class under consideration) is realized by an unmarked form in this language; under the reverse ranking, a marked form is used for plural. For fusional morphology, in which a single morpheme (like for instance the -*s* in English present tense verb forms) can realize person and number simultaneously, conjunctive faithfulness constraints to *sets* of inflectional categories have to be assumed besides faithfulness to an individual inflectional category, for instance FAITH$^{\text{PERS\&NUM}}$ (Bresnan 2001a: ex. (22)). The verb inflection paradigm for present tense in modern English can be predicted by the following ranking: *PL, *1, *2 ≫ FAITH$^{\text{PERS\&NUM}}$ ≫ *SG, *3 ≫ FAITH$^{\text{PERS}}$, FAITH$^{\text{NUM}}$. Plural as well as first and second person is never marked, since the markedness constraints for these feature values outrank all faithfulness constraints. For the combination of PERSON SINGULAR and NUMBER 3, a fusional form is used, since faithfulness to the combination of PERSON and NUMBER outranks the markedness constraints *SG and *3. For the other PERSON/NUMBER combinations, the fully unmarked form (*laugh*) is used, since FAITH$^{\text{PERS}}$ and FAITH$^{\text{NUM}}$ rank lower than the markedness constraints *SG and *3.

Given this characterization of the task of learning inflectional paradigms, the following typological spectrum is opened up by the interaction of faithfulness and markedness constraints: (i) When the markedness constraints outrank all faithfulness constraints, a paradigm with no inflectional distinctions follows; (ii) when faithfulness outranks all markedness constraints, all paradigm cells go along with an explicit form. (iii) Blocking effects occur when faithfulness is ranked in between certain markedness constraints. The account then predicts features (or feature combinations) whose markedness constraints outrank faithfulness to be realized by an unmarked form.

The OT-LFG framework makes it even possible to generalize the OT account of blocking to situations where it is not just alternative synthetic word forms that could be used to express an underlying feature bundle, but syntactically complex expressions are an additional alternative. Speakers of English have learned for instance when to use the analytical realization of a comparative adjective or adverb (such as *more quickly*) instead of a synthetic realizations (such as **quicklier*). In LFG's system of imperfect correspondence among parallel representations (Bresnan 2001b), such alternatives are just different surface realizations of the same f-structure. Now, the set-up in OT-LFG is to have such alternatives compete for

speakers; the speakers' underlying input has to be inferred from the situational context.

[26] To capture the finegrained differentiations inherent to inflectional paradigms, the faithfulness constraints have to be parametrized, for instance to specific verbs/verb classes for which learners have to learn distinct patterns.

the status of the most harmonic candidate. It is clearly possible for a theorist to find constraint sets that will lead to analytical realization of a phenomenon in one language an synthetic realization in another. This alone may not be considered a strong argument in favor of a competition-based framework using parallel representation structure like LFG. However, when it can be shown that having analytical and synthetic alternatives side-by-side in the candidate set for realizing an input (i.e. expanding the same partial f-structure) leads to a systematic explanation of variability in inflection paradigms that mix analytical and synthetic realizations – via a generalization of the morphological blocking effect – this constitutes persuasive evidence that the architecture of the theoretical account does capture aspects of the human cognitive system quite well. This is exactly what Bresnan (2001a) achieves by the account of negation in varieties of English she proposes.

The tableaux in Figure 6 illustrate competitions among different analytic c-structural realizations of negation; in Bresnan (2001a: ex. (43)), these tableaux serve to motivate the constraint set for an account in which an analytical form blocks a synthetic form. The analysis assumes two alternative realizations for the negation of verbs or auxiliaries in English: *not* can adjoin to the auxiliary itself, which is realized in I^0, or it can adjoin to the VP. (For the modal verb *can*, the orthographic rules of English happen to make the distinction visible in the written form.) By hypothesis, both alternatives can have the meaning of a wide-scope negation, but only the latter can mean negation of the VP. Bresnan assumes one markedness constraint for each of the two possible sites for adjoining negation, *NEG-VP and *NEG-I; in English *NEG-VP is ranked higher than *NEG-I. Faithfulness to the negation scope (i.e. the constraint $\text{FAITH}^{\text{NEG}}$) however is ranked higher than both markedness constraints. As an effect, the NEG-I option (*cannot*) arises as the optimal realization for a wide-scope reading of negation, whereas the more marked analytic form (*can not*) is required to express the VP scope of negation.

¬(POSS(work(he)))	$\text{FAITH}^{\text{NEG}}$	*NEG-VP	*NEG-I
☞ a. *he cannot have been working*			*
b. *he can not have been working*		*!	

POSS(¬(work(he)))	$\text{FAITH}^{\text{NEG}}$	*NEG-VP	*NEG-I
a. *he cannot have been working*	*!		*
☞ b. *he can not have been working*		*	

Figure 6: Two tableaux from (Bresnan 2001a: ex. (43))

For the realization of negated forms of the auxiliary *be* in various varieties of English, analytical forms compete with synthetic forms: the negated third person singular can be realized as *is not* or as *isn't*. Moreover, a synthetic form that is unmarked for person and number is available: *aren't*. Interestingly, although Standard English has a marked form for declarative first person singular (*am*), there is a lexical gap for the negated first person singular.[27] In negated interrogative clauses, this gap is – for many speakers – filled by the unmarked *aren't* (examples from Bresnan 2001a: (14)-(15)):

(19) a. * Am I not going?
  b. I am not going.

(20) a. Aren't I going?
  b. * I aren't going.

This effect can be derived in an OT-LFG analysis that assumes high-ranking markedness constraints punishing analytic negation adjoining either to C^0 (which would yield **Am not I going?*) or to VP (yielding (19a) (Bresnan 2001a: ex. (61)-(63)). These markedness constraints outrank the constraint $\textsc{faith}_{be}^{\textsc{pers\&num}}$, which regulates faithfulness to the person and number feature for the auxiliary *be*. A hypothetical synthetic form **amn't* marked for person and number is unavailable due to the idiosyncratic lexical gap in English that was just discussed.

How can the grammatical framework model that a person acquiring English learns about such an idiosyncratic gap? In the OT framework, it has been proposed to assume a constraint LEX parametrized for specific lexical material and incurring a violation whenever it is used. Learners of a language with an idiosyncratic gap will rank the respective LEX constraint above all other constraints (because adult speakers never use this material when one would expect them to, based on the context).[28] For the English speakers using (20a) rather than (19a) to fill the lexical gap in the interrogative case, the third analytical option, adjoining

[27]For synchronic learnability of such an idiosyncratic gap, it is not relevant how the gap came about. Bresnan (2001a: fn. 26) mentions stigmatization of an older synthetic form *ain't* as a potential explanation.

[28]Note that assuming constraints sensitive to specific lexical material in a language does not go against the principle of richness of the base, which excludes language-particular restrictions on the candidate set. However, it is not fully compatible with the assumption of a (finite) universal set of constraints. From the point of view of learning algorithms, it seems quite plausible however that instances of some constraint schema can be parametrized by lexical items that the learner has added to their inventory. See also van der Beek & Bouma 2004 for a discussion of language-particular lexicon properties within OT-LFG.

negation to I^0, is ranked lower than $\textsc{faith}_{be}^{\textsc{pers\&num}}$. This has the effect that the pattern in (19)/(20) is predicted: when *be* can be realized in I^0 as is the case in a declarative clause, the fully marked analytic form is the most harmonic; in a question however, where *be* is in C^0, the unmarked form *aren't* wins out.

This analysis demonstrates the explanatory potential coming from a competition-based account of grammaticality that makes distinct grammatical means available in candidate sets based on an input corresponding to the underlying content.

3.2 Harmonic alignment

In many languages, certain properties that argument phrases like subjects and objects can bear (e.g. first person vs. third person, full NP vs. pronoun, overt case marking, but also the choice of grammatical relation itself) are correlated with the availability of grammatical syntactic realization options, for instance in a clause with a transitive verbs. In the Australian language Dyirbal for example, the case marking patterns for transitive verbs are sensitive to such properties: "1st/2nd person pronouns are marked when they are objects, but not when they are subjects" (Aissen 1999: 674). When looking at the distribution of the relevant properties across languages, typologists have made the following observation: it is possible to organize these properties along markedness scales in such a way that the different scales tend to align with each other Silverstein 1976. In a relational scale, subjects are used for the most salient/central arguments, followed by objects for less salient arguments, followed by obliques. In terms of thematic roles, agents are more prominent than patients. Animacy hierarchies have first and second person pronominals at the top of the scale, followed by third person pronouns, common nouns referring to humans, to animate referents and finally inanimate referents. Aissen (1999, 2003) develops an influential OT syntax account[29] demonstrating that many fine-grained observations from typological studies can be explained when the following assumption is made: the OT constraints that make reference to the various markedness or prominence scales cannot be arbitrarily (re-)ranked, but there are universal subhierarchies that are imposed over families of related constraints. These subhierarchies, technically implemented by the mechanism of harmonic alignment, have the effect that the various different markedness or prominence scales are systematically aligned.[30]

[29] Aissen does not explicitly couch her account in an OT-LFG setting, but it is fully compatible and has greatly influenced subsequent OT-LFG work.

[30] The technique of harmonic alignment across prominence scales was already introduced by Prince & Smolensky (1993) for phonological features (sonority and syllable structure).

For certain pairs of constraints, the relative prominence is fixed *a priori*,[31] while their interaction with other factors can still be freely learned from the observations.

For instance, Aissen's (1999) account explains the split ergativity patterns in Dyirbal, where under specific conditions argument phrases are realized without case marking: the subject is generally unmarked when it is first or second person; the object is unmarked when it is third person. When the subject is third person, case has to be marked; likewise when the object is first or second person. (21) shows the OT subhierarchies that ensure the alignment of the relational scale and the person scale (combining first and second person as "local"). In essence, it is more marked to align a high element from one scale with a low element from another scale than to align either two high elements (Su/local) or two low elements (Ob/3rd).

(21) a. *Su/3rd ≫ *Su/local
b. *Ob/local ≫ *Ob/3rd

To capture the case marking patterns, each of the alignment constraints is locally conjoined with the constraint $*\varnothing_{\text{CASE}}$, which punishes expressions that do not use overt marking for the respective combination – similar to faithfulness constraints. Local conjunction (C_1 & C_2) of two distinct OT constraints C_1 and C_2 within a given local domain D is a mechanism that captures the fact that in certain cases, it can be more marked when the two constraints are violated within the same local domain, for instance the same argument phrase, than when there are two independent violations of C_1 and C_2 (Smolensky 1995). Since the local conjunction C_1 & C_2 can be ranked independent of the individual constraints, special markedness patterns that are sensitive to the conjunction can be learned.

In Aissen's 1999 account, the universal subhierarchies assumed for alignment constraints like *Su/3rd carry over to the family of their local conjunctions:

(22) $*\varnothing_{\text{CASE}}$ & *Su/3rd ≫ $*\varnothing_{\text{CASE}}$ & *Su/local

Learning the case marking patterns for a particular language amounts to learning where within the universal subhierarchy a structural markedness constraint for the relevant grammatical feature is placed in that language – here the constraint $*\text{STRUCT}_{\text{CASE}}$. In Dyirbal, $*\text{STRUCT}_{\text{CASE}}$ splits up both of the two hierarchies from (21):

[31] Zeevat & Jäger (2002) demonstrate that the effect of the subhierarchies may also follow empirically from a systematic skewedness in patterns of usage. To the extent that this skewedness follows from invariant aspects of human social interaction, etc., it is presumably hard to tell empirically whether *a priori* rankings should be assumed within in the language faculty.

(23) a. $^*\varnothing_{\textsc{case}}$ & *Su/3rd ≫ $^*\textsc{Struct}_{\textsc{case}}$ ≫ $^*\varnothing_{\textsc{case}}$ & *Su/local
b. $^*\varnothing_{\textsc{case}}$ & *Ob/local ≫ $^*\textsc{Struct}_{\textsc{case}}$ ≫ $^*\varnothing_{\textsc{case}}$ & *Ob/3rd

Hence, it is more harmonic to avoid using a structural case marking on an argument when this is the subject and first or second person, while for third person subjects the high ranking of the conjunction $^*\varnothing_{\textsc{case}}$ & *Su/3rd, excludes an unmarked subject in favor of a case-marked one.[32]

4 Extensions and debates

The competition-based definition of grammaticality in OT with its typological predictions attracted considerable attention in linguistic research communities, at the same time triggering debates about consequences of the new formal framework. Extensions were proposed to capture aspects of language(s) that the standard OT setup does not bring out. One such phenomenon is free variation, which one would expect not to exist under a plain OT approach. The stochastic OT framework, which does capture free variation, is discussed in Section 4.1.

Section 4.2 addresses a debate regarding the motivation of OT constraints. A substantial part of the OT community has been following the practice of providing a motivation for each constraint they assume which is grounded in functional considerations (for instance physiological considerations in OT phonology). This led to controversies, which are illustrative for the conceptual status different researchers assign to the constraints formulated in a theory of grammar. Section 4.3 steps back and discusses some of the cognitive considerations that led to the proposal of a competition-based account of knowledge of grammar in the first place. The section links this specifically to the question of learnability.

A second important extension of the basic architecture, bidirectional OT, is discussed in Section 4.4. It is motivated, *inter alia*, by the so-called phenomenon of word order freezing in languages with a (relatively) free order. In word order freezing, a clausal pattern that one would actually expect to be ambiguous *de facto* receives only one interpretation.

4.1 Extension I: Stochastic OT

One counterintuitive prediction of the competition-based definition of grammaticality is that languages should display very little (if any) free variation among

[32]We will come back to Aissen's harmonic alignment account in Section 4.2, as it forms the central target of Newmeyer's (2002) critique of functionally motivated constraint sets.

two equally grammatical ways of expressing the same thing – for example *we need a more catchy title* vs. *we need a catchier title*. Even when it is just some minor and low-ranking constraint in which the two realization options differ, the OT system will by definition predict one variant to be ungrammatical for the relevant input (unless the relevant constraint is tied with another constraint – which is not an option very often, since most constraints play a role in multiple constraint interactions). Most natural languages *do* however offer free variation for certain lexical or grammatical means – there is for instance a certain amount of free word order variation in many languages.

The circumstance that a standard OT system covering a non-trivial range of phenomena is extremely unlikely to ever predict free variation is concealed by the fact that linguistic studies commonly focus their attention on a particular family of phenomena. When isolating a specific set of constraints that is relevant for deriving the observations regarding these phenomena, the possibility of ranking two or more constraints at the same level for a given language (= a constraint tie) *does* create a basis for explaining the systematic occurrence of free variation. For instance, from Bresnan's (2001a) account of English auxiliaries discussed in Section 3.1 predicts variation among *cannot* and *can't*. However, for a more comprehensive account of grammatical knowledge, we have to assume that all the constraint sets posited for certain phenomena are combined in one larger constraint set. As a consequence, the effect of most constraint ties will go away – since the alternatives will differ in properties that are of relevance for some independent account (e.g. *cannot* has an extra syllable).

The problematic implications that standard OT has for free variation triggered several independent proposals for an extension of OT systems that will naturally predict free variation (see e.g. Müller 2014; Asudeh 2001). One of the most influential proposals is known as stochastic OT (Boersma 1998; Boersma & Hayes 2001).[33] It preserves most of the original OT architecture; the key modification is that the ranking of the OT constraints is no longer viewed as fixed and discrete, but (1) the rank of a given constraint is a value on a continuous scale, and (2) the constraints are assumed to oscillate stochastially around their (mean) rank. Hence, at the time of harmony evaluation for a particular OT competition, it is possible with a certain probability that two constraints that are close in rank effectively swap their relative position. As an effect, a stochastic OT system (SOT) can predict variation patterns and the learning algorithm for SOT puts the learner in a position to replicate the quantitative distribution pattern in the observed data. Asudeh (2001) shows that in combination with a harmonic

[33] A similar account was proposed by Anttila (1997).

alignment analysis following Aissen (1999), SOT can derive optionality patterns in Marathi (Indo-Aryan, India): for non-volitional transitive verbs like *saapaḍṇe* ('to find'), either of the two arguments can be realized as the subject (while the other is realized as the object).

In the standard OT learning algorithm, a learner reacts to an error that their preliminary constraint ranking makes (predicting a different candidate to be optimal than what the learner observes in an adult speaker) by reranking the constraints responsible for the incorrect prediction. In SOT, this is replaced by a gradual learning algorithm, which adjusts the numerical (mean) rank of the responsible constraints by a small increment. On this basis, exposure to a sufficiently large sample of adult speakers' data will lead to a replication of the stochastic distribution among the alternatives in the learner's grammar.

Bresnan et al. (2001) demonstrate with an analysis of corpus distribution of passives that the SOT approach, again combined with Aissen's (1999) harmonic alignment analysis, can explain strong quantitative effects in a language like English (which do not categorically enforce passive for certain person constellations among the arguments of transitive verbs) in parallel to strictly categorical patterns in Lummi (Straits Salish, British Columbia). In Lummi, the realization of a transitive verb with a third person agent and a first or second person patient in active voice is ungrammatical. The underlying input has to be realized in passive voice. Now, although in English a clause like *she invited me* is not ungrammatical, Bresnan et al.'s 2001 study showed that in corpora of spontaneous spoken English, there is a significant statistical effect of an elevated passive use in this constellation (i.e. *I was invited by her*) – a circumstance that is exactly what one would expect when the standard OT treatment of Lummi is extended to English in a stochastic OT framework.[34]

[34] In their criticism of Boersma & Hayes 2001, Keller & Asudeh (2002) argue that viewing a stochastic OT system not just as a model of variation/optionality, but of some notion of graded grammaticality tied to corpus frequencies, is conceptually problematic since it blurs the standard distinction between competence and performance. It should be noted however that certain systematic observations regarding quantitative distributions in corpus data seem to make it inevitable to revise some standard assumptions. Compare Bresnan's 2011 autobiographical notes: "Strikingly, the rare, marginal, and 'incorrect' construction types in large collections of English language usage parallel the rare grammatical phenomena that can be found across languages of the world. Moreover, judgments of ungrammaticality are often unstable and can be manipulated simply by raising or lowering the probability of the context. Most remarkably, language users have powerful predictive capacities, which can be measured using statistical models of spontaneous language use. From all these discoveries I have come to believe that our implicit knowledge of language has been vastly underestimated by theoretical linguistics of the kind I had practiced." (Bresnan 2011)

4.2 Functional motivation of the constraint set

In any linguistic account that makes use of abstract descriptive categories such as 'syllable', 'subject', 'passive', 'quantifier', etc., the symbolic expressions assumed to describe formal and content-related properties of linguistic utterances are theoretical constructs. They are not directly observable. However, most linguistic theories attempt to choose their central descriptive representations in a way that permits a mapping to empirically observable properties based on as few assumptions as possible: phoneme representations are chosen based on semantically distinguishable minimal pairs, for the definition of syntactic notions like subject, operationalized tests are advanced, etc. In the same vein, the candidate representations in most OT work (and definitely in OT-LFG) are chosen under an operationalized regime assuming that all candidate distinctions can be derived from the surface distribution and contextual clues reflecting semantic distinctions.

What about the choice of constraints used to drive the typological predictions? It is important to notice that even with perfectly uncontroversial, empirically grounded candidate representations, there are generally many extensionally equivalent choices of the constraint set for deriving the same distribution of optimal candidates (= the predicted language typology). For a given constraint set, an alternative set of *ad hoc* constraints with the same empirical prediction can be constructed. OT systems are empirically underdetermined in this respect (and this is no design fault; many formal systems employed as scientific models are underdetermined along certain dimensions). Therefore, to convince oneself that an OT analysis does indeed reflect a linguistically valid pattern, it is important to exclude that one or more of the constraints are ill-justified and merely play the role of getting the predictions right. Hence, starting out in phonological OT work, it has become good practice to provide a plausible motivation for each constraint that is independent from its role *within* the constraint set – a "functional" motivation. In phonology, many constraints can be given a motivation based on considerations of articulation, aerodynamics or perception, for instance "[…] a constraint against voiced obstruents, e.g., the Voiced Obstruent Prohibition (VOP, *[+voice], *Laryngeal) can be said to be functionally grounded, since vocal fold vibration is difficult to sustain if the outgoing airstream is blocked." (Krämer 2017: sec. 3.2.3). In syntax, the argumentation often needs to be more indirect (for instance by putting forward general considerations of economy to motivate a constraint against movement in a derivational framework, such as STAY in Grimshaw's 1997 and similar frameworks). But a wide-spread argumentation practice sees the need for independent justification of constraints beyond reaching a particular effect within a factorial typology. Many researchers welcomed that OT syntax triggered a confluence of the formalist and the functionalist perspective on language and grammars.[35] Still, Newmeyer (2002) presents a vigorous argument against the conception of "functionally-based optimality theory" (FOT) accounts, specifically targeting Aissen's (1999; 2003) harmonic alignment account. As the argumentation shows however (and as Bresnan & Aissen (2002) argue in detail in their rebuttal appearing in the same volume of *Natural Language & Linguistic Theory*), the concept of constraints that Newmeyer argues against when he criticizes that FOT "incorrectly locates the form-function interplay in the mental grammar itself, rather than seeing the response of form to function as emerging from language use and acquisition" (Newmeyer 2002: 43) is not the one inherent to the standard OT conception which completely shifts the definition of grammaticality away from a rule-based system to the interaction of violable constraints. Newmeyer draws into question whether "the claim that every constraint has a functional motivation" (Newmeyer 2002: 56) is empirically contentful for OT syntax. But if we take into account that constraints are abstract constructs which are not directly observable in any particular language, this formulation somehow reverses the need for justification that a theoretician should provide for their assumptions. Providing some plausible motivation beyond the technically desired effect responds to principles of scientific practice precluding arbitrariness in an underdetermined formal system – providing a functional motivation does not amount to an empirical *claim*, but it is part of the argumentation that the abstract choices made adhere to second-order principles of good scientific practice. It is only a full OT system that is empirically falsifiable; inadequacy may come to the surface when there is no way of extending a system which plausibly covers a core set of phenomena to clearly observable additional evidence.

4.3 Learnability in the context of the broader cognitive architecture

As has become clear from the discussions up to this point, a theorist's decision to move from a conventional generative grammar formalism to the competition-based formal model of grammaticality underlying OT does not merely mean that instead of working with hard constraints they now work with violable constraints. The status of familiar devices becomes fundamentally different with the different characterization of the set of well-formed analyses. Some of the debates that this development triggered have been already addressed in the previous sections – for instance the question how ineffability might be modeled and how op-

[35] Haspelmath (1999) argues for the additional need to take diachronic evolutionary processes in account.

tionality/free variation can be accounted for.[36] Not surprisingly, what is at the foreground of discussions in the linguistic literature are questions regarding the expressiveness of the formalism as a tool to describe and explain (constellations of) linguistic phenomena in particular languages. At the same time of course the typological predictions that go along with an OT system are appreciated.

But it is worthwhile considering the status of the components of an OT system as a model of the language faculty within our broader cognitive system. The competition-based frameworks of Harmonic Grammar (Legendre et al. 1990) and Optimality Theory (Prince & Smolensky 1993, 2004) were originally developed to reconcile (i) the potential of connectionist networks for learning complex input-output functions from exposure to data following up on work in the Parallel Distributed Processing framework (Rumelhart et al. 1986)[37] with (ii) the insights from linguistic theory in the generative tradition, which models systematic generalizations in a formal system. Optimality Theory is an attempt to gain ground towards resolving some of the most central challenges for the cognitive sciences: understanding how abstract systematic knowledge (which is best described using recursive symbolic systems, relying on logical inference) is implemented in connectionist architectures and how it blends with associative knowledge (which is best captured in subsymbolic terms). We know that the neurophysiological basis for all our cognitive systems, the human brain, is a large and complex connectionist network. For artificial connectionist networks, the potential to pick up complex patterns from empirical learning data has been convincingly demonstrated for many scenarios. However, the abstract symbolic concepts that are at the core of many linguistic accounts of grammar – allowing for a very compact characterization of very far-reaching generalizations – turn out to be hard learning targets for a bottom-up empirical learning procedure with the comparatively simple artificial networks available (Marcus 2001).[38]

Given the complexity of the brain and our very preliminary understanding of the interaction of cognitive subsystems involved, it is no surprise at all that there is still a gap between our current understanding of the level of abstract concepts capturing our systematic knowledge and what we know about the level of neurophysiological implementation. Yet, when looking at knowledge about language, there seems to be a certain degree of impatience in the research communities

[36] Wunderlich (2006), in his encyclopedia article on OT in morphology and syntax, provides a list of fundamental questions that arise when adopting an OT approach.

[37] The collection edited by Smolensky & Legendre (2006) is devoted to this perspective, but it does not seem to be very prominent in linguistic debates.

[38] Compare e.g. the systematicity debate started by Fodor & McLaughlin (1990) (Buckner & Garson 2019).

taking a conncectionist vs. an abstract symbolic approach – possibly because on both sides of the gap, mature theories and modeling frameworks have evolved substantially over the past decades, yet the key question of how the views go together remains rather open.

From the point of view of linguistics, it may seem that the concerns about the missing path across levels are just a problem for cognitive scientists who believe that a connectionist implementation of linguistic knowledge needs to be spelled out. Under the working assumptions of many linguistic frameworks, details of a technical implementation are of subordinate importance as long as one can convince oneself that a symbolic approach could be implemented in principle. This thinking ignores however that the established frameworks had (and have) a major weakness when it comes to explaining the learnability of a language in a somewhat realistic way (which is no embarassment *per se*, since it was never in the focus of research interests, but it is a fact). Formally, the Principles-and-Parameters Theory has an answer to the question of learnability: the framework assumes an articulate structural system to be innate, such that learning amounts to setting a small number of switches, the parameters. The Minimalist Program attempts to derive the complex structural displacement patterns (which are required to reduce the complex sound/meaning relationship available in natural languages to a uniform construction plan and which by assumption must be predetermined by universal grammar) as a consequence of a small set of assumptions that are justified on the grounds of simplicity of the theory. Through innate universal grammar, the learner of a language has access to the principles that govern the displacement patterns and thus the learner only needs to learn the language-particular choice of linearizing the displacement configurations (Chomsky et al. 2019). It is not clear however (and it is not part of the research question in such work) how this system would interact with other parts of the cognitive learning system, many of which are quite clearly responding to the statistical distribution of cues. But linguistic knowledge could not be acquired and put to use if it had no effective interfaces with the statistically sensitive parts of cognition. It is hard to imagine that a learner can find out about the space of possible constructions in a language when they cannot rely on expectations regarding *typical* (= highly frequent) ways of saying something; when the learner notices that certain expectations were incorrect, this will trigger a highly informative learning step regarding the relationship between the language system and contextual factors. The aspect of statistical learning is where connectionist approaches turn out to capture the empirical behavior quite realistically – but a theory that construes language acquisition as an entirely different process provides no grounds for capturing such triggers and any frequency-related patterns.

The classical constraint-based theories of grammar, such as LFG and Head-driven Phrase Structure Grammar (see Przepiórkowski forthcoming [this volume]), *do* emphasize the objective of finding psychologically realistic models of the cognitive processes of production and comprehension (i.e. parsing and generation algorithms that start from an information state that corresponds to what is available to a human listener or speaker). However, the goal of coming up with realistic algorithmic accounts capturing the information processing of real cognitive agents does not traditionally extend to the cognitive process of *learning* a language, i.e. acquiring the lexical and grammatical knowledge necessary to perform the tasks of production and comprehension. The grammar formalisms are not designed in such a way that there is a formal learning procedure that always starts from the same initial state, and then takes in observations from adult speakers of Bulgarian, English or Mandarin. But this would be required to have a falsifiable theory of the way grammatical knowledge is instantiated in cognitive agents. In the classical paradigm, the precise grammatical knowledge representation for a particular language is (still) specified by a scientific observer, the linguist, who is able to "look behind the scenes" and make far-reaching decisions about the use of certain descriptive means from a meta perspective – which is clearly not a realistic rendering of the information available to human learners (who however nevertheless reach the knowledge state robustly and fast). There are good reasons why the classical paradigm stops short of also trying to model the cognitive process of acquisition: The grammatical knowledge that is available to adult speakers of the languages of the world is complex and the intricate differences are far from being understood. So one might say that the community is still at the stage of clarifying what the exact targets for the acquisition process are, thereby avoiding a situation where it could not be truly judged whether a learning algorithm is on the right track from a theoretical point of view. An opponent could however argue that it is not clear whether the formalism that was designed so meta observers can specify a theory of adult linguistic knowledge provides the right concepts and interfaces to ever support a realistically learnable knowledge representation of grammar and the lexicon (for instance because it is unclear how associative knowledge merges in, as mentioned above). If there is an truth in this objection, the best strategy is probably to adopt a parallel strategy: advance systematic accounts of the adult linguistic knowledge and at the same time try to explore architectures that are better suited for modeling learning and for interfacing with knowledge that is more readily captured in a connectionist framework.

The design of Optimality Theory provides a link between the abstract symbolic level, tying in with established concepts from linguistic theory, and the level of connectionist implementation. From the point of view of LFG, the greatest conceptual gain from adopting an OT perspective might lie in the fact that this provides a fleshed-out learning algorithm (Tesar & Smolensky 1998), which is compatible with insights about the low-level generalizing behavior of connectionist approaches.

4.4 Extension II: Bidirectional optimization

The standard definition of grammaticality in an OT system is based on what is often called a speaker-oriented, or production-based competition. Here, the most harmonic candidate analysis realizing some underlying content representation is determined. The characterization of the competing candidates reflects the fundamental knowledge that a competent speaker of the language has: they know which is the grammatical way of expressing some thought in their language. When they learned their language, they had to exclude other ways that would be possible in principle. The candidate set and rerankable constraints are thus a straightforward rendering of the mathematical search space for learning an input-output function capturing the speaker's competence.

A competent language user however also has an additional ability: when other speakers make an utterance, the listener is (mostly) able to reconstruct what the speakers wanted to express in the given context. Since many expressions are ambiguous out of context, this is a non-trivial problem.

Formally, the disambiguation problem for a given surface form is construed as the mirror image of the task of realizing some underlying representation with the grammatical means of some particular language. Hence, it is tempting to explore to what extent the same competition-based architecture can be applied to model our ability to disambiguate. The representational setup of OT-LFG makes it particularly easy to implement the reverse competition: for a listener-oriented, or comprehension-based optimization, all that needs to be altered from the standard scenario is the basis for defining the candidate set. Here, all candidate analyses sharing the same surface string are compared.

A realistic model of our cognitive ability to make disambiguation decisions of course has to incorporate a lot more than just grammatical and lexical knowledge. On reading an ambiguous request/instruction like *Read in the library* (compare *Read in the comma-separated data file*), a reader may exploit frequency knowledge about the phrasal verb *read in* vs. the simple verb *read*, which one might argue is part of their extended lexical knowledge. But what will probably play a much more important role is the context of the request (Is there a physical library? Is the reader receiving programming hints?). And ultimately, reasoning about what the author meant will draw upon any available world knowledge.

So, taking all relevant knowledge sources into account, there is an asymmetry between production-based and comprehension-based optimization. But it is still a valid question what disambiguation decisions are made when there are no grammar-external clues. If we can isolate such cases, we might learn a lot about the way our grammatical knowledge is organized.[39]

A case in point are so-called word-order freezing phenomena, which have received considerable attention in OT frameworks. For languages with relatively free word order, the following type of effect is often reported: when there are no other disambiguating clues, listeners interpret examples as unambiguous that due to freedom of word order should actually have two possible interpretations. Bouma & Hendriks (2012), who provide a detailed discussion of OT treatments of word order freezing, give the Dutch example in (24):

(24) Dutch (Bouma & Hendriks 2012: (4))
Fitz zag Ella.
Fitz saw Ella
Only 'Fitz saw Ella' (SVO), although structurally compatible with 'Ella saw Fitz' (OVS)

The interpretive preference for cases like (24) is directly predicted if one assumes that not only the determination of grammatical (surface) outputs follows a (production-based) optimization, but also the distinction among potential underlying interpretations in comprehension – assuming the very same constraint set.[40] Accounts that make use of this idea are called bidirectional optimization accounts. Early discussions of such an account in an OT-LFG setting are found in Lee 2001 and in Kuhn 2000a, 2003.

Assuming that comprehension-based competition plays a role in a model of grammatical knowledge immediately raises questions about the relationship between the two directions: In a standard OT setting, one would not want to assume in general that listeners *only* apply constraint evaluation on the possible analysis candidates of a surface string to retrieve the semantic interpretation; they should also double check that the surface string is also optimal in the reverse

[39] Recall from Sections 1.2 and 2.6 that comprehension-based optimization is also predominantly used with the OT-style constraint ranking scheme implemented in the XLE system (Frank et al. 2001).

[40] Zeevat (2006) argues that bidirectional optimization is not an adequate account for word order freezing; but see Bouma & Hendriks 2012.

direction.[41] Demanding that a form-meaning pair has to be optimal both among all production-based and among all comprehension-based candidates is called strong bidirectional OT.

An additional variant of combining the outcome of the two optimizations was pioneered by Blutner (1998, 2000). With a so-called weak optimization, not only the optimal form-meaning pairs from a competition play a role in defining meaningful expressions. The "runner-up" form-meaning pair after removing the overall winner from the competition is defined to express a relationship too, and this can continue down a scale of expressions. The idea is best illustrated with the well-known example (25):

(25) a. John killed the sheriff.
 b. John caused the sheriff to die.

(25a) is the unmarked way of expressing that John killed the sheriff. Logically, (25b) is equivalent; however, on hearing this sentence (and not (25a)), most listeners will take the speaker to implicate that the sheriff died in an indirect way from John's actions, assuming Gricean maxims. Weak bidirectional optimality predicts this by assuming that the overall competition will make the form/meaning pair of sentence (25a) and a proposition with a plain instance of a killing event the most harmonic candidate in both directions. The more marked form (25b), paired up with a "specialization" of the meaning is weakly optimal.[42]

The fields of OT semantics and OT pragmatics which in essence build on top of the idea of bidirectional OT are among the most prolific areas in terms of publications (see for instance the *Stanford Encyclopedia of Philosophy* entry van Rooij & Franke 2020). It has to be noted that the role that the constraints play in OT systems modeling systematic patterns in pragmatics is slightly different from the grammaticality-defining role of the constraint set in OT phonology and OT syntax; nevertheless, there are many systematic connection points between OT syntax and OT semantics/pragmatics, discussed for instance by Beaver & Lee (2004).

[41]The simple comprehension-based optimization *does* play a role in work in OT phonology. The mechanism of lexicon optimization, assumed by Prince & Smolensky (1993), is such a competition. Also, Tesar & Smolensky (1998) propose a procedure of robust interpretive parsing during the learning process.

[42]The formal properties of bidirectional OT are discussed in Jäger 2002.

5 Developments after the mid-2000s

This section starts out with a discussion of the developments in the OT framework after the phase of the highest research activity in the late 1990s and early 2000s (Section 5.1) and furthermore asks how aspects of language learning are captured in OT vs. in mainstream work in Computational Linguistics and Natural Language Processing since the 2000s (Section 5.2). Section 5.3 addresses the relatively recent developments of neural modeling, i.e. the application of "deep" artificial neural networks.

5.1 Developments after the peak research activity

Müller (2012), in the conclusion to his survey paper on OT syntax, notes that after a phase of very high research activity, there have been comparatively few OT contributions in the area of syntax – contrary to the situation in morphology, semantics/pragmatics, and most notably phonology. To a certain degree, this development also holds for OT-LFG work on syntax. According to Müller's analysis of the situation, many ideas from derivationally based work in OT syntax live on in the research strand of the Minimalist Program, in which for instance the decision to apply one of a number of competing elementary operations like Agree, Merge, Move, Delete etc. is resolved in different ways across languages, which can be captured in a ranking account (Müller 2012: sec. 5).

In this section, we address the question: What was the further development in the part of the OT syntax community that assumes nonderivational candidate analyses – in particular OT-LFG? Section 4.3 above ended with the observation that the couching of LFG in an OT framework adds a plausible algorithmic account of learnability of grammars, which the original formalism was missing. Against this background, one might have expected an increase of activity rather than a reduction, for instance in the computationally oriented LFG subcommunity. One can only speculate about the exact reasons for trends in research communities, but there seem to have been framework-internal factors, which are discussed in this section, and external factors in the broader computational community, which will be discussed in Section 5.2.

Looking at the framework itself, the introduction of a competition-based definition of grammaticality in syntax research, or in a subcommunity of syntax research, did not go along with a fixed catalogue of new principles that readily leads to the formation of a homogeneous and focused research paradigm. Rather, the introduction of OT can be seen as the starting point for addressing a broad bundle of phenomena and conceptual aspects which could not be fully captured within the assumptions underlying a conventional formal grammar model – including quantitative aspects of linguistic knowledge that draw on a strictly binary notion of grammaticality and the established split of grammatical competence and performance into question.[43] One may see the fanning out of community activities into a whole range of development lines not as a failure of the new framework, but as a sign of the true complexity of empirical interrelations that the convential approach kept out of its scope by making strategic idealizing assumptions. Under such circumstances, the new framework, which is committed to lifting overly far-reaching idealizations, cannot provide firm ground in all respects. Regarding the quantitative aspects for instance, extensions like stochastic OT (Section 4.1) provide certain alternatives, but in particular work by Bresnan (2007); Bresnan & Ford (2010); Bresnan & Nikitina (2010) went on to explore a broader range of modeling options. For other aspects, bidirectional optimization and other extensions of the architecture promise the most explanatory account. In parallel to these architectural considerations, linguistic work continued to study empirically robust grammatical patterns for which symbolic abstractions have been established – exploring a phenomenon in new languages and/or interactions among phenomena. It is hence a quite natural that depending on the focus of particular studies, authors would choose to use an OT-LFG framework or a classical LFG setting.

5.2 The shift of computational syntax research in the 2000s

Besides the fact that the OT approach unlocked many new conceptual and empirical questions, there were presumably framework-external factors that precluded a streamlining of activities into a single coherent research paradigm. One of them could be that the linguistically motivated exploration of the formal OT architecture was not accompanied by major synchronized efforts in a computationally oriented community[44] – contrary to the situation in the 1980s and early 1990s where formal and computational results for linguistic grammar formalisms received significant attention at the computational linguistics conferences. At first glance, this is surprising. OT as a move to a competition-based framework in theoretical work in linguistics *did* lead to an architecture that is structurally very

[43]In her acceptance speech of the Lifetime Achievement Award by the Association for Computational Linguistics, Joan Bresnan shares her view of the development of the various descriptive frameworks that points in this direction (Bresnan 2016).

[44]The use of the idea of violable constraint for postfiltering the readings predicted by a classical grammar discussed in Sections 1.2 and 2.6 is so different conceptually from OT as a grammaticality-defining device that many of the implications of the latter do not arise for the former.

similar to the architectures underlying the dominant data-driven approaches in computational linguistics and natural language processing (NLP) (Eisner 2000: 287): A high-level objective in both contexts lies in developing a model framework capable of using empirical ("training") data from a natural language (in principle an arbitrary one) to induce a language-particular instantiation of the framework that replicates the competence/language behavior of a speaker of that language. This is reflected by the fact that the same families of learning algorithms could be employed (which led to interesting discussions between theoretical and computational linguistics, see Keller & Asudeh 2002; Goldwater & Johnson 2003; Jäger 2007). What differs across the research contexts despite the great similarities are three interleaved points:

First, linguistic OT work by design uses a strict constraint ranking to limit the degrees of freedom in theory development and to have abstract symbolic connecting points to work on linguistic generalizations.[45] In practically motivated machine learning approaches in NLP however, there is no reason for positing such restrictions on the learnable input-output functions. Hence, the actual learning processes for some given language input are quite different in the two research areas. Key insights from decades of research on the theory of grammar are nevertheless incorporated in most NLP modeling efforts of the time, since supervised learning on annotated corpora ("treebanks") is the most effective approach, and the annotation categories reflect many of the important distinctions. Again however, the generation of new insights in theoretical vs. computational work is somewhat disconnected: improvements in NLP are based on enrichments of the competition and evaluation models – which is excluded by assumption in linguistic OT work.

Second, while both in linguistics and in NLP, competitions in both processing directions play a role, it is the production-based view that plays the defining role (for grammaticality) in linguistics, whereas it is the comprehension-based view (=disambiguation in parsing) that is the most fundamental in NLP.[46]

Finally, third – and related to the first and second point – a linguistic optimization approach controls for context factors which would discriminate among candidates but are considered extralinguistic (or orthogonal to the phenomenon under consideration). This is important to be able to isolate the effect of the abstract grammatical categories responsible for systematic generalizations.[47] NLP work however pursues the objective of maximizing the predicted system scores

[45] Stochastic OT assumes oscillating rank values for a constraint, but each evalution is still based on the strict ranking from standard OT.

[46] Compare Sections 1.2 and 2.6 and the discussions in Frank et al. 2001, Kuhn 2001, 2003.

[47] The considerations from Section 5.1 regarding the rethinking of long-standing conceptual as-

for the target discrimination task (such as replicating the structure assignment decision human annotators made on real-life corpus sentences, making use of their linguistic knowledge as much as exploiting any explicit or implicit contextual clue). It is therefore not only legitimate, but good practice to try to exploit correlations of the actual target task with any other trends reflected in the empirical training data. For instance, a syntactic parser, say, of English, trained with NLP techniques on a treebank from a certain domain will presumably make its decisions to a large part because it has picked up grammatical knowledge regarding the positioning of grammatical subjects and objects from the treebank. To a certain degree, it will however also exploit non-linguistic domain knowledge reflected in the distribution of factual statements (e.g. a certain brand name being mentioned a lot more than others in a context like *buy a … watch*). The different objectives imply very different foci in the modeling work.

The three differences in the configuration details and application of competition-based learning architectures lead to distinct key research challenges in theoretical vs. computational work and explain how separate agendas have evolved. Nevertheless, there are numerous connecting points that become particularly relevant when simplifying assumptions underlying the respective standard approaches are relaxed. For example, the idea of bidirectional optimization from theoretical work can be translated straightforwardly to machine learning models applied to disambiguation in parsing and choice in generation (Cahill & Riester 2009; Zarrieß et al. 2011; Yu et al. 2019). An example of developments in which linguistic insights and considerations re-gained attention in the past years is the Universal Dependencies project (Nivre et al. 2016), which assembles treebanks for a growing number of languages in a cross-linguistically uniform dependency format (compare Haug forthcoming [this volume]).

5.3 OT and recent neural models in natural language processing

Section 4.3 addressed the connectionist motivation behind the original proposal of OT. Interestingly, in the 2010s successes in machine learning with artificial neural networks (Henderson 2020) and the broad availability of high-capacity computing resources brought about a major shift in NLP research (as well as most other areas of applied machine learning), often associated with the buzz word of "Deep Learning": neural network models replaced the conventional machine learning (ML) architectures discussed here in Section 5, which required the

sumptions showed there is a certain danger of circularity in a theoretically motivated exclusion of factors as extralinguistic or irrelevant for a given phenomenon.

design and optimization of the set of ML features to reach the best generalizations from training data. (The machine learning features are the equivalent to the rankable constraint in OT – the only difference is that they are used in more general mathematical functions than in the strict relative ranking of OT.)

The training of "deep" neural networks does not rely on pre-designed ML features. Instead, it employs hidden layers of neurons that are densely connected with a neuron-based input or output representation and with other hidden layers. Weight parameters on all the connections are iteratively adjusted by supervised training based on input/output data, where weight adjustments for links involving hidden layers are percolated from the ends using backpropagation (which can be thought of as spreading out the activation of some neuron to all connected neurons in the neighboring layers). The effect is that during empirical learning, the hidden layers receive the role of couching emerging internal representations of systematic patterns observed in the training data. For instance, when certain input neurons tend to be activated when a specific output occurs, some neuron in the next hidden layer can take over the role of an internal "feature" recording the constellation. With appropriately chosen network architectures that capture task-specific properties of the input and output representation (e.g. cross-talk among neighboring elements in a sequential input such as character or word token sequences), the neural approach led to substantial improvements over conventional ML approaches for a great variety of learning tasks.

The conceptual relationship between recent computational neural networks and the OT approach as popularized in the 1990s has not been discussed prominently in the research literature, but it is worthwhile comparing the major components in these architectures – in particular since the inherent black-box characteristics of neural models have prompted wide-spread efforts into making the emerging model representations scientifically interpretable (see e.g. Belinkov & Glass 2019). At first glance, the move away from human-designed ML feature sets, which parallel OT constraint sets, seems to have widened the gap between NLP work and a linguistic notion of OT competition. Moreover, the effective ability of neural models to induce task-relevant generalizing representations, which bridge between some input and some output representation, has made it possible to train neural architectures for complex input/output mapping tasks for which a conventional approach would have crucially involved a linguistically informed intermediate representation. An illustrative example is machine translation. Conventionally, it was unquestionable that the best possible machine translation approach would for instance exploit a parser on the source side and a generator on the target side, which take advantage of all accumulated insights regarding the grammatical systems in these languages (potentially mediated through ML-based parsers/generators trained on treebanks that capture the linguistic knowledge). Now, neural models for translation can be trained on very large numbers of pairs of input/output sentences without providing any abstract characterization of linguistic properties or composing a pipeline of substeps such as parsing, transfer and generation (for an overview, see Zhang & Zong 2015). The model parametrization resulting for the end-to-end translation process captures many of the relevant grammatical regularities in the hidden layers. The great advantage of the free induction of internal representations as systematic patterns occur in the observed input/output relation is that the model will capture not only regularities along the major dimensions underlying established modular descriptions of knowledge of language, but any other trends that are manifest in the data. And since all aspects are incorporated in a single, densely connected model architecture, cross-talk among the various knowledge sources and contextual clues is captured very effectively.[48]

But how could then a comparison with the classical OT architecture be of any use or lead to new insights? OT was designed to have an abstract symbolic level for an underlying connectionist architecture that implements an input/output function with supervised training. But with the assumed fixed set of grammatically relevant constraints, one of the key strengths of current connectionist models – the free induction of internal representations at hidden levels, without preconceived features or violable constraints – is excluded by design. Still, the tension between OT's fixed set of symbolic constraints and the current neural model's completely unrestricted representational space for capturing cross-talk among any factors influencing the observed input/output behavior could be the key to advances on both the linguistic and the computational side. The unrestricted neural models lack scientific interpretability, while the OT approach enforces the exclusion of non-linguistic factors as an influence for calculating the competitions. It is conceivable to reconcile both issues by building up symbolically informed diagnostic tools operating on trained neural models: such tools would in a first step allow researchers to test to what extent and with what representational means a model trained on corpus data captures known symbolic generalizations. In a second step, the ways in which this linguistically well-understood generalization interacts with other knowledge sources reflected in

[48]Recent advances in the neural modeling of text, in particular transformer-based contextual word embeddings like BERT (Devlin et al. 2019) have led to an often surprising level of emerging generalizations. Even before the advent of this generation of models, Loula et al. (2018) reported "impressive generalization capabilities" of neural sequence-to-sequence models in phrasal composition underlying the meaning of commands like *turn left twice*.

the training data could be explored. Recent work in NLP has started looking at various techniques (probing tasks, causal intervention) that can serve as such diagnostic tools (Ettinger et al. 2016; Gulordava et al. 2018). To give an idea of the potential that lies in symbolically guided analysis of the performance of trained neural models, Figure 7 shows an analysis of the behavior of a neural syntactic parser from Falenska & Kuhn 2019. The parser, a neural graph-based dependency parser, is trained in a supervised way to predict dependency arcs (essentially representing the grammatical relation between the words in a sentence). It can exploit a rich internal contextual representation to record emerging systematic dependencies, for example to capture the influence of intervening words (*some prefer water* vs. *some cold water*: the most likely grammatical relation between *some* and *water* is different).

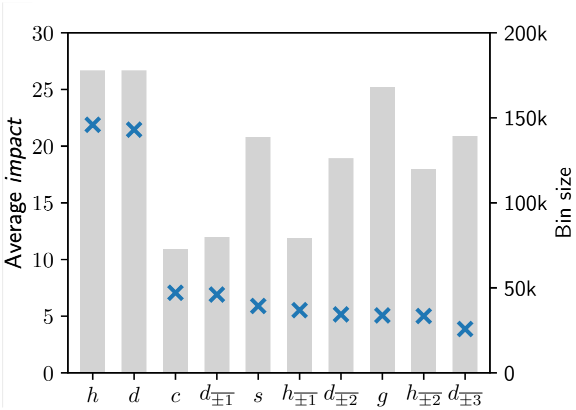


Figure 7: Analysis of trained neural dependency parsers from Falenska & Kuhn 2019, aggregating results from nine treebank training experiments (Ancient Greek, Arabic, Chinese, English, Finnish, Hebrew, Korean, Russian and Swedish): impact (blue crosses) of the information from words in various functional positions (occurring with a frequency plotted by the gray bars) on the prediction score of a multi-layer perceptron for a dependency arc in a graph-based dependency parser. The position types compared are the following: heads (h), dependents (d), children of d (c), siblings (s), grandparents (g), h,d$_{\mp i}$ tokens at distance $\pm i$ from h or d which are none of h, d, c, s, or g.

In the experiments depicted in Figure 7, gold-standard treebank information from a development dataset is used for analytical purposes: for a neural model trained without pre-defined ML features, the effective impact[49] of various different functional positions on the parser's predictive behavior is determined. The

[49]The paper introduces a special normalized measure for the impact of specific word representations on the prediction.

impact is calculated on the basis of the model's sensitivity to diagnostic manipulations of the input: information about the token located in the position in question is provided in one application experiment and left out in another. The diagram reveals that the parser has learned from exposure to data to be substantially more sensitive to the children of dependents (c) than to siblings (s), although the latter occur much more frequently in the training data. (It should be noted that there is no delimitable locus in the model's parameter space where this information is represented – but based on the theoretically informed diagnostic tests, we can observe that whatever the model has learned correlates with the distinction.)

Diagnostic tools of this kind provide a view on the ability of a connectionist model to pick up patterns for which a theory of syntax has posited symbolic abstractions – capturing their systematic significance. Since direct information about the representational configuration is not in the input representation that the parser receives in training or application, we have an indication that the emerging internal, connectionist representation does encode implicit knowledge about important aspects of functional structure.

Powerful neural modeling frameworks are hence on the way of providing a solution to certain aspects that motivated the use of symbolic formalisms with soft constraints. A central challenge for theoretical interpretability of the models' predictive capacity is to further develop the diagnostic machinery for disentangling the overlaid effects of very different knowledge sources – including various types of "non-linguistic" knowledge that are responsible for patterns in the composition of texts and communicative exchanges.[50]

[50] To make progress in this process of disentangling, it is necessary to cross long-established disciplinary boundaries that provided clear-cut subspaces in the study of language, text and corpora of discourses from specific contexts: theory of grammar with its subdisciplines, psycholinguistics, sociolinguistics, but also literary studies, media studies, etc.; many relevant patterns fall into the realm of scholarly disciplines that do not study language and text *per se*, such as cultural studies, history and social science. The delimited subspaces have so far justified convenient idealizations in the working assumptions – in the theory of grammar for example the idealizing assumption of a shared body of grammatical and lexical knowledge that makes up the linguistic competence of all native speakers of language X. An extra challenge comes from the divergent methodologies that have developed in the subfields as a response to the very distinct idealizations; this becomes clear in work in subareas of digital humanities and computational social science which has recently explored corpus-based modeling techniques for addressing research questions from literary studies (Kuhn 2019) or political science (Padó et al. 2019). Considerable effort is needed to appropriately incorporate findings from computational models in the respective question contexts and theoretical frameworks (Reiter et al. 2020).

6 Conclusion: epistemological considerations

Optimality Theory was introduced to the study of language in the 1990s as a symbolic description framework for linguistic representations which can make stronger empirically testable predictions about the human language faculty than a classical formal grammar. It can be seen as an attempt to overcome an epistemological limitation of the established formalisms underlying most work in linguistics.

With a classical framework (plain rewrite systems from the Chomsky hierarchy as much as extensions such as for instance transformational systems, tree-adjoining grammars and unification-based grammars), the predictions that can be tested "directly"[51] against empirical observations are always tied to the formal grammar for a particular language, i.e. a specific instantiation of the class of formal systems, for example, a specific context-free grammar $G_{137} = \langle N_{137}, \Sigma_{137}, P_{137}, S_{137} \rangle$ with a concrete set of non-terminal symbols, terminal symbols, rewrite rule productions and start symbol (and similarly for more expressive grammar formalisms).

One such instantiation is typically viewed as a scientific model for (aspects of) the grammatical competence of speakers of, say, Japanese, Swahili or English. The formal system predicts a set of terminal strings, which can be experimentally compared against the linguistic behavior of competent speakers. A grammatical theory about a range of phenomena, e.g. in Japanese, is then falsifiable in the sense of Popper 1959[52] because it is conceivable that relevant types of terminal strings predicted to be excluded from the formal language according to the theory *do* in fact occur in an experiment (potentially using corpus studies that compare against very similar string types predicted to be included).[53] The theoretical scope is however rather limited: If G_{137} fails to predict an empirically observable opposition of acceptable vs. unacceptable data in Japanese, all that

[51] "Directly" is set in scare quotes, since what goes through as direct empirical evidence in linguistic work always depends on methodological preassumptions that a community agrees on. The nature of language data in communication is such that very few event types is truly observable in a direct way. However, appeal to certain theoretical constructs and certain contextually triggered inferences is typically considered uncontroversial since they are orthogonal to research questions under debate.

[52] Falsifiability is a prerequisite for a theory with any predictive power.

[53] Most theories of grammar work with the stronger assumption that (some part of) the internal symbolic structures used by the formalism also represent relevant aspects of meaning, i.e. they can be viewed as logical forms. This provides an additional basis for testable predictions; the experiment then has to access speakers' (and listeners'/readers') interpretation of given strings. What is directly testable are however still the system's predictions for one fully parametrized (language-particular) instance of the grammar formalsm.

the theorist can conclude is that some aspect in the specification of this formal grammar has been inappropriate. When there are multiple ways of fixing the problem (e.g. by different strategies of introducing additional non-terminal symbols in a variant N'_{137} and by modifying some existing productions and adding some new ones in P'_{137}), there is no theoretically forceful way of distinguishing between them. Assume for the sake of the argument that intuitively only one of the modified grammars makes similarities in Korean, which is maybe modeled in some grammar G_{214}, explicit. As long as all potential modifications predict the same formal language, a predictive theoretical statement differentiating between the options cannot be made (for lack of falsifiability).

With an OT system, testable predictions are made for quite a different type of experiment than what we just saw for classical formal grammars: the theory implemented in a particular OT system (including a spelled-out candidate generation function *Gen* and a set of rankable constraints *Con*) does not predict a single formal language, but – via the factorial typology implied by all possible rankings of *Con* – a whole class of formal language approximations of natural languages. Fully formalized OT systems come with a spelled-out empirical learning algorithm (Tesar & Smolensky 1998 and subsequent work in the community), which provides the basis for falsifiability of a theory about the human language faculty as such (of course, focusing on some selection of linguistic phenomona): a formal OT system predicts how a learner of any specific language in their learning behavior responds to exposure to language behavior by adult speakers of the language in question.[54] A concrete OT system can thus be falsified by evidence from *any* natural language: a particular observed adult language behavior could in principle trigger a sequence of constraint rerankings that make it impossible for the learner to converge on the constraint ranking needed for this language. If this is the case for some conjectured theory (with a constraint set etc.) and observed language data, the theory of some part of the human languaage faculty counts as falsified.

Of course, most research communities employing classical formalisms, including the LFG community, have established agreed-upon meta principles and methodological research practices which effectively ensure that empirical evidence from a particular natural language is accepted by the community as evidence affecting *all* formal systems adhering to the shared conventions (to continue with our context-free grammar illustration, the theory could be characterized as the set of formal grammars $\{G_i \mid$ the components of G_i satisfy all meta

[54] Adult behavior is assumed to be observed in contexts that provide enough extra-linguistic clues to support inferences about the intended meaning where there is ambiguity.

principles }). There are countless examples of such meta principles: X-bar theory, extended projections, the concept of lexical redundancy rules, constraints on legitimate transformations in a derivational framework, a universal layering of functional projections, etc. With such principles it becomes possible that the grammatical theory constituted by the meta principles is falsified by language-particular evidence. Even the language learning/acquisition problem has been formulated on the basis of such framework, presumably most prominently in the Principles-and-Parameters theory (Chomsky 1981): the learner (guided by some language acquisition device that is assumed to be part of the cognitive equipment) has the task of adjusting a number of free parameters in an otherwise highly constrained innate grammatical system – in response to observed linguistic behavior by adult speakers. (A rigorous formalization as for OT learning algorithms is typically not provided.)

However, although sophisticated research practices have been established that ensure a far-reaching consensus about plausible meta principles assumed in a community, the epistemological relationship between a particular instantiation of the meta framework (say, formal grammar G_{137} for Japanese) and the framework itself, with its meta principles, remains contestable. Typically, the representational constructs developed to capture generalizations across natural languages have been established in a long process of cautious plausible reasoning – yet there are almost always alternative, empirically indistinguishable ways of predicting surface-level divergences across languages. To explain unexpected patterns in some language X, a modification in the formulation of one or another meta principle could be made, or idiosyncratic lexical knowledge could be posited. The falsifiability issue of general theoretical statements is not completely resolved by the assumption of meta principles, which explains in part why there have been many controversial debates about the representational locus for capturing cross-linguistic variation in a phenomenon – take for instance Binding Principles, which one might construe along a configurational tree structure or along a functional hierarchy (Asudeh & Dalrymple 2006). In the same vein, a reoccurring type of argument against specific linguistic accounts is the accusation for overly strong theory-internal assumptions. In other words, it happens quite frequently that members of a research community develop reservations with respect to the falsifiability of parts of the established consensus framework.[55]

[55]Given the underdetermination of theories of grammar by direct empirical evidence, aesthetic arguments regarding the simplicity of a theory are often advanced, most notably in Minimalism (Chomsky 1995). But even this strategy cannot escape controversies, since there are different possible starting points for seeding a theoretical accounts in fundamental propositions.

As just noted, OT can be technically seen as the move towards an approach that meets higher standards of falsifiability (when fully formalized). One may ask oneself then why it has not replaced classical formalisms in mainstream linguistic research? Section 5.1 discussed this question under the perspective of the concrete course of research activities in the 1990s and 2000s. But there are also relevant meta-theoretical considerations: since most aspects of representational choice in linguistic modeling are not directly observable (only the linear sequence of surface units of expression and semantic entailments of the content of utterances are directly observable), the space of possible OT theories remains vastly underdetermined. This means that (unless a community decides to change their research paradigm completely), plausible argumentation for abstract intermediate representations remains an important part of linguistic theorizing. And since substantial groundwork in linguistic research has always lied in the systematic capturing of regularities in variation patterns for particular languages, the justification for the use of classical formalisms has not disappeared. By choosing a strict ranking approach over constraints captured in terms of established symbolic representations, the OT endeavor was from the outset designed to stay connected with work using the classical formalisms; the effect of relevant constraints on a phenomenon under consideration can be calculated in manually constructed tableaux.[56] Insights from some OT account may thus feed back into the more general debate of what are appropriate theoretical constructs for systematically capturing a particular aspect of linguistic knowledge.

Against this background, the most important contribution of OT to generative linguistics might have been to increase the awareness in (part of) the community that a comprehensive, falsifiable account of the human language faculty has to include a formalized account of learnability of language from exposure to data.

References

Aissen, Judith L. 1999. Markedness and subject choice in optimality theory. *Natural Language & Linguistic Theory* 17(4). 673–711.

Aissen, Judith L. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21. 435–483.

[56] As was discussed in Section 5, recent machine learning approaches open up alternative avenues for modeling human language behavior. Computational models incorporating a very large parameter space can be trained on large corpora to achieve higher prediction accuracy on most tasks; however, the emerging representations (in the hidden layers of "deep" neural network models) provide no direct basis for a falsifiable theory of aspects of linguistic competence. Some connection to hypotheses that can be expressed symbolically seems to be indispensible.

Anderson, Stephen R. 1969. *West Scandinavian vowel systems and the ordering of phonological rules*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).

Andrews, Avery D. 1982. The representation of case in modern Icelandic. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 427–503. Cambridge, MA: The MIT Press.

Andrews, Avery D. 1990. Unification and morphological blocking. *Natural Language & Linguistic Theory* 8(4). 507–557. DOI: 10.1007/bf00133692.

Anttila, Arto. 1997. *Variation in Finnish phonology and morphology*. Stanford: Stanford University. (Doctoral dissertation).

Asudeh, Ash. 2001. Linking, optionality, and ambiguity in Marathi. In Peter Sells (ed.), *Formal and empirical issues in Optimality Theoretic Syntax*, 257–312. Stanford: CSLI Publications.

Asudeh, Ash & Mary Dalrymple. 2006. Binding theory. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn., 23–31. Amsterdam: Elsevier. DOI: 10.1016/b0-08-044854-2/01955-6.

Beaver, David I. & Hanjung Lee. 2004. Input-output mismatches in OT. In Reinhard Blutner & Henk Zeevat (eds.), *Optimality Theory and pragmatics*, 112–153. London: Palgrave Macmillan.

Belinkov, Yonatan & James Glass. 2019. Analysis methods in neural language processing: A survey. *Transactions of the Association for Computational Linguistics* 7. 49–72. DOI: 10.1162/tacl_a_00254.

Blutner, Reinhard. 1998. Lexical pragmatics. *Journal of Semantics* 15(2). 115–162. DOI: 10.1093/jos/15.2.115.

Blutner, Reinhard. 2000. Some aspects of optimality in natural language interpretation. *Journal of Semantics* 17(3). 189–216. DOI: 10.1093/jos/17.3.189.

Boersma, Paul. 1998. *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. Amsterdam: University of Amsterdam. (Doctoral dissertation).

Boersma, Paul, Joost Dekkers & Jeroen van de Weijer. 2000. Introduction — Optimality Theory: Phonology, syntax, and acquisition. In Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer (eds.), *Optimality Theory: Phonology, syntax, and acquisition*. Oxford: Clarendon Press.

Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32(1). 45–86. DOI: 10.1162/002438901554586.

Bögel, Tina, Miriam Butt, Ronald M. Kaplan, Tracy Holloway King & John T. Maxwell III. 2009. Prosodic phonology in LFG: A new proposal. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '09 conference*, 146–166. Stanford: CSLI Publications.

Bouma, Gerlof & Petra Hendriks. 2012. Partial word order freezing in Dutch. *Journal of Logic, Language, and Information* 21(1). 53–73.

Bresnan, Joan. 1996. LFG in an OT setting: Modelling competition and economy. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '96 conference*. Stanford: CSLI Publications. Handout from keynote address.

Bresnan, Joan. 2000. Optimal syntax. In Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer (eds.), *Optimality Theory: Phonology, syntax, and acquisition*, 334–385. Oxford: Oxford University Press.

Bresnan, Joan. 2001a. Explaining morphosyntactic competition. In Mark Baltin & Chris Collins (eds.), *Handbook of contemporary syntactic theory*, 11–44. Oxford: Blackwell.

Bresnan, Joan. 2001b. *Lexical-Functional Syntax*. Oxford: Blackwell.

Bresnan, Joan. 2002. The lexicon in Optimality Theory. In Paola Merlo & Suzanne Stevenson (eds.), *The lexical basis of sentence processing: Formal, computational and experimental issues*, 39–58. Amsterdam: John Benjamins. DOI: 10.1075/nlp.4.03bre.

Bresnan, Joan. 2007. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In Sam Featherston & Wolfgang Sternefeld (eds.), *Roots: Linguistics in search of its evidential base* (Studies in Generative Grammar), 77–96. Berlin: Mouton de Gruyter.

Bresnan, Joan. 2011. Linguistic uncertainty and the knowledge of knowledge. In Roger Porter & Robert Reynolds (eds.), *Thinking Reed: Centennial essays by graduates of Reed College*, 69–75. Portland, OR: Reed College.

Bresnan, Joan. 2016. Linguistics: The garden and the bush. *Computational Linguistics* 42(4). 599–617. DOI: 10.1162/COLI_a_00260.

Bresnan, Joan & Judith L. Aissen. 2002. Optimality and functionality: Objections and refutations. *Natural Language & Linguistic Theory* 20(1). 81–95.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-Functional Syntax*. 2nd edn. (Blackwell Textbooks in Linguistics 16). Malden, MA: Wiley-Blackwell.

Bresnan, Joan, Shipra Dingare & Christopher Manning. 2001. Soft constraints mirror hard constraints: Voice and person in English and Lummi. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '01 conference*, 13–32. Stanford: CSLI Publications.

Bresnan, Joan & Marilyn Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86. 168–213. DOI: 10.1353/lan.0.0189.

Bresnan, Joan & Tatiana Nikitina. 2010. The gradience of the dative alternation. In Linda Ann Uyechi & Lian-Hee Wee (eds.), *Reality exploration and discovery: Pattern interaction in language and life* (CSLI Lecture Notes 197), 161–184. Stanford: CSLI Publications.

Buckner, Cameron & James Garson. 2019. Connectionism. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosophy*, Autumn 2019. Stanford: Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2019/entries/connectionism/.

Cahill, Aoife, Tracy Holloway King & John T. Maxwell III. 2007. Pruning the search space of a hand-crafted parsing system with a probabilistic parser. In *ACL 2007 workshop on deep linguistic processing*, 65–72. Prague: Association for Computational Linguistics. DOI: 10.3115/1608912.1608924.

Cahill, Aoife & Arndt Riester. 2009. Incorporating information status into generation ranking. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, 817–825. Suntec, Singapore: Association for Computational Linguistics. DOI: 10.3115/1690219.1690261.

Cahill, Aoife & Andy Way. Forthcoming. Treebank-driven Parsing, Translation and Grammar Induction using LFG. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris Publications. DOI: 10.1515/9783110884166.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/9780262527347.001.0001.

Chomsky, Noam, Ángel J. Gallego & Dennis Ott. 2019. Generative grammar and the faculty of language: Insights, questions, and challenges. *Catalan Journal of Linguistics Special Issue: Generative Syntax: Questions, Crossroads, and Challenges*. 229–261. DOI: 10.5565/rev/catjl.288.

Collins, Michael. 1997. Three generative, lexicalised models for statistical parsing. In *Procedings of the 35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European chapter of the Association for Computational Linguistics*, 16–23. Madrid: Association for Computational Linguistics. DOI: 10.3115/976909.979620.

Crouch, Richard, Mary Dalrymple, Ronald M. Kaplan, Tracy Holloway King, John T. Maxwell III & Paula Newman. 2008. *XLE Documentation*. Xerox Palo Alto Research Center. Palo Alto, CA. https://ling.sprachwiss.uni-konstanz.de/pages/xle/doc/xle_toc.html.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III & Annie Zaenen (eds.). 1995. *Formal issues in Lexical-Functional Grammar*. Stanford: CSLI Publications.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee & Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 17th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (naacl-hlt 2019)*, 4171–4186. Association for Computational Linguistics. DOI: 10.18653/v1/N19-1423.

Dione, Cheikh M. Bamba. 2014. LFG parse disambiguation for Wolof. *Journal of Language Modelling* 2(1). 105–165. DOI: 10.15398/jlm.v2i1.81.

Eisner, Jason. 2000. Book reviews: Optimality Theory. *Computational Linguistics* 26(2). https://aclanthology.org/J00-2014.

Engels, Eva & Sten Vikner. 2014. *Scandinavian object shift and Optimality Theory*. Basingstoke: Palgrave Macmillan. DOI: 10.1057/9781137431646.

Ettinger, Allyson, Ahmed Elgohary & Philip Resnik. 2016. Probing for semantic evidence of composition by means of simple classification tasks. In *Proceedings of the 1st workshop on Evaluating Vector-Space Representations for NLP*, 134–139. Berlin: Association for Computational Linguistics. DOI: 10.18653/v1/W16-2524.

Falenska, Agnieszka & Jonas Kuhn. 2019. The (non-)utility of structural features in BiLSTM-based dependency parsers. In *Proceedings of the 57th annual meeting of the Association for Computational Linguistics*, 117–128. Florence: Association for Computational Linguistics. DOI: 10.18653/v1/P19-1012.

Fanselow, Gisbert & Caroline Féry. 2019. Ineffability in grammar. In *Resolving conflicts in grammars: Optimality theory in syntax, morphology and phonology. Linguistische Berichte / Sonderhefte 11*. Helmut Buske Verlag.

Fodor, Jerry A. & Brian P. McLaughlin. 1990. Connectionism and the problem of systematicity: Why Smolensky's solution doesn't work. *Cognition* 35. 183–204.

Forst, Martin. 2007. Filling statistics with linguistics: Property design for the disambiguation of German LFG parses. In *ACL 2007 workshop on deep linguistic processing*, 17–24. Prague: Association for Computational Linguistics. DOI: 10.3115/1608912.1608916.

Forst, Martin & Tracy Holloway King. Forthcoming. Computational implementations and applications. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Frank, Anette, Tracy Holloway King, Jonas Kuhn & John T. Maxwell III. 2001. Optimality Theory style constraint ranking in large-scale LFG grammars. In Peter Sells (ed.), *Formal and empirical issues in Optimality Theoretic syntax*, 367–397. Stanford: CSLI Publications.

Frank, Robert & Giorgio Satta. 1998. Optimality theory and the generative complexity of constraint violation. *Computational Linguistics* 24(2). 307–316. https://aclanthology.org/J98-2006.

Goldwater, Sharon & Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. In Jennifer Spenader, Anders Eriksson & Östen Dahl (eds.), *Proceedings of the Stockholm workshop on variation within Optimality Theory, Stockholm University*, 111–120.

Grimshaw, Jane. 1997. Projection, heads, and optimality. *Linguistic Inquiry* 28. 373–422.

Gulordava, Kristina, Piotr Bojanowski, Edouard Grave, Tal Linzen & Marco Baroni. 2018. Colorless green recurrent networks dream hierarchically. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, volume 1 (long papers)*, 1195–1205. New Orleans: Association for Computational Linguistics. DOI: 10.18653/v1/N18-1108.

Haspelmath, Martin. 1999. Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18(2). 180–205. DOI: 10.1515/zfsw.1999.18.2.180.

Haug, Dag. Forthcoming. LFG and Dependency Grammar. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Heck, Fabian & Gereon Müller. 2000. Successive cyclicity, long distance superiority, and local optimization. In Roger Billerey & Brook D. Lillehaugen (eds.), *Proceedings of the 19th West Coast Conference on Formal Linguistics*, 218–231. Somerville, MA: Cascadilla Press.

Henderson, James. 2020. The unstoppable rise of computational linguistics in deep learning. In *Proceedings of the 58th annual meeting of the Association for Computational Linguistics*, 6294–6306. Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.561.

Huang, C.-T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).

Jäger, Gerhard. 2002. Some notes on the formal properties of bidirectional Optimality Theory. *Journal of Logic, Language and Information* 11(4). 427–451.

Jäger, Gerhard. 2007. Maximum entropy models and stochastic Optimality Theory. In Annie Zaenen, Jane Simpson, Tracy Holloway King, Jane Grimshaw, Joan Maling & Chris Manning (eds.), *Architectures, rules, and preferences: Variations on themes by Joan W. Bresnan*, 467–479. Stanford: CSLI Publications.

Johnson, Mark. 1998. *Optimality-theoretic Lexical Functional Grammar*. Unpublished manuscript, Brown University; Paper presented at the 11th Annual CUNY Conference on Human Sentence Processing, Rutgers University.

Kager, René. 1999. *Optimality Theory*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/cbo9780511812408.

Kaplan, Ronald M. & Jürgen Wedekind. 2000. LFG generation produces context-free languages. In *COLING 2000 volume 1: The 18th International Conference on Computational Linguistics*, 425–431. DOI: 10.3115/990820.990882.

Karlsson, Fred. 1990. Constraint grammar as a framework for parsing running text. In *COLING '90 volume 3: Proceedings of the 13th International Conference on Computational Linguistics*, 168–173. DOI: 10.3115/991146.991176.

Karttunen, Lauri. 1998. The proper treatment of optimality in computational phonology. In *Finite state methods in natural language processing*, 1–12. https://aclanthology.org/W98-1301.

Keller, Frank & Ash Asudeh. 2002. Probabilistic learning algorithms and Optimality Theory. *Linguistic Inquiry* 33(2). 225–244. DOI: 10.1162/002438902317406704.

King, Tracy Holloway, Stefanie Dipper, Anette Frank, Jonas Kuhn & John T. Maxwell III. 2000. Ambiguity management in grammar writing. In *Proceedings of the Linguistic Theory and Grammar Implementation workshop at European Summer School in Logic, Language, and Information (ESSLLI-2000)*.

King, Tracy Holloway, Stefanie Dipper, Anette Frank, Jonas Kuhn & John T. Maxwell III. 2004. Ambiguity management in grammar writing. *Research on Language and Computation* 2. 259–280. DOI: 10.1023/b:rolc.0000016784.26446.98.

Krämer, Martin. 2017. Current issues and directions in optimality theory. In *The routledge handbook of phonological theory*. Routledge. DOI: 10.4324/9781315675428-3.

Kuhn, Jonas. 2000a. Faithfulness violations and bidirectional optimization. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '00 conference*, 161–181. Stanford: CSLI Publications.

Kuhn, Jonas. 2000b. Processing optimality-theoretic syntax by interleaved chart parsing and generation. In *Proceedings of the 38th annual meeting of the association for computational linguistics*, 368–375. Hong Kong: Association for Computational Linguistics. DOI: 10.3115/1075218.1075264.

Kuhn, Jonas. 2001. Generation and parsing in Optimality Theoretic syntax – Issues in the formalization of OT-LFG. In Peter Sells (ed.), *Formal and empirical issues in Optimality Theoretic Syntax*, 313–366. Stanford: CSLI Publications.

Kuhn, Jonas. 2002. OT syntax: Decidability of generation-based optimization. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, 48–55. Philadelphia: Association for Computational Linguistics. DOI: 10.3115/1073083.1073094.

Kuhn, Jonas. 2003. *Optimality-Theoretic Syntax – A declarative approach*. Stanford: CSLI Publications.

Kuhn, Jonas. 2019. Computational text analysis within the Humanities: How to combine working practices from the contributing fields? *Language Resources and Evaluation* 53(4). 565–602. DOI: 10.1007/s10579-019-09459-3.

Kuhn, Jonas. Forthcoming. LFG, Optimality Theory and learnability of languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Lee, Hanjung. 2001. Markedness and word order freezing. In Peter Sells (ed.), *Formal and empirical issues in Optimality Theoretic Syntax*, 63–128. Stanford: CSLI Publications.

Legendre, Géraldine, Jane Grimshaw & Sten Vikner (eds.). 2000. *Optimality-Theoretic Syntax*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/5161.001.0001.

Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky. 1990. *Can connectionism contribute to syntax? Harmonic grammar, with an application*. Tech. rep. 467. Institute of Cognitive Science, University of Colorado at Boulder. CU-CS-485-9, Computer Science Technical Reports.

Legendre, Géraldine, Paul Smolensky & Colin Wilson. 1998. When is less more? Faithfulness and minimal links in wh-chains. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky (eds.), *Is the best good enough? Optimality and competition in syntax*, 249–289. Cambridge, MA: The MIT Press & MITWPL.

Loula, João, Marco Baroni & Brenden M. Lake. 2018. Rearranging the familiar: Testing compositional generalization in recurrent networks. *CoRR* abs/1807.07545. http://arxiv.org/abs/1807.07545.

Marcus, Gary. 2001. *The algebraic mind: Integrating connectionism and cognitive science*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/1187.001.0001.

Maxwell, John T., III & Ronald M. Kaplan. 1989. An overview of disjunctive constraint satisfaction. In *Proceedings of the 4th International Workshop on Parsing Technologies (IWPT 1995)*, 18–27. Also published in Tomita (1991) as 'A Method for Disjunctive Constraint Satisfaction', and reprinted in Dalrymple, Kaplan, Maxwell & Zaenen (1995: 381–402).

Maxwell, John T., III & Ronald M. Kaplan. 1996. Unification-based parsers that automatically take advantage of context freeness. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '96 conference*, 1–31. Stanford: CSLI Publications.

McCarthy, John J. & Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill N. Beckmann, Laura Walsh Dickey & Suzanne Urbanczyk (eds.), *Papers in Optimality Theory*, 249–384. Amherst, MA: GLSA.

Müller, Gereon. 2003. Local vs. global optimization in syntax: A case study. In Jennifer Spenader, Anders Eriksson & Östen Dahl (eds.), *Proceedings of the Stockholm workshop on variation within Optimality Theory, Stockholm University*, 82–91.

Müller, Gereon. 2012. Optimality-theoretic syntax. Unpublished manuscript, Universität Leipzig.

Müller, Gereon. 2014. Optionality in optimality-theoretic syntax. In *The second GLOT international state-of-the-article book*, 289–322. Berlin: De Gruyter Mouton.

Müller, Gereon. 2020. *Inflectional morphology in harmonic serialism*. Sheffield, UK: Equinox Publishing.

Newmeyer, Frederick J. 2002. Optimality and functionality: A critique of functionally-based optimality-theoretic syntax. *Natural Language & Linguistic Theory* 20(1). 43–80.

Nivre, Joakim, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajič, Christopher D. Manning, Ryan McDonald, Slav Petrov, Sampo Pyysalo, Natalia Silveira, Reut Tsarfaty & Daniel Zeman. 2016. Universal Dependencies v1: A multilingual treebank collection. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC'16)*, 1659–1666. Portorož: European Language Resources Association (ELRA). https://aclanthology.org/L16-1262.

Padó, Sebastian, Andre Blessing, Nico Blokker, Erenay Dayanik, Sebastian Haunss & Jonas Kuhn. 2019. Who sides with whom? Towards computational construction of discourse networks for political debates. In *Proceedings of the 57th annual meeting of the Association for Computational Linguistics*, 2841–2847. Florence: Association for Computational Linguistics. DOI: 10.18653/v1/P19-1273.

Pesetsky, David. 1998. Some optimality principles of sentence pronunciation. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky (eds.), *Is the best good enough? Optimality and competition in syntax*, 337–383. Cambridge, MA: The MIT Press & MITWPL.

Popper, Karl. 1959. *The logic of scientific discovery*. London: Hutchinson. Translation of *Logik der Forschung* (1935), by the author.

Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. Tech. rep. 2. New Brunswick, NJ: Rutgers University Center for Cognitive Science. Published as Prince & Smolensky (2004).

Prince, Alan & Paul Smolensky. 2004. *Optimality theory: Constraint interaction in generative grammar*. Oxford/Malden: Blackwell.

Przepiórkowski, Adam. Forthcoming. LFG and HPSG. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/3225.001.0001.

Reiter, Nils, Axel Pichler & Jonas Kuhn (eds.). 2020. *Reflektierte Algorithmische Textanalyse. Interdisziplinäre(s) Arbeiten in der CRETA-Werkstatt [Reflected Algorithmic Text Analytics: Interdisciplinary Works from the CRETA Workshop]*. Berlin: De Gruyter.

Riezler, Stefan, Tracy Holloway King, Ronald M. Kaplan, Richard Crouch, John T. Maxwell III & Mark Johnson. 2002. Parsing the Wall Street Journal using a Lexical-Functional Grammar and discriminative estimation techniques. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*. Philadelphia: Association for Computational Linguistics.

Riezler, Stefan, Detlef Prescher, Jonas Kuhn & Mark Johnson. 2000. Lexicalized stochastic modeling of constraint-based grammars using log-linear measures and EM training. In *Proceedings of the 38th annual meeting of the Association for Computational Linguistics*, 480–487. Hong Kong: Association for Computational Linguistics. DOI: 10.3115/1075218.1075279.

Rumelhart, David E., James L. McClelland & the PDP Research Group (eds.). 1986. *Parallel distributed processing, volume 1: Explorations in the microstructure of cognition: Foundations*. Cambridge, MA: The MIT Press.

Samek-Lodovici, Vieri. 1996. *Constraints on subjects: An Optimality Theoretic analysis*. New Brunswick, NJ: Rutgers University. (Doctoral dissertation).

Sannasgala, P. B. 1976. *Sinhalese vocables of Dutch origin*. Colombo: Kularatne.

Sells, Peter. 1999. Constituent ordering as alignment. In Susumu Kuno et al. (eds.), *Harvard studies in Korean linguistics 8*. Harvard University.

Sells, Peter. 2001. *Structure, alignment and optimality in Swedish*. Stanford: CSLI Publications.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R. M. W. Dixon (ed.), *Grammatical categories in Australian languages*. Australian Institute of Aboriginal Studies.

Smolensky, Paul. 1995. On the internal structure of the constraint component Con of UG. Handout of talk, UCLA, 7 April, Rutgers Optimality Archive (http://roa.rutgers.edu), ROA 87.

Smolensky, Paul. 1996. *The initial state and 'richness of the base' in Optimality Theory*. Tech. rep. 4. Department of Cognitive Science, Johns Hopkins University.

Smolensky, Paul & Géraldine Legendre (eds.). 2006. *The harmonic mind: From neural computation to optimality-theoretic grammar*. Cambridge, MA: The MIT Press.

Stabler, Edward P. 2011. Computational perspectives on Minimalism. In Cedric Boeckx (ed.), *Oxford handbook of Linguistic Minimalism*, 617–642. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199549368.013.0027.

Stabler, Edward P. 2013. Two models of Minimalist, incremental syntactic analysis. *Topics in Cognitive Science* 5(3). 611–633. DOI: 10.1111/tops.12031.

Tesar, Bruce & Paul Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29(2). 229–268. DOI: 10.1162/002438998553734.

Tomita, Masaru (ed.). 1991. *Current issues in parsing technology*. Dordrecht: Kluwer Academic Publishers. DOI: 10.1007/978-1-4615-3986-5.

van der Beek, Leonoor & Gerlof Bouma. 2004. The role of the lexicon in Optimality Theoretic syntax. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '04 conference*, 60–80. Stanford: CSLI Publications.

van Rooij, Robert & Michael Franke. 2020. Optimality-theoretic and game-theoretic approaches to implicature. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosophy*, Summer 2020. Stanford: Center for the Study of Language & Information (CSLI). https://plato.stanford.edu/archives/sum2020/entries/implicature-optimality-games.

Vikner, Sten. 2001. The interpretation of object shift and Optimality Theory. In Gereon Müller & Wolfgang Sternefeld (eds.), *Competition in syntax*, 321–340. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110829068.

Wedekind, Jürgen. 1999. Semantic-driven generation with LFG- and PATR-style grammar. *Computational Linguistics* 25(2). 277–281. https://www.aclweb.org/anthology/J99-2006.

Wunderlich, Dieter. 2006. Syntax: Optimality Theory. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn., 408–418. Amsterdam: Elsevier. DOI: 10.1016/b0-08-044854-2/04213-9.

Yu, Xiang, Agnieszka Falenska, Ngoc Thang Vu & Jonas Kuhn. 2019. Head-first linearization with tree-structured representation. In *Proceedings of the 12th International Conference on Natural Language Generation*, 279–289. Tokyo: Association for Computational Linguistics. DOI: 10.18653/v1/W19-8636.

Zaenen, Annie & Richard Crouch. 2009. OBLs hobble computations. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '09 conference*, 644–654. Stanford: CSLI Publications.

Zarrieß, Sina, Aoife Cahill & Jonas Kuhn. 2011. Underspecifying and predicting voice for surface realisation ranking. In *Proceedings of the 49th annual meeting of the Association for Computational Linguistics: Human Language Technologies*, 1007–1017. Portland, OR: Association for Computational Linguistics. https://aclanthology.org/P11-1101.

Zeevat, Henk. 2006. Freezing and marking. *Linguistics* 44(5). 1095–1111. DOI: 10.1515/LING.2006.035.

Zeevat, Henk & Gerhard Jäger. 2002. A reinterpretation of syntactic alignment. In *Proceedings of the fourth international Tbilisi Symposium on Language, Logic and Computation*.

Zhang, Jiajun & Chengqing Zong. 2015. Deep neural networks in machine translation: An overview. *IEEE Intelligent Systems* 30(5). 16–25. DOI: 10.1109/MIS.2015.69.

Chapter 4

LFG and Continental West-Germanic languages

Gerlof Bouma
University of Gothenburg

1 Introduction

This chapter is concerned with the Lexical-Functional Grammar treatment of the present-day West-Germanic languages *sans* English or Scots. This group is sometimes referred to as Continental West Germanic (Zwart 2008), as it is mostly comprised of Germanic languages[1] spoken in countries on the European mainland: from Belgium and the Netherlands, through Luxembourg and Germany, to Switzerland, the very northern parts of Italy, and Austria; and in addition in smaller regions bordering these countries. In spite of the label 'continental', the group of Continental West Germanic languages (henceforth: CWG) also contains languages like Yiddish (outside of Europe at least spoken in Israel and North America), Afrikaans (South Africa and Namibia), Dutch in the Antilles and Suriname (in terms of language politics part of the same standardization body as Belgium and the Netherlands), and German heritage variants in the Americas (Putnam 2011) and Siberia (Andersen 2016), to name but a few group members outside of continental Europe.

[1]Unless it is relevant to the discussion, I will use the term *language* in a broad sense that ignores matters like the language/language variety/regiolect/dialect/etcetera's political status, whether a language label covers a homogeneous or heterogeneous group of subvarieties, or whether it is mutually intelligible with languages that do not fall under the same label.

Gerlof Bouma. Forthcoming. LFG and Continental West Germanic languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

In terms of L1 speakers, the largest of these languages are standard varieties of German and Dutch, with circa 75 and 25 million speakers, respectively.[2] Their status as standardized national languages in multiple countries also means they are supported by strong academic infrastructures. It is therefore not surprising that most of the work on CWG in LFG is done on these two languages. Standard German and Dutch[3] will figure prominently in this chapter. This is merely a reflection of their salience in the LFG literature, and should not be interpreted linguistically, for instance as a sign of them being more typical CWG languages than other members of the language group.

A comprehensive inventory of West-Germanic languages with demographic and linguistic information can be found in Ethnologue[4] (Eberhard et al. 2019). Bibliographic data on West-Germanic can be found in Glottolog[5] (Hammarström et al. 2019). Note that neither of these resources distinguishes a CWG branch in their taxonomies. For an overview of the syntactic traits of Continental West Germanic, I refer the reader to Zwart (2008). An accessible description of how German distinguishes itself syntactically from English and from the North-Germanic languages can be found in the introductory chapters of Haider (2010), with arguments that in many cases carry over to the other CWG languages.

1.1 A general picture of Continental West Germanic, with a focus on the clause

In this subsection we will go through some of the syntactic traits of CWG. Our focus will be on the clause/verbal domain, since this has been the main interest of the LFG literature on CWG. We will discuss the nominal domain more briefly. The purpose of this subsection is twofold. First, it gives a very general impression of CWG syntax and indicates how it distinguishes itself from its North- and West-Germanic neighbours. Secondly, it introduces some of the language-particular background needed to understand the individual LFG analyses discussed in the rest of the chapter.

1.1.1 Clause layout

A prominent syntactic feature of CWG languages is the combination of *asymmetric verb second* together with *verb final* (Zwart 2008; Haider 2010). The former

[2]Counts based on Eberhard et al. (2019).

[3]Unless the context requires otherwise, I will use *German* and *Dutch* without further modifiers to refer to the standardized, national language varieties of these two CWG languages.

[4]https://www.ethnologue.com/subgroups/west-0, consulted July 2022

[5]https://glottolog.org/resource/languoid/id/west2793, consulted July 2022

label characterizes a clause structure in which the finite verb in main clauses, but *not* in subordinate clauses, is preceded by exactly one constituent, which can have a wide range of grammatical functions. The latter label covers the generalization that any verbal material that is not in second position – finite verbs in subordinate clauses and non-finite verbs in general – clusters together towards the end of the clause, potentially following arguments and adjuncts. As such, CWG contrasts with Modern English, which lacks both pervasive verb second in main clauses and verb finality, and follows a more rigid subject-verb-complement order. CWG also differs from the present-day North-Germanic languages, which can be characterized as combining verb second with subject-verb-complement order.[6]

For the discussion of the layout of a CWG clause, it is helpful to make use of the so-called *topological field model* of the clause, which can be found in traditional descriptions of German and Dutch and in reference grammars like Zifonun et al. (1997) and Haeseryn (1997). In this model, the layout of a clause is described in terms of linearly ordered fields, and different word order variants associated with different clause types are obtained by assigning constituents to different fields. The field schema we use in this chapter is given in (1).

(1) lead ‖ Vorfeld | left bracket | Mittelfeld | right bracket | Nachfeld ‖ tail

In a main clause, the *left bracket* (lb) and *right bracket* (rb) are reserved for verbal material: a single, second-position finite verb is in the left bracket, and any other verbs are in the verb cluster in the right bracket. Between the brackets there is the *Mittelfeld* 'middle field' (Mf), which may contain any number of constituents. The *Vorfeld* 'prefield' (Vf) is the designated place for the single constituent preceding the finite verb, whereas the *Nachfeld* 'postfield' (Nf) may contain several items, and is typically reserved for heavier constituents like clausal arguments, extraposed relative clauses and adverbial prepositional phrases. The *lead* and *tail* fields[7] host material that is more loosely connected to the clause, such as

[6]Two remarks are in order with respect to this characterization of CWG clause layout. First, as it can be used to demarcate CWG from English as well as from North-Germanic, it gives some linguistic substance to the pooling of CWG languages into one group, as we do in this chapter. Secondly, and somewhat weakening the first point, once we start to look closer at individual CWG languages, we find deviations from the general pattern. Modern Yiddish in particular fits the description poorly, both in terms of the verb-second pattern and the verb-final pattern. It is beyond the scope of this chapter to go into all the exceptions, but some of them will be discussed towards the end of this section and in the context of LFG analyses of these exceptions.

[7]The terminology around these last two fields is not as established as for the fields that are part of the clause proper. The lead is for instance also known as *Vorvorfeld* 'pre-prefield' or

vocatives and hanging topics. Note that not every field needs to contain material. Examples of different declarative main clause types are given in (2).

(2) Dutch

a. Subject-initial declarative:

Vf———	lb—	Mf———	rb———
De draken	doen	Doris dadelijk	duizelen.
the dragons	make.PRS.PL	Doris immediately	feel.dizzy.INF

'The dragons immediately make Doris feel dizzy.'

b. Object-initial declarative:

Vf—	lb—	Mf———	rb———
Doris	doen	de draken dadelijk	duizelen.
Doris	make.PRS.PL	the dragons immediately	feel.dizzy.INF

c. Sentence adverb-initial declarative:

Vf—	lb—	Mf———	rb———
Dadelijk	doen	de draken Doris	duizelen.
immediately	make.PRS.PL	the dragons Doris	feel.dizzy.INF

The verb-second constraint is clear here: the finite verb is always in the left bracket and precedes its subject if a non-subject is in the Vorfeld (2b,c). The subject and object can appear in identical positions – contrast the OVS order in (2b) with SVO in (2a). Linear order is therefore not fully determined by grammatical function, or vice versa. The Vorfeld is also the target of long-distance dependencies, like fronting of wh-constituents out of embedded clauses (not shown here). The Mittelfeld may contain a collection of (nominal) arguments and (simple) adverbials, which are typically local to the clause. The extent to which the order of material within the Mittelfeld is fixed differs between languages (see Section 2.4.1). Grammatical-function assignment under word order variation, long-distance dependencies, and the order of elements in the Mittelfeld are all basic CWG phenomena that the LFG models discussed below must address.

Other clause types have empty Vorfeld regions, such as the polar interrogative (3a), which is a verb-first construction, and the subordinate clause (3b), in which the left bracket is filled by the complementizer.

(3) Dutch

linkes Außsenfeld 'left outfield', and the tail as *Nachnachfeld* 'post-postfield' or *rechtes Außenfeld* 'right outfield'.

a. Polar interrogative:

lb—	Mf—				rb—
Doen	de	draken	Doris	dadelijk	duizelen?
make.PRS.PL	the	dragons	Doris	immediately	feel.dizzy.INF

b. Subordinate:

lb–	Mf—				rb—	
…dat	de	draken	Doris	dadelijk	doen	duizelen.
COMP	the	dragons	Doris	immediately	make.PRS.PL	feel.dizzy.INF

Example (3b) has the finite verb in the right bracket, in the verb cluster. This shows the asymmetry of the verb-second phenomenon: unlike in a main clause, the finite verb in a subordinate clause can be preceded by any number of constituents in the Mittelfeld. The topological model accommodates the complementary distribution of the finite verb of a main clause and the complementizer of a subordinate clause by locating both in the left bracket.

The right bracket in (3b) contains two verbs: first the finite verb, then the non-finite verb. This is considered to be the default order in Standard Dutch, but there is considerable variation in this ordering, both between and within CWG languages. An extensive overview of ordering possibilities in CWG verb clusters is given in Wurmbrand (2004).

The topological schema based on the combination of verb second and verb final is widely applicable to the CWG languages, but, as with any generalization, there are cases where it does not apply. To start, we must keep separate the notion of main clause vs. subordinate clause *word order* from the notion of unembedded and embedded clause *uses*. This is because German, amongst others, allows embedded clauses to have verb second under bridge verbs in the absence of a complementizer (see Section 2.2.1); that is, it allows main clause word order for certain embedded clauses. Furthermore, the separation of non-verbal material in the Mittelfeld and verbal material in the right bracket is not always as clean as the topological model suggests, as languages may allow for material from the two fields to be mixed, blurring the border between them (see Section 2.1.2). Finally, Afrikaans and Yiddish have clause structures that deviate further. Spoken Afrikaans optionally allows the combination of a complementizer and verb second in subordinate clauses (Biberauer 2009). Modern Yiddish has verb second in main as well as in subordinate clauses, and in addition its status as a verb-final language is debated (Diesing 1997). Historical stages of Yiddish did however follow CWG's characteristic pattern more closely (Santorini 1992).

1.1.2 Clause union

The examples in (2) and (3) each contain two verbs. We have discussed the topological model as a schema of the clause, without questioning whether we are dealing with mono-clausal structures here. Since Bech (1955/7), it is common to distinguish between *coherent* and *incoherent* verb combinations. The former describes a combination of two verbs into a single clause, *clause union*, whereas the latter results in a biclausal structure. The contrast is illustrated below. In (4a), the coherently combining *durfde* 'dared' shares the verb cluster with its embedded verb *te kopen* 'to buy', and the embedded object sits in the Mittelfeld of the clause headed by *durfde*. Example (4b) contains the incoherently combining *beloofd* 'promised', and here we see that both the embedded verb and its object appear after the matrix verb, in the Nachfeld of the matrix clause. As shown in (4c), this word order is not available for the coherently combining *durfde* 'dared'.

(4) Dutch

a.	lb——	Mf————			rb——————	
…omdat	hij	geen	auto	durfde	te kopen.
because	he	no	car	dared	buy.TEINF[8]

'…because he didn't dare to buy a car.'

b.	lb——	Mf	rb———	Nf———————		
…omdat	hij	beloofde	geen	auto	te kopen.
because	he	promised	no	car	buy.TEINF

'…because he promised not to buy a car.'

c. * …omdat hij durfde geen auto te kopen.
because he dared no car buy

The second sign that we are dealing with one clause in (4a) and two in (4b) is the scope of the negation, as evident from the translations. In both examples, negation is marked on the embedded object through the negative determiner but it nevertheless scopes over the finite verb in (4a). The same negation marking in the biclausal (4b) yields a narrow scope negation.

A third phenomenon associated with clause union is the potential to trigger *infinitivus pro participio* (IPP; German: *Ersatzinfinitiv* 'replacement infinitive'). IPP refers to realization of a verb in the infinitive when a participle is expected on the basis of the selecting auxiliary. For this to occur, the clause itself must

[8]The abbreviations TEINF and ZUINF in the glosses are used for the verb forms in Dutch and German that combine the infinitive marker (*te* in Dutch, *zu* in German) with an infinitive. Unlike corresponding forms in for instance English, these combinations are generally not separable.

also contain a further, lower verb in the infinitive.[9] The occurrence of IPP is therefore evidence of the middle and lower verb combining coherently. Below, IPP is triggered in (5a), affecting the coherently combining *durven* 'dare', but not in (5b) for the incoherently combining *beloofd* 'promised'.

(5) Dutch[10]

a. Hij heeft geen auto {durven / *gedurfd} te kopen.
   he has no car dare.INF dare.PTCP buy.TEINF
   'He didn't dare to buy a car.'

b. Hij had {beloofd / *beloven} geen auto te kopen.
   he had promise.PTCP promise.INF no car buy.TEINF
   'He promised not to buy a car.'

Example (5a) additionally shows that a clause can contain more than two verbs. In principle, there is no limit to the number of verbs involved in clause union, since the same couple of coherently combining verbs can appear at multiple levels of embedding.[11]

A wide range of verbs allow for coherent combination. For instance, for Dutch, the reference grammar Haeseryn (1997) lists over 100 verbs that always combine coherently, and an additional 20 that do so optionally. In this list we find auxiliaries; evidential, modal and aspectual verbs; but also verbs with a clearer lexical contribution such as causal and perceptual verbs, and for instance verbs corresponding to *help*, *learn*, *try* or *forget*. In theoretical syntactic work, combining behaviour is commonly taken to be an underived, lexical property of the embedding verb, but see Cook (2001) for an explanation of coherence in German in terms of information structure.

1.1.3 Crossing dependencies

When we have coherently combining verbs that also introduce their own object, we can end up with a clause in which a sequence of objects in the Mittelfeld is

[9] Further conditions may apply, for instance on the order of the auxiliary and the middle verb.

[10] A note on the use of brackets and parentheses in examples in this chapter: I will use curly brackets to indicate choice. The choice is either between several forms in one position, such as in the current example (5), or between several positions for one form, such as in the example in (28). Square brackets delimit constituents when this is relevant, such as in (6). Parentheses indicate optionality as usual.

[11] In practice, it seems that three-verb combinations are common, but more complex clauses are rare. For instance, Coussé & Bouma (2022) report numbers for a mixed corpus of written and spoken Dutch: about 3% of coherence domains contain 3 verbs, but only 0.1% contain 4 verbs.

followed by a corresponding sequence of verbs in the cluster. In languages like German or West Frisian, the unmarked order of the objects is by increasing order of embedding $O_1O_2...O_n$, whereas the order of verbs is by decreasing order of embedding $V_n...V_2V_1$. This gives rise to a pattern of nested dependencies between the objects and their verbs (6).

(6) Standard German

		O_1		O_2		V_2	V_1
...dass	wir	[dem	Hans]	[das	Haus]	streichen	halfen
COMP	we	the.DAT	Hans	the.ACC	house	paint	helped

'...that we helped (V_1) Hans (O_1) paint (V_2) the house (O_2).'

In Dutch and Swiss German, however, objects and verbs can *both* be ordered by increasing level of embedding $O_1O_2...O_n$ and $V_1V_2...V_n$. This creates cross-serial dependencies between objects and verbs, as in (7).

(7) Swiss German (Shieber 1985: §2, example 1)

		O_1		O_2		V_1	V_2
...das	mer	[em	Hans]	[es	huus]	hälfed	aastriiche.
COMP	we	the.DAT	Hans	the.ACC	house	helped	paint

The phenomenon of cross-serial dependencies has received ample interest in the literature, because it requires more than context-free power to model (Bresnan et al. 1982; Pullum & Gazdar 1982; Shieber 1985).

1.1.4 In and around the nominal domain

We end the overview of CWG syntax by briefly discussing the main characteristics of the nominal domain and, even more briefly, adpositions. This is to give a general sense of what these domains look like in CWG languages. Most of what is discussed below resembles what we find in the North-Germanic languages and English.

The nominal domain in CWG generally follows a determiner–adjective–noun pattern, with further adnominal material (relative clauses, PPs, etc.) realized postnominally. This is exemplified in (8).

(8) Dutch

de	mooiste	plek	van	Europa
the	beautiful.SUPERLATIVE	place	of	Europe

'the most beautiful place in Europe'

Present-day CWG languages have at most four cases (NOM, GEN, DAT and ACC; for example Standard German), but many make fewer case distinctions (see Kasper 2014 for a compact description of the situation in German varieties and references), and several have only a subject-object form distinction remaining in the pronominal paradigm (for example, Afrikaans, Dutch, and West Frisian).[12] There are two numbers (SG, PL). Any of three genders (M, F and N; Alemannic[13], Low Saxon,[14], Standard German, West Flemish), two genders (COMMON gender and N; Dutch, West Frisian) or no distinction (Afrikaans) may occur. Gender agreement distinctions only show up in the singular. The different paradigm sizes with respect to gender are illustrated in (9). Note the form contrasts in the definite determiners.

(9) a. Alemannic

d	Frau	dr	Maa	s	Chind
the.F.NOM	woman	the.M.NOM	man	the.N.NOM	child

b. West Frisian

de	frou	de	man	it	bern
the.COMMON	woman	the.COMMON	man	the.N	child

c. Afrikaans

die	vrou	die	man	die	kind
the	woman	the	man	the	child

Adjectives can be associated with multiple inflectional paradigms – see Section 3.2.2 for a discussion of these *declension classes* in Standard German. Even in languages with more elaborate paradigms, there is typically a great level of syncretism between forms across inflectional dimensions for determiners, pronouns, adjectives and nouns. The consequences of syncretism for grammatical modelling are discussed in Section 3.2.4.

Adpositions are overwhelmingly prepositional (10a), but the sporadic postposition (10b) and circumposition (10c) occur as well.

(10) Gronings (Low Saxon)

[12] See taalportaal.org for linguistic descriptions of Afrikaans, Dutch, and West Frisian. Consulted September 2022.

[13] The term Alemannic (German) covers amongst others Alsace German, Swabian and Swiss German.

[14] Low Saxon is by some authors used interchangeably with Low German. Our use of the term here comprises regional languages of the North of Germany and the East/North-East of the Netherlands. Our choice for the term Low Saxon is partly driven by the fact that LFG work on this language uses this term: see Section 3.2.1.

a. **op** de grìns
on the border
'on the border'

b. t haile joar **deur**
the whole year through
'the whole year through'

c. **om** de provìnzie **tou**
around the province around
'around the province'

1.2 Overview of the rest of the chapter

Thus far, I have talked about the geographic distribution and the syntactic characteristics of the Contintental West-Germanic languages, to define the scope of the chapter, and to give a background for what is to come below.

The remainder of this chapter is devoted to LFG analyses of different aspects of CWG syntax. In Section 2, I discuss LFG accounts of the clause and the verbal domain, and in Section 3, I discuss LFG studies of the nominal domain. These sections are structured in a parallel fashion: they start with analyses of the overall structure of their respective domains, and then continue with a discussion of more specific LFG accounts organized by topic. In the LFG literature on CWG, the clausal and verbal domains have received by far the most attention, which means that the corresponding section dominates this chapter in terms of size.

The chapter ends with concluding remarks in Section 4, in which I briefly touch upon some LFG and LFG-related work that was not included in detail here, and give pointers for further reading.

2 LFG analyses in the clausal and verbal domains

This section deals with phenomena at the level of the clause. I will start in Section 2.1 with a discussion of the variety of ways in which the overall shape of the clause has been modelled, mostly in terms of c-structure. I then look at specific topics that have been prominent in the LFG literature on CWG languages. The topics are divided into thematic sections. Phenomena at the left and right periphery are discussed in Section 2.2 and Section 2.3, respectively. Studies dealing with the ordering of dependents are discussed in Section 2.4. Finally, mapping-based analyses of areas of CWG clause syntax are presented in Section 2.5.

2.1 The overall shape of the clause

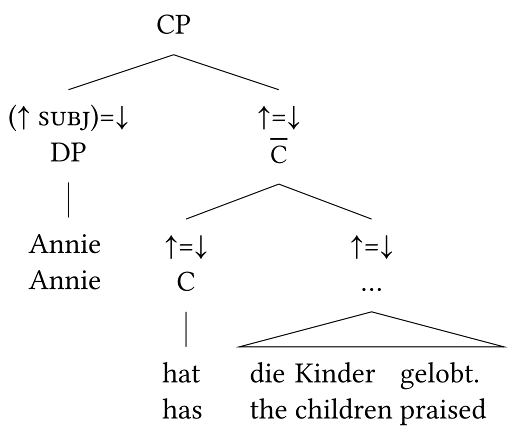
The discussion of the different LFG conceptions of the overall shape of the clause is organized according to the topological field model. I first consider the top level of the clause (directly containing the Vorfeld and left bracket) in Section 2.1.1, and then the lower level of the clause (the Mittelfeld and right bracket) in Section 2.1.2. The Nachfeld is discussed in Section 2.1.3.

2.1.1 Vorfeld and left bracket

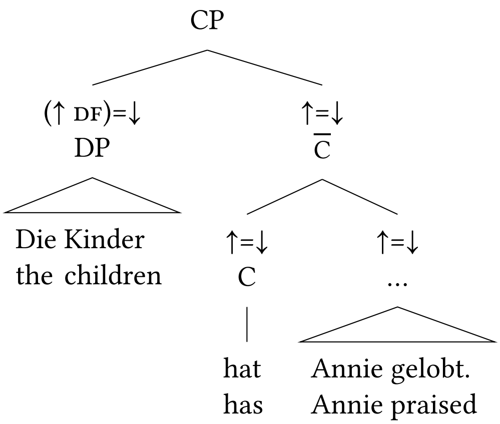
Berman & Frank (1996), Choi (1999), Berman (2003), and Frank (2006) model the German verb-second clause as a CP. The finite verb sits in C irrespective of whether the initial position is occupied by the subject of a declarative clause (11a) or by some other element, like the object in (11b). The complementary distribution between the finite verb and the complementizer in the left bracket follows as well: the complementizer can only appear in C, and when it is realized, the finite verb must occur in another, lower position (11c).

(11) German

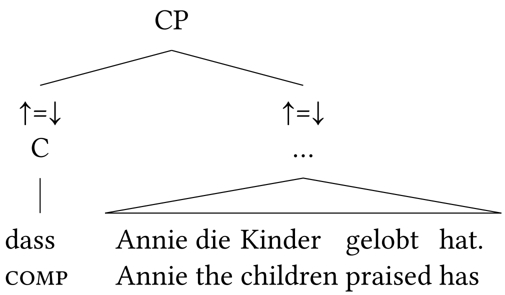
a.



b.



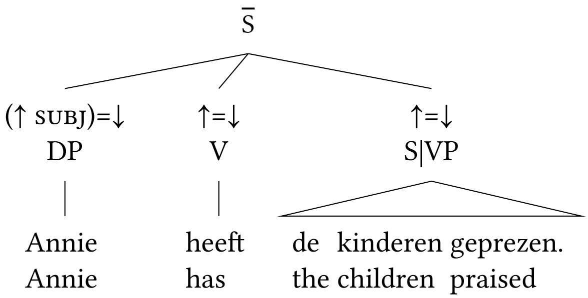
c.



The nature of Comp-CP, the node dominating the combined Mittelfeld and right bracket, differs between these authors, however, and will be discussed below. Van der Beek (2005) and Jones (2020), on Dutch, consider only main clauses, which they posit to be IPs.

Zaenen & Kaplan (1995, on Dutch; 2002 on German) prefer a slightly flatter structure, exemplified in (12). The label 'S|VP', a convention from the cited papers, is used to show that the authors do not wish to choose between these categories.

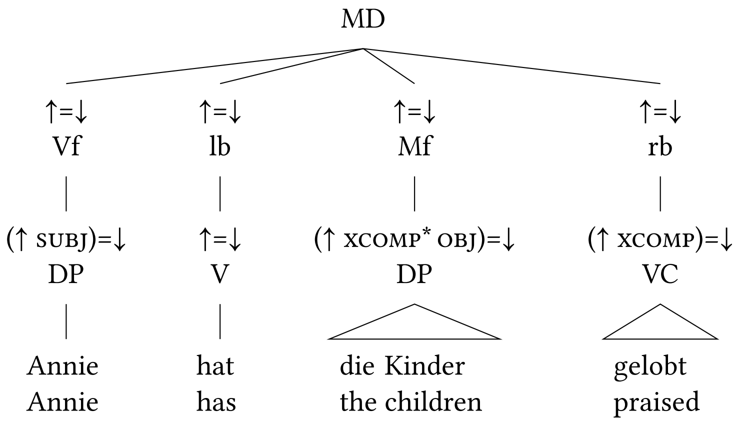
(12) Dutch



Zaenen & Kaplan's subordinate clauses are isomorphic to those in (11c), but are labelled $\overline{\text{S}}$ instead of CP.

An even flatter structure appears in the computational grammar fragment discussed in Clément et al. (2002), who model the topological field schema directly in LFG. All topological fields are c-structure nodes and direct descendants of the MD node ('main domain') that represents the whole sentence. Example (13) gives a somewhat simplified tree, using the abbreviations for topological fields I introduced in Section 1.1.1.

(13)



A very similar flat structure can be found in Rohrer (1996).[15]

2.1.2 Mittelfeld and right bracket

The Mittelfeld and right bracket form the lower c-structure level in the clause. This is the unlabelled Comp-CP in (11) and the S|VP node in (12). All authors agree that this part of the tree does *not* involve an IP.[16] This choice against an

[15] Rohrer (1996) however also writes "Diese flache Struktur läßt sich problemlos in eine binäre rechtsverzweigende Struktur umwandeln. [...] Wir behalten das flache Mittelfeld hier primär aus expositorischen Gründen bei" (p96, fn 3). [This flat structure can be converted to a binary right-branching structure without problems. We maintain the flat Mittelfeld here primarily for reasons of exposition.]

[16] In fact, in LFG, the assumption of an IP anywhere in CWG c-structure is rare. We mentioned Van der Beek (2005) and Jones (2020), on Dutch, who use it as the category at the top level, for the whole V2 declarative clause. The choice is not further motivated in these works, and moreover it is peripheral to the respective discussions. Bresnan et al. (2016) posit that Comp-CP contains an IP in one of the book's exercises on German. However, since this is a textbook, it is unclear whether the authors are theoretically committed to this choice, or whether it was made for other reasons, for instance pedagogical ones.

intermediate IP can also be found in analyses of German in the Chomskyan tradition, for instance in the line of work summarized in Haider (2010: see §2.2 therein for an overview of the arguments).

A salient question in the analysis of this part of the clause is the order of the verbs and the arguments, and the concomitant contrast between nested versus cross-serial dependencies. We will focus first on the two polar opposites: verbs ordered after increasing level of embedding (cross-serial dependencies) and verbs ordered after decreasing level of embedding (nested dependencies). The following pair, a variation on (6–7) above, illustrates the difference with three verbs in the verb cluster:

(14) a. Standard German
...dass wir [die Kinder] [dem Hans] [das Haus] streichen helfen lassen
comp we the.acc children the.dat Hans the.acc house paint help let
'...that we let the children help Hans paint the house.'

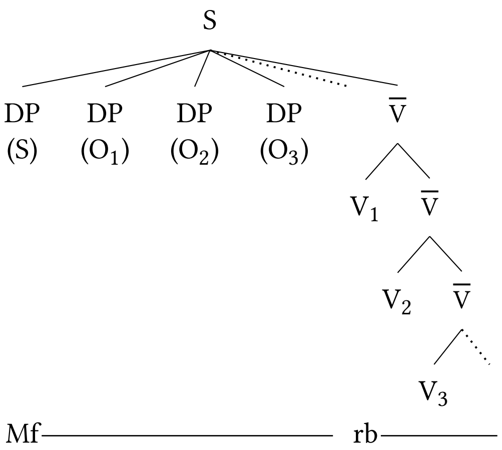
b. Swiss German (Shieber 1985: §2, example 5)
...das mer [d' Chind] [em Hans] [es huus] lönd hälfe aastriiche.
comp we the.acc children the.dat Hans the.acc house let help paint.

As mentioned in the introduction, there is considerable variation in the order of the verbal elements beyond these two opposites, and there is even variation in the extent to which the nominal material in the Mittelfeld and verbal material in the right bracket is kept separated, both between and within CWG languages. This variation will be briefly discussed at the end of this subsection.

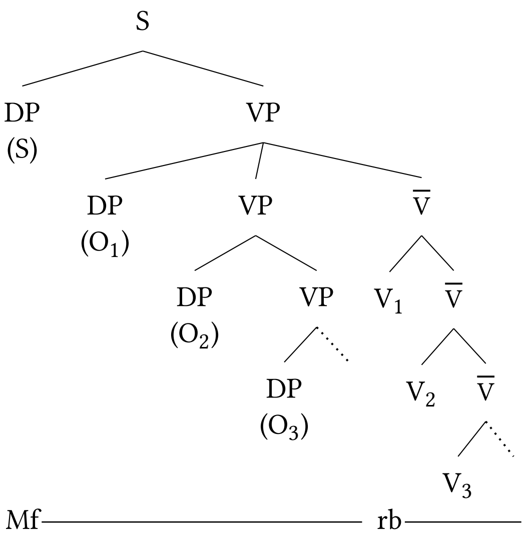
2.1.2.1 Cross-serial dependencies An early LFG analysis of Dutch cross-serial dependencies is found in Bresnan et al. (1982), which was a prominent demonstration of how LFG's formalism has the power needed for linguistically valid analyses of such dependencies.[17] Starting from a proposal by Evers (1975), schematically in (15a), with a flat Mittelfeld and a right-branching verb-cluster, Bresnan et al. argue that a structured Mittelfeld is to be preferred, as in (15b).

[17]The paper played a central role in the discussion of the context-freeness of natural language syntax. See e.g. Pullum & Gazdar (1982) and Shieber (1985) for more discussion of the issues involved and the kind of evidence considered.

(15) a.



b.



The tree in (15b) contains two parallel embedding structures: one for the objects in the Mittelfeld and one for the verbs in the right bracket. This is captured in the c-structure definitions in (16).

(16) a. $\mathrm{S} \longrightarrow \underset{(\uparrow \textsc{subj})=\downarrow}{\mathrm{DP}} \quad \underset{\uparrow=\downarrow}{\mathrm{VP}}$

b. $\mathrm{VP} \longrightarrow \left(\underset{(\uparrow \textsc{obj})=\downarrow}{\mathrm{DP}}\right) \left(\underset{(\uparrow \textsc{xcomp})=\downarrow}{\mathrm{VP}}\right) \left(\underset{\uparrow=\downarrow}{\overline{\mathrm{V}}}\right)$

c. $\overline{\mathrm{V}} \longrightarrow \underset{\uparrow=\downarrow}{\mathrm{V}} \quad \left(\underset{(\uparrow \textsc{xcomp})=\downarrow}{\overline{\mathrm{V}}}\right)$

For the objects, each level of VP embedding adds a level of XCOMP embedding at f-structure. For the verbs, each level of $\overline{\mathrm{V}}$ embedding does the same. A compatible stacking of XCOMPs is thus built up in both parts of the tree. The optionality of the object DP in (16b) allows for verbs that do not introduce their own object. It is essential that an XCOMP level is introduced for these in both parts of the tree, too, to maintain the parallel structure.

The accusative with infinitive verbs involved in the cross-serial construction are analyzed as raising-to-object verbs. The inflected verb form *zag* 'saw', for instance, receives a lexical entry along the lines of (17)

(17) *zag* V (↑ PRED)='see⟨SUBJ,XCOMP⟩OBJ'
(↑ XCOMP SUBJ)=(↑ OBJ)
(↑ XCOMP FORM)=INF
(↑ SUBJ NUM)=SG
(↑ TENSE)=PST

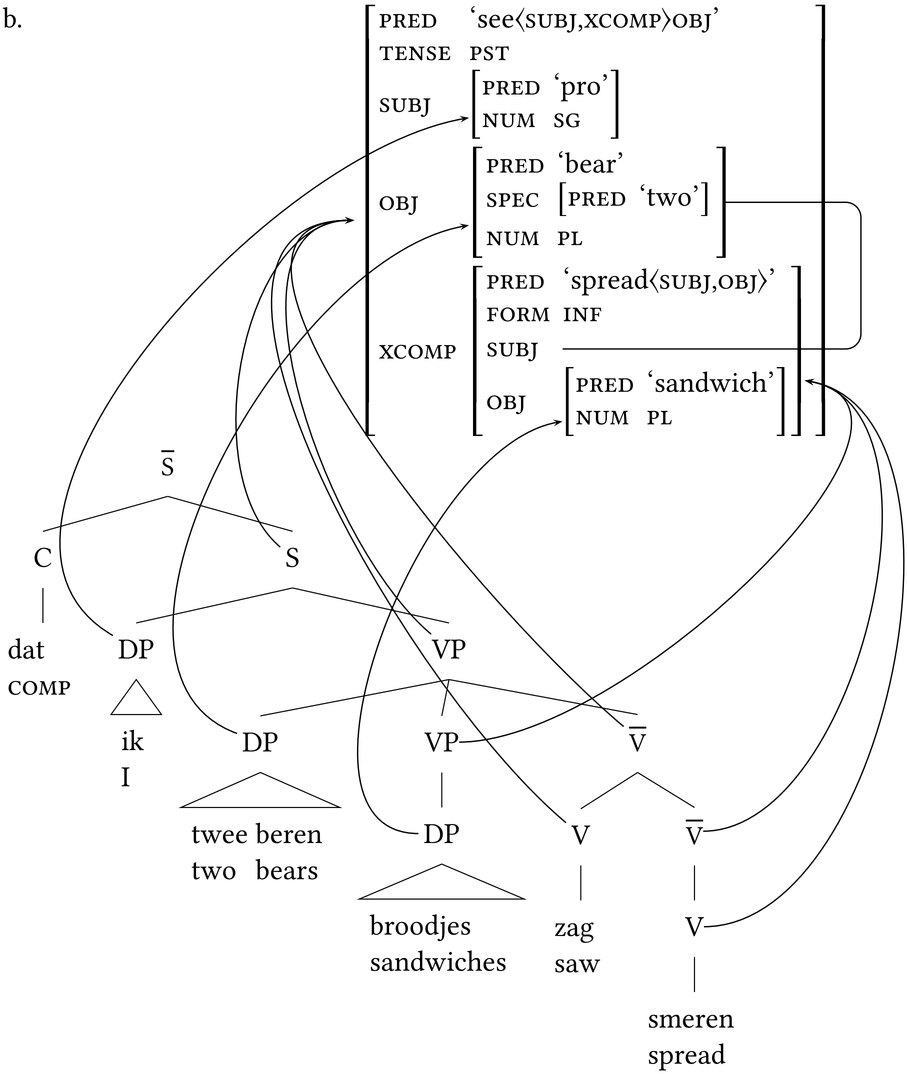
Complemented with rules for DPs and additional lexical entries, this grammar fragment gives us analyses like the one in (18).

(18) a. Dutch

	S	O_1		O_2	V_1	V_2
...dat	ik	[twee	beren]	broodjes	zag	smeren.
COMP	I	two	bears	sandwiches	saw	spread

'...that I saw (V_1) two bears (O_1) prepare (V_2) sandwiches (O_2).'

b.



However successful in capturing cross-serial dependencies, this analysis runs into descriptive problems if taken more generally as a model of the Dutch sentence. Zaenen & Kaplan (1995) give the example in (19), which involves the coordination of two V̄s that each require a different level of XCOMP embedding for the object supplied in the Mittelfeld.

(19) Dutch (Zaenen & Kaplan 1995: §2.3, example 9)
...dat Jan een liedje [[$_{\overline{\mathrm{V}}}$ schreef] en [$_{\overline{\mathrm{V}}}$ trachtte [$_{\overline{\mathrm{V}}}$ te verkopen]]].
COMP Jan a song wrote and tried sell.TEINF
'...that Jan wrote and tried to sell a song.'

Since different levels of XCOMP embedding of the object correspond to different c-structures in the model of Bresnan et al. (1982), example (19) cannot receive an analysis if we use the standard treatment of constituent coordination in LFG. It would require the shared material to receive two different c-structures at the same time. Zaenen & Kaplan's (1995) alternative relies on functional uncertainty to connect the objects to predicates at the required level of XCOMP embedding, and on functional precedence rules to make sure that the linear order of objects reflects their level of embedding. They replace the VP and $\overline{\mathrm{V}}$ rules of (16) with those in (20).

(20) a. $$\mathrm{VP} \longrightarrow \begin{array}{cc} \mathrm{DP}^* & \overline{\mathrm{V}} \\ (\uparrow \textsc{xcomp}^* \ \textsc{obj}){=}\downarrow & \uparrow{=}\downarrow \end{array}$$

b. $$\overline{\mathrm{V}} \longrightarrow \begin{array}{c} \mathrm{V} \\ \uparrow{=}\downarrow \end{array} \left(\begin{array}{c} \overline{\mathrm{V}} \\ (\uparrow \textsc{xcomp}){=}\downarrow \\ \neg((\uparrow \textsc{xcomp}^+ \ \textsc{obj}) <_f (\uparrow \textsc{obj})) \end{array} \right)$$

This analysis abandons the nested c-structure of the VP in favour of a flat one, which moves us back in the direction of (15a). The functional uncertainty equation on the object DP in (20a) allows connecting the object to a predicate at any depth of XCOMP embedding, and the general principles of f-structure coherence and completeness make sure each object is matched to exactly one predicate. The functional precedence constraint on the $\overline{\mathrm{V}}$ node in (20b) prevents more embedded objects from preceding less embedded ones. Together, these f-structure constraints force the same relation between Mittelfeld objects and right bracket verbs as the c-structures subtrees in Bresnan et al.'s analysis. Moreover, the interaction between functional uncertainty and the standard LFG approach to constituent coordination lets us handle sentences like (19) correctly.

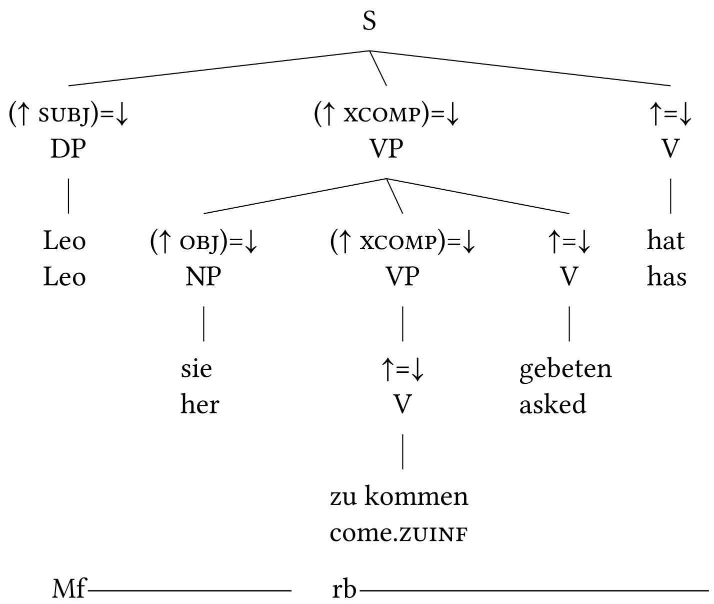
Zaenen & Kaplan (1995) also apply the combined use of functional uncertainty and functional precedence to Zürich German, where cross-serial dependencies are observed as well.

2.1.2.2 Nested dependencies The analyses of the structures that would give rise to consistently nested dependencies all come from LFG work on Standard German. However, explicit discussion of constructions with multiple objects at different levels of embedding is rare in this part of the literature – perhaps because the modelling of these dependencies is not seen as particularly problematic. We therefore do not always fully know how the relevant nested dependencies are to be derived in these LFG models.

Some authors assume nested VPs, which rather naturally correspond to nested dependencies between objects and verbs, even when this consequence is not a central concern. One example is the grammar fragment of Netter (1988), who gives annotated c-structures like the one in (21).

(21) a. German (Netter 1988: §1, example C4)
...dass Leo sie zu kommen gebeten hat.
COMP Leo her come.ZUINF asked has
'...that Leo asked her to come.'

b.



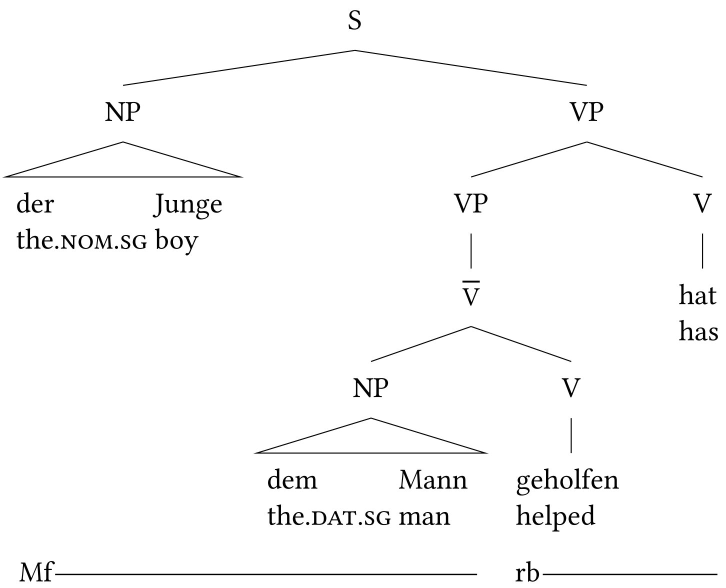
We also find nested VPs in Choi (1999), where the combined Mittelfeld and verb cluster of the subordinate clause in (22a) would get the structure given in (22b).[18]

(22) German

[18] Choi (1999) does not provide a tree for this exact sentence, but does show a more complex example with a comparable structure. In addition, Choi (1999) never explicitly motivates the specific c-structure associated with embedded verbs in the VP. Nevertheless, we can infer the structure given here from the examples and discussion there.

a. ...dass der Junge dem Mann geholfen hat.
COMP the.NOM.SG boy the.DAT.SG man helped has
'...that the boy has helped the man.'

b.



In addition to a nested VP structure, the tree in (22b) shows the subject appearing in S and the object inside the VP. Any deviations from the canonical word order implied by this structure are modelled using optional adjunction of objects to higher positions. Choi (1999) motivates this partially configurational structure for German by appealing to contrasts like the following: a verb and its object can be realized together in the Vorfeld (23a), whereas – it would appear – a verb and its subject cannot (23b).

(23) German(Choi 1999: §2.1, example 12)

a. [Dem Mann geholfen] hat der Junge.
the.DAT.SG man helped has the.NOM.SG boy
'Help the man, the boy did.'

b. * [Der Junge geholfen] hat dem Mann.
the.NOM.SG boy helped has the.DAT.SG man

Under Choi's analysis, this contrast follows straightforwardly by assuming that a VP, unlike an S, can be put in the Vorfeld.

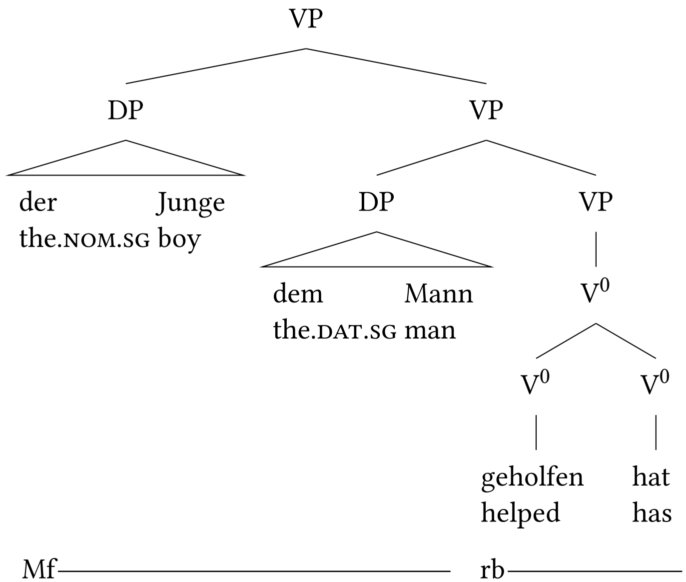
These analyses with nested VPs, which in principle could directly yield the pattern of nested dependencies we find in German, do *not* have a single node containing the whole right bracket and nothing else. Put differently, they do not include the verb cluster as such. This contrasts with the analyses we saw for the cross-serial dependency languages (Dutch, Zürich German) above, where the verb cluster exactly matched a $\overline{\mathrm{V}}$ node.

Proposals for Standard German that have a c-structure node corresponding to the verb cluster do exist in the LFG literature. One prominent such proposal is made by Berman (2003), who rejects Choi's claim that the German VP includes the object but excludes the subject, on the basis of data like (24), which, in contrast to (23b), is a successful example of Vorfeld realization of a verb with its subject.

(24) German (Berman 2003: §3.2.3.2, example 28a)
[Kinder gespielt] haben hier noch nie.
children played have here yet never
'Children have never played here.'

Instead of assuming that there are canonical positions for subjects and objects in S and VP respectively, and that only scrambled objects are adjoined, Berman does away with S completely, and always adjoins Mittelfeld arguments to VP. Furthermore, Berman posits that verbs in the verb cluster are combined by head adjunction. The structure of (23a) under this model is then (25).

(25)



Since the association of arguments with their predicates can no longer rely on positional grammatical function annotations, Berman (2003) argues that case is responsible for this association in German. This is modelled using the standard approach of conditional expressions that relate specific cases to specific functions, for instance, $(\downarrow \textsc{case})=\textsc{acc} \Rightarrow (\uparrow \textsc{obj})=\downarrow$. However, Berman does not discuss how this standard approach should be extended to allow embedded objects of coherently combined verbs in general, and it is not clear how one would correctly constrain the projection of multiple Mittelfeld objects onto f-structures that are embedded under one or more layers of XCOMP without resorting to nested VPs or functional uncertainty.[19]

We have seen that the (idealized) Dutch and German patterns are mirror images in terms of the order of the verbs in the verb cluster. The approach of Zaenen & Kaplan (2002) capitalizes on this by taking the mirror image of the $\overline{\text{V}}$ rule for Dutch in (20b) as the basis for their analysis of the German verb cluster (26).

(26)
$$\overline{\text{V}} \longrightarrow \begin{pmatrix} \overline{\text{V}} \\ (\uparrow \textsc{xcomp})=\downarrow \\ \neg((\uparrow \textsc{xcomp}^{+}\ \textsc{obj}) <_f (\uparrow \textsc{obj})) \end{pmatrix} \begin{matrix} \text{V} \\ \uparrow=\downarrow \\ \ \end{matrix}$$

As before, in the proposal for Dutch, the highest $\overline{\text{V}}$ node corresponds directly to the right bracket and functional uncertainty solves the relation of Mittelfeld material to embedded verbs without having to assume nested VPs.

2.1.2.3 Variation I already mentioned at the the beginning of this subsection that characterizing languages as having either cross-serial or nested dependencies is an oversimplification. For instance, both German and Dutch allow further variation in the ordering of elements in the verbal cluster. Moreover, in Zürich German – amongst other CWG languages – Mittelfeld and right bracket material can mix to some extent.

In German, *Oberfeldumstellung* (also known as *auxiliary flip*) can occur with three-verb combinations where V_1 is a perfect or passive auxiliary, and V_2 is itself a coherently combining verb that selects an infinitive. In this construction, the verb cluster has the order $V_1V_3V_2$, and IPP is triggered for V_2. Contrast the "regularly ordered" (27a) with the Oberfeldumstellung in (27b).[20]

[19] Berman (2003) partially sidesteps the issue by (tacitly) assuming that auxiliaries add features and do not create XCOMP embeddings. This means that, for instance, *lobte* 'praised' and *hat gelobt* 'has praised' both have their objects directly in the containing f-structure as OBJs. However, since not all coherently combining verbs can be analysed as auxiliaries and some clearly have enough lexical content to warrant their own PRED values, this does not completely address the problem.

[20] *Oberfeldumstellung* also occurs with longer verb clusters. Furthermore, there is a (possibly regional) construction called *Zwischenstellung* that has $V_3V_1V_2$. See Cook (2001) for empirical discussion and an analysis.

(27) German (Cook 2001: §1.4, example 1.31)

a. …dass ich dich kommen gesehen habe.
   COMP I you.ACC.SG come.INF seen have.PRS
   '…that I have seen you come.'

b. …dass ich dich habe kommen sehen.
   COMP I you.ACC.SG have.PRS come.INF see.INF

This word order variant is problematic for the nested VP models mentioned above (namely, Netter 1988; Choi 1999), since the verb cluster-initial finite verb "interrupts" the embedded VP. Models in which a c-structure node corresponds to the verb cluster (namely, Zaenen & Kaplan 2002; Berman 2003; Clément et al. 2002) have an easier time capturing such variation. An analysis of this variation can be found in Clément et al. (2002: in terms of c-structure) and in Cook (2001: in terms of the interaction between syntax and information structure). An OT-LFG analysis of verb order in Swiss German dialects is outlined in Seiler (2007).

Dutch verb clusters have so-called *participle climbing* and *particle climbing*, which refer to the realization of participles and particles to the left of the position expected from the principle of ordering by increasing embedding. Example (28) shows the different positions a particle can occupy in a three-verb cluster.

(28) Dutch
…dat Jan het liedje {mee} zal {mee} hebben {mee-}gezongen.
COMP Jan the song along will along have along-sung
'…that Jan will have sung along to the song.'

Kaplan & Zaenen (2003) adapt their earlier model of the Dutch Mittelfeld and verb-cluster to allow these and further variants, and to capture the IPP effect. Poortvliet (2015) is a further development of this model.

Zürich German has cross-serial dependencies, like Dutch, but in addition allows the nominal and verbal material to mix, as in (29).

(29) Zürich German
…das er sini chind wil mediziin la schtudiere.
COMP he his children wants medicine let study
'that he wants to let his children study medicine.'

Zaenen & Kaplan (1995) use the combination of functional precedence and functional uncertainty developed for Dutch to capture these data. Another case of mixing verbal and non-verbal material can be found in Standard German, which allows a variant of Oberfeldumstellung where V_1 precedes a collocational nominal complement of V_3. An analysis of this construction can be found in Cook (2001).

2.1.3 Nachfeld

The two options for adding Nachfeld material to the different c-structures of the clause given above are 1) adjunction to any of the nodes at the right periphery and 2) inclusion of one or more optional daughters on the right hand side of the relevant c-structure rules. Adjunction is used by Berman (2003), who assumes Nachfeld occupants (typically PPs, VPs or CPs) are right-adjoined at the VP level. Rohrer (1996), Clément et al. (2002) and Van der Beek (2005) model the Nachfeld as an optional daughter in the node covering the whole clause. Zaenen & Kaplan (1995; 2002), and Kaplan & Zaenen (2003) insert the optional daughter in the node covering the Mittelfeld/right bracket.

In Dutch and German, the non-finite complement of an incoherently combining verb appears in the Nachfeld. In Dutch, Mittelfeld placement of such a complement is ruled out (30), but in German it is allowed (see for instance Rohrer 1996: for examples).

(30) Dutch

Vf-	lb–	Mf———————				rb——	Nf———————			
Hij	had	{*geen	auto	te	kopen}	beloofd	{geen	auto	te	kopen}.
he	had	no	car		buy.TEINF	promised	no	car		buy.TEINF

'He had promised not to buy a car.'

To facilitate lexical specification of whether a verb combines coherently or not, and the formulation of placement restrictions on the non-finite verbal complement, Rohrer (1996), Zaenen & Kaplan (1995; 2002), and Kaplan & Zaenen (2003) associate coherence with selecting an XCOMP and incoherence with selecting a COMP. The relevant c-structure rule from Kaplan & Zaenen (2003) is an extension of (20a) and is given here in slightly simplified form as (31).

(31) $$\text{VP} \longrightarrow \underset{(\uparrow\ \text{XCOMP}^*\ (\text{COMP})\ \text{OBJ}){=}\downarrow}{\text{DP}^*}\quad \underset{\uparrow{=}\downarrow}{\overline{\text{V}}}\quad \left(\underset{(\uparrow\ \text{XCOMP}^*\ \text{COMP}){=}\downarrow}{\text{VP}}\right)$$

The optional rightmost daughter contains a non-finite complement in the Nachfeld, assigned COMP.[21]

The rule in (31) also allows for the so-called *third construction*, a marked construction in Dutch and German in which a dependent of an incoherently combined non-finite complement in the Nachfeld is realized in the Mittelfeld of the

[21]The distinction between COMP and XCOMP is that of complements that supply their own subject (closed complements) and complements that do not supply their own subject (open complements). Since non-finite COMPs do not have an overt subject, they therefore must have an f-structure subject PRED 'pro', whose interpretation is equated to one of the arguments of the selecting verbs using anaphoric control. See Dalrymple (2001: Chapter 12, §3) for a discussion of anaphoric control.

containing clause. In terms of word order, this construction therefore mixes properties of coherent and incoherent combination. An example of the third construction is in (32). Note that the lack of an IPP effect on *geprobeerd* 'tried' shows that we are dealing with incoherent combination.

(32) Dutch
Hij had een auto geprobeerd te kopen.
he had a car tried buy.TEINF
'He had tried to buy a car.'

The c-structure rule in (31) captures the third construction by the functional uncertainty-based grammatical function assignment of DPs in the Mittelfeld: the optional COMP in the path allows it to reach into an incoherently combined complement.[22] LFG analyses of the the German third construction are discussed in Rohrer (1996) and Kaplan & Zaenen (2003).

2.2 Topics related to the left periphery

2.2.1 Topicalization

In the context of the verb-second CWG languages, we mean by topicalization Vorfeld placement of material, in particular material that is *not* put there by default. Roughly, then, topicalization is Vorfeld placement of anything but the local subject. The term topicalization is used irrespective of whether the Vorfeld occupant is a topic or not. In both German and Dutch, the Vorfeld may be occupied by a categorially and functionally wide range of constituents. It is also a target position for material extracted from embedded clauses and phrases.

Berman (2003: Chapter 6) formally distinguishes two different types of topicalization for German, depending on whether the Vorfeld constituent is local to the matrix clause or whether a long-distance dependency is involved. Berman introduces this distinction on the basis of observations from weak cross-over, which will be discussed in Section 2.4.2, below. In either case, the Vorfeld is Spec-CP, and its definition is part of the straightforward c-structure rule in (33).

(33) $$\text{CP} \longrightarrow \underset{(\uparrow \text{DF})=\downarrow}{\text{XP}} \quad \underset{\uparrow=\downarrow}{\overline{\text{C}}}$$

[22] Kaplan & Zaenen (2003) are not concerned with CP complements – that is, finite complement clauses – but if these are assigned COMP as well, the analysis of the third construction sketched here will need to be further constrained to prevent lifting dependents from finite subordinate clauses into the Mittelfeld.

When material local to the f-structure projected from CP is put in the Vorfeld, Berman assumes information like case and agreement drives the association with grammatical function, just as it does in the Mittelfeld – see the earlier discussion of Berman's model in Section 2.1.2, around example (25). For long-distance dependencies, Berman posits the presence of a trace at the extraction site, annotated with an inside-out functional uncertainty equation to incorporate the f-structure of the Vorfeld constituents – which by (33) is the DF of the f-structure of the whole clause – into the extraction site's f-structure.[23]

German has embedded verb-second clauses with bridge verbs, provided the complementizer is absent, as in (34a). Extraction out of such embedded clauses is also allowed, on the condition that none of the clauses involved in the long-distance dependency has material in Spec-CP (that is, no intermediate clause has a Vorfeld occupant). This is shown in the contrast (34b,c).

(34) German (b,c from Berman 2003: §6.2.4, examples 23, 24)

a. Ich glaube, (*dass) der Hans sagte gestern, (*dass) die
   I believe COMP the.NOM Hans said yesterday, COMP the.NOM
   Maria hat den Peter eingeladen.
   Maria has the.ACC Peter invited
   'I think Hans said yesterday that Maria invited Peter.'

b. Den Peter glaube ich, sagte der Hans gestern, hat
   the.ACC Peter think I said the.NOM Hans yesterday has
   die Maria eingeladen.
   the.NOM Maria invited

c. * Den Peter glaube ich, gestern sagte der Hans, hat
   the.ACC Peter think I yesterday said the.NOM Hans has
   die Maria eingeladen.
   the.NOM Maria invited

Berman captures this restriction with an off-path constraint ¬(→ DF) on the functional uncertainty equation for extractions. Since only Spec-CP introduces DF in Berman's model,[24] this effectively rules out examples like (34c).[25]

[23]See Kaplan & Zaenen forthcoming [this volume] for more information on modelling long-distance dependencies using inside-out functional uncertainty.

[24]For this to hold, we need to understand DF as not including SUBJ, since subjects can be introduced in other positions in the clause, too. Indeed, as the example shows, there is no ban on subjects occurring anywhere in the path of a long-distance dependency, as long as they do not occur in the Vorfeld. An unfortunate side effect of taking DF as not including SUBJ would be that the special Vorfeld privileges of subjects, see Section 2.2.2, remain unmodelled.

[25]However, consider the following data:

In non-LFG work, Reis (1996) argues that sentences like (34b) are only apparent cases of extraction, and that they involve parenthetical constructions, instead. An LFG analysis of German parentheticals with bridge verbs is given in Fortmann (2006), although he does not consider the exact type of sentence discussed here.

German allows topicalization of VPs (as discussed above in Section 2.1.2.2, example 23) and, in the case of coherent combination, topicalization of partial VPs. For instance, in (35), the main verb and its accusative object are realized in the Vorfeld, whereas the dative object is in the Mittelfeld.

(35) Ein Märchen erzählen wird er ihr.
a fairy tale tell.INF will he her.DAT
'He will tell her a story.'

VP topicalization can in principle be modelled using the standard mechanisms. For instance, under the assumption that coherently combined verbal complements are XCOMPs and if we use outside-in functional uncertainty, a c-structure rule like in (36) implements topicalization of coherently combined VPs.

(36) $$\mathrm{CP} \longrightarrow \begin{array}{cc} \mathrm{VP} & \overline{\mathrm{C}} \\ (\uparrow \textsc{topic}){=}\downarrow & \uparrow{=}\downarrow \\ (\uparrow \textsc{xcomp}^{+}){=}\downarrow & \end{array}$$

If the rule for VPs allows partial VPs, rule (36) says very little about which material is required to be present in the fronted VP and which material can be left to be realized in-situ. Potentially, then, this also captures partial VP topicalization. Zaenen & Kaplan (2002) problematize two aspects of such a straightforward implementation: First, in the case of partial VP topicalization, the resulting f-structure for the whole VP contains the combined topicalized and in-situ material, and

(i) a. Ich denke hier, Sie sollten etwas präziser sein.
I think here you should somewhat precise.COMPARATIVE be
'I think here(:) you should be a bit more precise.'
b. Hier denke ich, Sie sollten etwas präziser sein.

(ii) a. Ich denke, Sie sollten hier etwas präziser sein.
'I think(:) you should be a bit more precise here.'
b. Hier denke ich, sollten Sie etwas präziser sein.

Although the off-path constraint against DFs gets rid of the form (i b) as a realization of the meaning of (ii), it does not block (ii b) as a realization of the meaning of (i). In other words, the off-path constraint itself leaves unexplained why the embedded V1 of (ii b) signals that it is involved in an extraction. A possible line of defence against this criticism is to appeal to a form of Economy of Expression: the embedded V1 is a slightly more complex structure than embedded V2, a complexity that is not needed for the relational information expressed in (i).

therefore there is no way to see at f-structure which part of the VP was topicalized. This is problematic for approaches to information structure that associate information status with f-structures. Secondly, the approach would erroneously allow examples like (37).

(37) * [Ihr ein Märchen] wird er erzählen.

[Ihr	ein	Märchen]	wird	er	erzählen.
her.DAT	a	fairy tale	will	he	tell.INF

'He will tell her a fairy tale.'

The preverbal material is here analysed as a headless VP, which is generally only allowed postverbally.[26]

Zaenen & Kaplan (2002) solve these problems by replacing unification with *subsumption* in the functional uncertainty annotation of Spec-CP, along the lines of (38).

(38) $$\mathrm{CP} \longrightarrow \begin{array}{c} \mathrm{VP} \\ (\uparrow \textsc{topic}){=}\downarrow \\ \downarrow \sqsubseteq (\uparrow \textsc{xcomp}^{+}) \end{array} \quad \begin{array}{c} \overline{\mathrm{C}} \\ \uparrow{=}\downarrow \\ \end{array}$$

This directly solves the first problem, since the information in TOPIC now no longer contains the f-structure for the whole VP, but only information projected from the material in the Vorfeld. It also solves the second problem, since, as shown in (39), the f-structure for the example in (37) under TOPIC now no longer meets LFG's *coherence* condition – it contains arguments but no predicate to select them.

(39) F-structure for (37), which violates the coherence condition:

$$\left[\begin{array}{ll}
\textsc{topic} & \left[\begin{array}{ll} \textsc{obj}_{\theta} & [\textsc{pred}\ \text{‘pro’}] \\ \textsc{obj} & [\textsc{pred}\ \text{‘fairy tale’}] \end{array}\right] \\
\textsc{pred} & \text{‘will}\langle\textsc{xcomp}\rangle\textsc{subj}\text{’} \\
\textsc{subj} & [\textsc{pred}\ \text{‘pro’}] \\
\textsc{xcomp} & \left[\begin{array}{ll} \textsc{pred} & \text{‘tell}\langle\textsc{subj,obj,obj}_{\theta}\rangle\text{’} \\ \textsc{subj} & \\ \textsc{obj}_{\theta} & [\textsc{pred}\ \text{‘pro’}] \\ \textsc{obj} & [\textsc{pred}\ \text{‘fairy tale’}] \end{array}\right]
\end{array}\right] \quad \sqsubseteq$$

[26] I am aware that this claim is too broad. See for instance Müller et al. (2012), who use headless Vorfeld VPs in their analysis of apparent multiple fronting in German. However, a discussion of exceptions to this rule would take us too far away from the main discourse here.

Finally, a reformulation of the *completeness* condition to take subsumption relations into account[27] allows the f-structures resulting from topicalizing a partial VP, as illustrated in (40).

(40) a. Ein Märchen erzählen wird er ihr.
a fairy tale tell.INF will he her.DAT
'He will tell her a story.'

b.
$$\begin{bmatrix} \textsc{topic} & \begin{bmatrix} \textsc{pred} & \text{`tell}\langle\textsc{subj},\textsc{obj},\textsc{obj}_\theta\rangle\text{'} \\ \textsc{obj} & \begin{bmatrix} \textsc{pred} & \text{`fairy tale'} \end{bmatrix} \end{bmatrix} \\ \textsc{pred} & \text{`will}\langle\textsc{xcomp}\rangle\textsc{subj}\text{'} \\ \textsc{subj} & \begin{bmatrix} \textsc{pred} & \text{`pro'} \end{bmatrix} \\ \textsc{xcomp} & \begin{bmatrix} \textsc{pred} & \text{`tell}\langle\textsc{subj},\textsc{obj},\textsc{obj}_\theta\rangle\text{'} \\ \textsc{subj} & \text{———} \\ \textsc{obj}_\theta & \begin{bmatrix} \textsc{pred} & \text{`pro'} \end{bmatrix} \\ \textsc{obj} & \begin{bmatrix} \textsc{pred} & \text{`fairy tale'} \end{bmatrix} \end{bmatrix} \end{bmatrix} \quad \sqsubseteq$$

Berman (2003: §3.3) solves the ungrammaticality of examples structurally similar to (37) by appealing to the endocentricity principles as formulated for LFG, which include the constraint that every lexical category must have an extended head (Bresnan et al. 2016: §7.2). In (37), neither the main verb *erzählen* nor the auxiliary *wird* c-command the material in the Vorfeld VP, which leaves the fronted VP without an extended head. Note that this solution would not be compatible with Zaenen & Kaplan's conception of the German clause, as they do not assume the German VP is necessarily endocentric.[28]

2.2.2 The Vorfeld subject-object asymmetry

In both German and Dutch, the main clause subject is privileged when it comes to realization in the Vorfeld. In LFG work, this can be modelled directly by annotating the Vorfeld position at c-structure explicitly with (↑ SUBJ)=↓ (Theiler & Bouma 2012), or by annotating it with (↑ DF)=↓, under the assumption that the grammaticalized discourse functions include the subject (Berman 2003: §3.2.1).[29] In an OT-LFG setting, Choi (2001) posits a high-ranking constraint SUBJECT-LEFT

[27] "An f-structure *g* is complete if and only if each of its subsidiary f-structures is either locally complete or subsumes a subsidiary f-structure of *g* that is locally complete" (Zaenen & Kaplan 2002: [24]).

[28] As mentioned in Section 2.1, above, Zaenen & Kaplan use the label S|VP for this category.

[29] I refer, however, to the comment in footnote 24, where I point out that the analysis of German embedded verb-second clauses discussed there relies on the contradicting assumption that SUBJ is *not* part of DF.

that prefers early realization of the subject, which includes realization in the Vorfeld.

One of the reflexes of this special relation between the Vorfeld and the main clause subject is a contrast like the following:

(41) German (Meinunger 2007)
Wo ist das Geld?
'Where is the money?'

a. Es liegt auf dem Tisch.
it lies on the table
'It is on the table.'

b. * Es hat Bernd auf den Tisch gelegt.
it has Bernd on the table put
'Bernd (has) put it on the table.'

Although the referent of the weak pronoun *es* has the same information status in both cases, it appears it can only occur in the Vorfeld as a subject, and not as an object. This would fit in with any of the approaches sketched above: being a subject alone is enough reason to be allowed in the Vorfeld, but – apparently – the weak pronoun *es* is incompatible with any of the other information structural functions of Vorfeld constituents.

The ban on object *es* in the Vorfeld is not categorical, however. Meinunger (2007: and references therein) gives many examples, and shows that the conditions under which object *es* can appear in the Vorfeld coincide with the conditions for the use of the homonymic Vorfeld expletive used in the presentational construction. In particular, the subject of the clause should not be topic (42).

(42) a. Es hat {jemand / *er} geklaut
it has someone he stolen
'Someone / *he has stolen it'

b. Es hat {jemand / *er} das Geld geklaut
EXPL has someone he the money stolen
'Someone / *he has stolen the money.'

Theiler & Bouma (2012) capture this behaviour by assuming that the common source of sentences with a Vorfeld object and of those with Vorfeld expletive *es* is the latter, the presentational construction. The presence of *es* in the Vorfeld signals exactly that the main clause subject is not topic. This construction is modelled using a c-structure rule that explicitly mentions the form of its first daughter.

(43) CP ⟶ NP $\overline{\text{C}}$
(↓ FORM) $=_c$ ES ↑=↓
((↑ SUBJ)$_\sigma$ DF) ≠ TOPIC
((↑ XCOMP*OBJ)=↓)

The optional assignment to object anywhere in the coherence domain of the clause is what allows object *es* to appear in the Vorfeld, under the same circumstances as the presentational construction's expletive. Expletive *es* also shows up in other situations; see Section 2.5.1 below.

2.2.3 Left dislocation

Thus far, we have not considered the *lead* field, which is positioned before the Vorfeld in the topological model given in (1) and which we characterized as reserved for material more loosely connected to the clause proper. We can distinguish several *left dislocation* phenomena that target the lead.[30] Two questions raised by this broadened view of the clause are 1) how tightly is left-dislocated material coupled to the clause, and 2) are there phenomena that we have treated as Vorfeld occupation that are better analysed as left dislocation with an empty Vorfeld?

The first question is central in Zaenen (1996), who studies contrastive left dislocation (44b) in Dutch and Icelandic, and asks whether this should be treated as topicalization (44a) or as a hanging topic (44c). The former counts as a well-integrated part of the clause, the latter has a looser relation to the clause.

(44) Dutch

a. Jan wil ik hier nooit meer zien.
   Jan want I here never again see

b. Jan die wil ik hier nooit meer zien.
   Jan DEM want I here never again see

c. Jan, ik wil hem hier nooit meer zien.
   Jan I want him here never again see
   'Jan, I never want to see (him) here again'

On the basis of categorial constraints on different kinds of left-dislocated material and on the basis of binding data, Zaenen concludes that contrastive left-dislocation patterns with topicalization in both languages. She proposes an analysis in which the contrastively left-dislocated material is connected to the clause

[30] We can likewise talk of right-dislocated material, positioned in the tail, but since we are not aware of any LFG discussions of right dislocation, we will ignore this phenomenon in this chapter.

using the same functional uncertainty equations we normally assume for topicalized material in the Vorfeld. The pronominal element in the Vorfeld in a contrastive left-dislocation is taken to be an (f-structure) adjunct to the left-dislocated material, and does not itself engage in the long-distance dependencies directly.

The second question underlies the discussion in Berman (2003: §7.4), which revolves around the contrast illustrated in (45).

(45) German (Berman 2003: §7.4, examples 58–61)

a. Dass die Erde rund ist, (das) hat ihn gewundert.
COMP the earth round is DEM.NOM has him.ACC surprised
'That the earth is round(, that) surprised him.'

b. Dass die Erde rund ist, (das) hat er nicht gewusst.
DEM.ACC has he not known
'That the earth is round, (that) he didn't know.'

c. Dass die Erde rund ist, *(dessen) war sie sich nicht bewusst.
DEM.GEN was she REFL not aware
'That the earth is round, of that she wasn't aware.'

d. Dass die Erde rund ist, *(darüber) hat sie sich gewundert.
about.DEM has she REFL surprised
'That the earth is round, that she was surprised about.'

In (45a,b), the fronted CP appears to alternate between being left-dislocated (with resumption) and appearing in the Vorfeld (without), whereas in (45c,d), the fronted CP must be left-dislocated. Berman gives an LFG interpretation of an existing approach in which this alternation is only apparent, and the CP is *always* left-dislocated. The difference in (45a,b) is that in German, nominative and accusative topics may be dropped from the Vorfeld. Whether the resumptive demonstrative pronoun is realized at c-structure or not, its f-structure presence is constant, and it is this which is assigned a grammatical function. The left-dislocated CP is connected anaphorically to the resumptive pronoun.

2.2.4 Split NPs

The split NP construction in German involves multiple NPs at different positions in the clause which together describe one argument. The first NP occurs in the Vorfeld of the top level clause, and a further NP occurs somewhere further down in the Mittelfeld of a possibly embedded clause. An example is (46a).

(46) German (Kuhn 2001: §1)

a. [Ein Schwimmbad] hat er sich noch [keins] gebaut.
   a swimming pool has he REFL yet none built.

b. Er hat sich noch [kein Schwimmbad] gebaut.
   he has REFL yet no swimming pool built
   'He hasn't built a swimming pool yet.'

A striking property of the two NPs *ein Schwimmbad* and *keins* is that they both have the form of complete NPs: the first NP includes a determiner for the head count noun, the second NP involves the independent form *keins* 'none', rather than the form *kein* 'no', which is used when a nominal head is realized in the NP itself (46b).

Kuhn (2001) proposes a solution in terms of an LFG variant with linear logic-based semantics. Semantically, the clause-internal NP is a regular elliptical NP; the job of the Vorfeld NP is to supply a property as antecedent. By assuming that the form of the NP can be syntactically determined completely in terms of c-structure, treating the two NPs as c-structurally independent, but projecting to the same f-structure, the form-related characteristics of the NPs can be made to follow.

2.2.5 Asymmetric coordination

Frank (2006) gives an analysis of asymmetric coordination puzzles in German, like the *subject gap with fronted finite verb* (SGF) coordination in (47a).

(47) German (Frank 2006: §3.2)

a. In den Wald [ging der Jäger] und [fing einen Hasen]
   in the.ACC forest went the.NOM hunter and caught a.ACC hare
   'The hunter went into the woods and caught a hare.'

b. * In den Wald ging der Jäger und einen Hasen fing.
   in the.ACC forest went the.NOM hunter and a.ACC hare caught

At first sight, this looks like a run-of-the-mill symmetric $\overline{\mathrm{C}}$ coordination. However, this is not the case, since the PP in the Vorfeld is unambiguously a directional PP, which is incompatible with the verb in the second conjunct. Furthermore, what *is* shared between the two conjuncts is the subject in the Mittelfeld of the first conjunct, which is not in a c-structural position that would lead us to expect this possibility.

Frank models SGF coordination using an optional annotation on the rule for symmetric CP coordinations, which shares the (grammaticalized) discourse function of the first conjunct with the coordination as a whole, and therefore, with the second conjunct.[31]

(48)
$$
\begin{array}{llccc}
\text{CP} & \longrightarrow & \text{CP} & \text{Coord} & \text{CP} \\
 & & \uparrow\in\downarrow & \uparrow=\downarrow & \uparrow\in\downarrow \\
 & & ((\uparrow \text{GDF}) = (\downarrow \text{GDF})) & &
\end{array}
$$

This extra annotation makes sure the completeness requirements in the second conjunct can be met. Frank also shows that this approach makes correct predictions with respect to the interpretation of the scope of quantified subjects in an SGF coordination. However, the formal account leaves unexplained why the second conjunct cannot have a fronted object, like the ungrammatical (47b). For this, Frank appeals to the discourse structure of SGF coordination: the second conjunct is conceptualized as part of the discourse-functional domain of the first. If the second conjunct were to have a Vorfeld TOPIC or FOCUS, this would signal that it sets up its own discourse-functional domain.

2.3 Topics related to the right periphery

2.3.1 Clefts

The it-cleft construction in Dutch involves a neuter weak pronoun (typically *het* 'it'), a copula, focused material, and a backgrounded finite clause in the Nachfeld. Van der Beek (2005: Chapter 2) shows that, with these ingredients, there are in fact two distinct cleft constructions: the intransitive cleft (an existential copula with extraposed complement clause, 49a) and the transitive cleft (an identificational copula with extraposed relative clause, 49b).

(49) Dutch (Van der Beek 2005: Figures 2.4 and 2.1)

a. Het is aan hem dat ze denkt.
   EXPL is on him COMPL she thinks.
   'It is of him that she is thinking.'

b. Het zijn jouw kinderen die huilen.
   it are your children REL cry
   'It is your children who are crying.'

Van der Beek shows that the two cleft types differ further in whether they involve an expletive or referential neuter pronoun, whether they tolerate pseudo-copulas

[31] Here, too, the grammaticalized discourse functions include the subject. In fact, in this construction, the shared material will always turn out to be the subject.

like *bleken* 'seem', or only forms of the verb *zijn* 'be', and whether the neuter pronoun is obligatorily the cleft subject or not.

Van der Beek models both cleft types with dedicated c-structure rules to capture the fixed position of the backgrounded clause, and to introduce the construction-specific annotations. This ensures, for instance, that the clause must be in the Nachfeld, and cannot be realized as one constituent with the pronoun or the focused material. In the intransitive cleft, both the expletive and the complement clause project to the SUBJ's f-structure, and the focused material is linked to an adjunct or oblique position in f-structure (*aan hem* 'on him' in 49a). In the transitive cleft, the relative clause is an adjunct of the referential weak neuter pronoun, which is the construction's TOPIC, and which is subject or object depending on properties of the focused material. The top level c-structure rules for the two constructions are given in (50).

(50) a. intransitive cleft:

$$\text{IP} \longrightarrow \underset{(\uparrow \text{SUBJ})=\downarrow}{\text{NP}} \quad \underset{\substack{(\uparrow \text{SUBJ})=\downarrow \\ (\downarrow \text{TYPE}) =_c \text{THAT} \\ (\uparrow \text{FOCUS})=(\uparrow \text{XCOMP}^* \{\text{ADJ} \in | \text{OBL}_\theta\})}}{\text{CP}}$$

b. transitive cleft:

$$\text{IP} \longrightarrow \underset{(\uparrow \text{SUBJ})=\downarrow}{\text{NP}} \quad \underset{\uparrow=\downarrow}{\bar{\text{I}}} \quad \underset{\substack{\downarrow \in (\uparrow \text{TOPIC ADJ}) \\ (\uparrow \text{TOPIC PRONTYPE})=\text{COP} \\ (\downarrow \text{TYPE}) =_c \text{REL} \\ (\uparrow \text{FOCUS})=(\uparrow \{\text{SUBJ}|\text{OBJ}\})}}{\text{CP}}$$

The TYPE feature of the CP-projected f-structures distinguishes relative clauses from complement clauses headed by *dat* 'that'. The PRONTYPE=COP restriction singles out a class of special copular pronouns which are N.3SG in form, but which show a wider range of agreement, whose existence can be argued for on independent grounds.

2.3.2 Correlative *es* and extraposed CPs

The constructions discussed in Berman et al. (1998) and Berman (2003: Chapter 8) also contain a neuter pronoun and a finite clause in the Nachfeld. In this case, the pronoun and the finite clause realize a propositional argument of the clause's main verb, and they can either appear on their own (51a,b) or together (51c), in which case the pronoun is referred to as a *correlative pronoun*.

(51) German (Berman et al. 1998: §1, examples 1 and 2)

a. Hans hat bedauert, daß er gelogen hat.
   Hans has regretted that he lied has
   Hans regretted that he lied.

b. Hans hat es bedauert.
   Hans has it regretted
   Hans regretted it.

c. Hans hat es bedauert, daß er gelogen hat.
   Hans has it regretted that he lied has

The central modelling assumptions made in both analyses are that the pronoun is referential – whether it occurs on its own or as a correlative together with the finite clause – and that the pronoun and the finite clause when they appear alone (51a,b), are OBJs. In the correlative pronoun construction (51c), however, it is the pronoun that has this grammatical function. The finite clause is then either seen as supplying further semantic restrictions to the interpretation of this pronoun (Berman et al. 1998: see also Section 2.2.4 above) or as an apposition to the pronoun (Berman 2003).

Berman (2003) also goes on to show that there is a range of correlative *es* data, and that despite superficial resemblances, different syntactic analyses are called for. For instance, Berman argues that in contrast to the data above, the psych verb *stören* 'disturb' in (52) has a different argument structure for the cases with correlative *es*: it either takes the finite clause as subject when the correlative is absent, or it takes *es* as subject and the finite clause as object when the correlative is present.

(52) German (Berman 2003: §8.1, example 2d)
     …weil (es) mich stört, dass sie den Hans liebt.
     because EXPL me bothers COMP she the.ACC Hans loves.
     '…because it bothers me that she loves Hans.'

2.4 Topics related to the ordering of dependents

2.4.1 Scrambling

Material in the Mittelfeld can be reordered to a certain extent. For instance, Dutch allows different orders of object and adverb (53).

(53) Dutch
     Anna heeft {de was} gisteren {de was} gedaan.
     Anna has the laundry yesterday the laundry done
     'Anna did the laundry yesterday.'

In German, the order of arguments themselves is free, as well. Example (54) shows one order for the arguments of a ditransitive, but the other five possible argument orders are grammatical, too.

(54) German (Haider & Rosengren 2003: §1, example 1)
…dass [das Objekt] [dem Subjekt] [den ersten Platz]
COMP the.NOM object the.DAT subject the.ACC first place
streitig macht.
competes.for
'…that the object competes for first position with the subject.'

In general, both scrambling over adjuncts and scrambling of arguments is sensitive to information structural effects, and – related to this – things like the referential form of the material involved. Choi (1999; 2001) explains German scrambling and clause-local fronting facts using an OT-LFG model in which constraints on canonical ordering of grammatical functions conflict with constraints on information structurally induced ordering. An information structural account of clause-local word order variation and quantifier scope in German is given in Cook & Payne (2006). The explanation of a deviating unmarked word order for a small group of ditransitives found in Cook (2006) is discussed below, in Section 2.5.2.

As far as the order of arguments in the Mittelfeld is concerned, Dutch is much more restricted than German. Nevertheless there is some variation. An OT account of the Dutch dative alternation, which also covers variation in the ordering of direct object and indirect object, is presented in Van der Beek (2005). Zaenen (1989) discusses scrambling of objects over subjects with Dutch experiencer verbs and passives of ditransitives, and argues for an effect of thematic role.

The cross-serial dependency pattern comes about when objects and verbs are in separate groups and both are ordered in the same fashion according to level of embedding. The verb cluster rule given in (20) above (Zaenen & Kaplan 1995), sorts embedding verbs before embedded ones, and explicitly forces the same order on the objects with the help of an f-precedence constraint. In her work on the order of objects in Dutch, Van der Beek (2005: §3.8) argues that this constraint should be treated as a violable OT constraint. An optionally higher ranking constraint prefers early realization in the Mittelfeld of a third person, inanimate pronoun. This constraint explains examples like (55), in which the object pronoun belonging to the embedded verb precedes the object belonging to the finite verb.

(55) Ik $\text{zag}_{\textsc{obj}:1}$ 't_2 Jo_1 $\text{doen}_{\textsc{obj}:2}$.
I saw it Jo do
'I saw Jo do it.'

As with all scrambling, this type of scrambling is less constrained in German, and may also apply to full NPs, and even involve scrambling of an embedded object over the main clause subject. I am however not aware of any LFG-related work on this.[32]

2.4.2 Weak cross-over

In German, scrambling and topicalization interact with binding between arguments, which results in data like (56). Note that the grammaticality judgements are relative to the co-indexations given in the examples.

(56) German (Berman 2003, §5.2, examples 10a, 11b, 10d, 11d, 27, 31; examples a–d below originally from Choi 1995)

a. …dass jeden$_i$ seine$_i$ Mutter mag.
   COMP everyone.ACC his mother likes
   '…that their$_i$ mother likes everyone$_i$.'[33]

b. Jeden$_i$ mag seine$_i$ Mutter.
   everyone.ACC likes his mother
   'Their$_i$ mother likes everyone$_i$.'

c. * …dass seine$_i$ Mutter jeden$_i$ mag.
   COMP his mother everyone.ACC likes

d. * Seine$_i$ Mutter mag jeden$_i$.
   his mother likes everyone.ACC

e. Jeden$_i$ sagte sie, habe seine$_i$ Mutter getröstet.
   everyone.ACC said she has.SBJV his mother consoled
   'Everyone$_i$, she said their$_i$ mother had consoled.'

f. * Jeden$_i$ sagte seine$_i$ Mutter, habe sie getröstet.
   everyone.ACC said his mother has.SBJV she consoled
   'Everyone$_i$, their$_i$ mother said she had consoled.'

[32] The term *embedded object shift* is van der Beek's term for this type of word order variation. In the literature on German, the phenomenon is sometimes discussed as a kind of *long(-distance) scrambling*, that is, scrambling across clause boundaries, although the view that the embedded object leaves its clause goes against the conception of coherent combining as clause union. In fact, in Lee-Schoenfeld (2007), this type of scrambling is taken as one of the hallmarks of coherence and thus of monoclausality.

[33] English seems to require the passive to achieve the intended bindings. The intended reading in (56a–d) is therefore more naturally given as *(that) everyone is liked by their own mother.*

Between dependents of the same predicate, an object may bind into the subject, provided it precedes it. It does not matter whether it precedes it in the Mittelfeld (56a) or by being moved into the Vorfeld (56b), even from an embedded clause (56e). However, as (56f) shows, an object cannot bind into an upstairs subject, even when it precedes it.

Berman (2003), using the framework of Bresnan (1998) and observations from Choi (1995), shows that the data in (56a–d) is straightforwardly explained by assuming that to bind a pronoun, an operator must either outrank it in terms of grammatical function – this isn't the case in any of these examples – or linearly precede it. The linear precedence constraint is satisfied in (56a,b), but not in (56c,d). However, example (56f) is problematic under this simple account, since the operator precedes the pronoun, but cannot bind it.

Berman therefore proposes to analyse long-distance dependencies using a trace, and to interpret the linear precedence requirement as if it includes this trace. The sentences in (56e,f) are then as in (57).[34]

(57) a. Jeden$_i$ sagte sie, habe ϵ_i seine$_i$ Mutter getröstet.
everyone.ACC said she has.SBJV his.NOM mother consoled

b. *Jeden$_i$ sagte seine$_i$ Mutter, habe ϵ_i sie getröstet.
everyone.ACC said his.NOM mother has.SBJV she consoled

In (57a), the operator's trace precedes the bound pronoun, so that the linear order requirement is met. In (57b), however, the trace follows the pronoun, which – under Berman's definition – means the operator as a whole does not precede it This results in the unavailability of the indexed reading.

Bresnan et al. (2016: §9.5) discuss the same data using a near identical framework. Although the difference in linear order of the bound pronoun and the operator trace between (57a) and (57b) is noted, the ungrammaticality of (57b) is ultimately explained by taking the binding domain of the operator to be the f-structure for the predicate *getröstet* 'consoled', irrespective of the operator's DF role in the matrix f-structure.[35] There is therefore no need to refer to the position of the trace to explain the long-distance dependency data. Under that analysis, it would appear that weak cross-over in German alone is not a reason to assume long-distance dependencies involve traces.

[34] Berman (2003) assumes that local arguments are adjoined to VP, in any order. This also applies to traces – the object trace may therefore appear before its clause-mate subject. In the examples in (57) we have inserted the trace as early as c-structurally possible.

[35] In contrast, Berman (2003: §5.2.6) explicitly considers the binding domain of the operator to be "extended to the matrix clause" because "it functions as a discourse function in the matrix clause" (p86).

Dalrymple et al. (2001)[36] give a trace-less account of the German cross-over data. Rather than considering the linear order of the binding operator and the bound pronoun, they consider f-precedence between two f-structures that are dependents of the same predicate, such that one contains the operator and the other the pronoun. In (56f), these f-structure siblings are the SUBJ (containing the pronoun) and the COMP (containing the operator) of *sagte*. Since the latter does not f-precede the former, the linear precedence requirement on binding is not met.

2.5 Topics related to mapping

2.5.1 Sentences "without a subject" in German

A recurring debate in German clausal syntax concerns the existence of true subjectless sentences. Berman (2003: Chapter 4) points out that it would appear that German has such sentences, given that 1) under her analysis, German does not have a dedicated subject position, 2) there are no oblique subjects in German (a common view, following for instance Zaenen et al. 1985, but contra the later Eythórsson & Barðdal 2005) and 3) there are sentences without nominatives, such as (58).

(58) German (Berman 2003: §4.2, examples 10a, 16b, 10d; indication of optionality of expletive mine)

a. …weil (*es) getanzt wurde.
   because EXPL danced was
   '…because people were dancing.'

b. …weil (*es) dem Mädchen geholfen wurde.
   because EXPL the.DAT girl helped was
   '…because the girl was being helped.'

c. …weil (es) mich friert.
   because EXPL me.ACC freezes
   '…because I'm cold.'

Note that each of these *can* occur without the expletive pronoun *es*, and the first two *must* occur without it.

Berman models clauses without a subject using argument structures without a SUBJ, and shows that these cases can be given an analysis in terms of Lexical

[36] This paper is a response to the trace-based proposals of Bresnan (1998) and Berman (2003). The latter was also published/circulated on earlier occasions, which explains the apparent anachronism.

Mapping Theory (LMT).[37] For the predicates involved in the examples above, we have the following LMT derivations:

(59) a. getanzt ⟨*agent*⟩ (lit. 'danced', impersonal passive)
[−*o*]
∅

b. geholfen ⟨*agent*, *beneficiary*⟩ ('helped', passive, lexical case)
[−*o*] [+*o*]/DAT
∅ OBJ_θ

c. frieren ⟨*experiencer*⟩ ('be cold', active intransitive, lexical case)
[+*o*]/ACC
OBJ_θ

The question remains, then, why the expletive is not allowed in (58a,b), whereas it is in (58c). Berman adopts the analysis that German verbal agreement morphology is distinct enough to contribute subject features. Thus, the f-structures for the sentences in (58) all contain subjects. This way, German can be analysed as meeting the *Subject Condition*, which says that every f-structure with a predicate must contain a SUBJ.[38] Inserting a subject expletive would then be ruled out as a violation of Economy of Expression. It follows that the optional *es* in cases like (58c) is selected for: verbs like *frieren* have an alternative specification like the one in (60).

(60) *frieren* (↑ PRED) = 'be-cold⟨OBJ_θ⟩SUBJ'
(↑ SUBJ FORM) $=_c$ ES_

2.5.2 Mapping explanations of variation

Zaenen (1993) is concerned with (the nature of) the unaccusative/unergative distinction in Dutch. One of the challenges in the characterization of unaccusativity in Dutch is that it not only applies to intransitives, but also to a subset of transitive experiencer verbs. Consider the examples in (61), which shows two intransitives, two transitives with the experiencer as the object, and a transitive with the experiencer as the subject. The selection of a form of *zijn* 'be' instead of *hebben* 'have' as the perfect auxiliary is given here as the reflex of unaccusativity.

[37] See Findlay & Kibort forthcoming [this volume] for more information on Lexical Mapping Theory.

[38] This Subject Condition formulation pertains to f-structure. In other contexts, for instance in Bresnan et al. (2016: §14.4), the Subject Condition is taken to be a constraint on argument structures. It is clear that under Berman's view such a constraint does not hold for German.

(61) Dutch

a. Zij *is / heeft gewerkt.
   she is has worked

b. Zij is / *heeft gestorven.
   she is has died

c. Zij *zijn / hebben haar geirriteerd.
   they are have her irritated

d. Zij zijn / *hebben haar bevallen.
   they are have her pleased

e. Zij *is / heeft hen gevreesd.
   She is has them feared.

Zaenen shows that it is possible to give semantic correlates of unaccusativity, and discusses which phenomena can be related directly to unaccusativity in Dutch (namely, auxiliary selection, prenominal attributive use of perfect participle) and which only relate indirectly (impossibility of impersonal passive). Her analysis is formalized in terms of a variant of LMT that does not rely on thematic roles to determine the intrinsic classifications of a predicate's arguments. Instead, Zaenen incorporates Dowty's (1991) proto-roles into LMT using the following simple rule: a participant that has more proto-agent than proto-patient properties is marked [$-o$], otherwise the participant is marked [$-r$]. The LMT alternative is further spelled out to allow derivation of grammatical function assignments for the data in (61): the subjects in examples (61a,c,e), with *hebben* 'have', come from intrinsic [$-o$] markings, whereas the subjects in (61b,d), with *zijn* 'be', come from intrinsic [$-r$]. The choice of auxiliary can be correctly modelled by referring to the intrinsic markings of the subject. Kordoni (2003) discusses analysing the German locative alternation in terms of Zaenen's mapping account.[39]

Another variation which is shown to be driven by lexical semantic differences that affect mapping are the so-called "high" versus "low" datives in German. Although arguments in the German Mittelfeld are readily scrambled, there is an unmarked order, which can be detected by studying information structural and quantifier scoping properties. Between objects, the unmarked order is generally DAT before ACC (62; "high dative"). However, a smaller number of verbs show ACC before DAT (63; "low dative"), and for a couple of verbs both orders appear to be unmarked. In the examples, superscript M marks the marked variant.

[39] It should be noted that Dowty (1991) talks about the *English* locative alternation in terms of proto-roles in depth.

(62) German (Cook 2006: §1, examples 1–2)

a. Es hat ein Mann [einem Kind] [ein Buch] geschenkt.
EXPL has a.NOM man a.DAT child a.ACC book given
'A man gave a book to a child (as a present).'

b. ᴹ Es hat ein Mann [ein Buch] [einem Kind] geschenkt.

(63) a. ᴹ Es hat ein Polizist [einer Gefahr] [einen Zeugen]
EXPL has a.NOM policeman a.DAT danger a.ACC witness
ausgesetzt.
exposed
'A policeman has exposed a witness to a danger.'

b. Es hat ein Polizist [einen Zeugen] [einer Gefahr] ausgesetzt.

Cook (2006) demonstrates that the different unmarked orders can be related to differences in lexical semantics, which in turn give rise to thematic alternations. For the alternating verbs, it is shown that the different word orders prefer different readings in line with the general lexical semantic observations. All meanings/word orders involve an agent and a patient/theme, which under standard LMT assumptions are mapped to SUBJ and (accusative) OBJ respectively. In addition, the DAT-ACC order is associated with a bene-/maleficiary role, which is mapped to a (dative) OBJ_θ. The ACC-DAT order, however, involves a third participant which is a goal or a location and which gets mapped to a (dative) OBL_θ. Cook argues that the unmarked order of complements in the German Mittelfeld is OBJ_θ-OBJ-OBL_θ. The apparent word order variation is thus a fixed word order seen in the light of the unmarked order of grammatical functions. Cook extends her account to explain the compatibility of the different datives with the *kriegen*-passive, which can be used with a selection of verbs to promote the dative argument to subject.

2.5.3 Transitivity of reflexives

Lexically conditioned reflexives in German and Dutch show up in a range of situations. The simplex reflexives *sich* in German and *zich* in Dutch appear for instance in clauses with transitive verbs with co-referring arguments (64a)/(65a),[40] in anticausatives (64b)/(65b), and in inherent reflexives (64c)/(65c).

[40] The class of grooming verbs is part of a larger class of transitive verbs that, exceptionally, allows the simplex reflexive. In general, the complex reflexive, *zichzelf* / *sich selbst* is available to realize reflexive objects with transitive verbs. This exception is what justifies treating these reflexives as being lexically specified.

(64) a. Max rasiert sich.
Max shaves REFL
'Max shaves himself.'

b. Die Tür öffnet sich.
the door opens REFL
'The door opens.'

c. Max schämt sich.
Max is.ashamed REFL
'Max is ashamed.'

(65) a. Max scheert zich.
Max shaves REFL

b. De deur opent zich.
the door opens REFL

c. Max schaamt zich.
Max is.ashamed REFL

In a contrastive study of reflexivization, Sells et al. (1987) distinguish three kinds of transitivity: 1) c-structure transitivity – the reflexive is an independent constituent, 2) f-structure transitivity – the syntactic predicate selects an OBJ, 3) semantic transititivity – the referential identity of the arguments is accidental. Interestingly, German and Dutch simplex reflexives receive different analyses: they are both considered to be transitive in terms of c-structure, and intransitive in terms of semantics, but Sells et al. analyse the German reflexives as f-structurally intransitive, and the Dutch reflexives as f-structurally transitive. This is based upon the contrast in (66): the German reflexive can appear in an impersonal passive, whereas the non-reflexive counterpart is ruled out. The reflexive thus patterns with intransitives. The Dutch counterpart is not well-formed, which would suggest Dutch reflexives pattern with transitives.

(66) a. German (Sells et al. 1987: §2.4, example 74)
Jetzt wird {sich / *ihn} aber gewaschen!
now is REFL him however washed
'But now it is time to wash yourself!'
Not: '…to wash him!'

b. Dutch
* Nu wordt (er) {zich / hem} gewassen!
now is EXPL REFL him washed

Sells et al. (1987) model reflexives of verbs like (64a), (65a) with a lexical rule which maps a transitive verb (in all the three senses above) to a reflexive verb. For German, this involves leaving out the object slots, and marking the predicates with '[a] feature F, forcing them to combine with the reflexive element.'[41] The resulting intransitive can then serve as input for the lexical rule for the impersonal passive. For Dutch, the reflexivization rule involves moving the thematic object to a non-thematic object slot, marked to be filled by a simplex reflexive.

The analysis of Sells et al. crucially relies on the use of lexical rules to take care of mapping. Modern LFG work would rely on a variant of LMT. Data like those in (66) then also receive a different status, as LMT does not model the (impersonal) passive as a rule to be applied on the output of another rule. Alencar & Kelling (2005) propose an analysis of the whole range of data in (64a–c) in terms of LMT, and explicitly reject the importance of the contrast in (66). Instead, they offer additional data to support the conclusion that the German reflexive *is* transitive at f-structure. Their LMT analysis of the data in (64), above, is summarized in (67).

(67)
a. rasieren/öffnen ⟨*agent*, *theme*⟩ ('shave'/'open', transitive)
$\qquad$ [−*o*] [−*r*]
$\qquad$ SUBJ OBJ

b. (sich) rasieren ⟨$agent_i$, $theme_i$⟩ ('shave', reflexive of transitive)
$\qquad$ [−*o*] [−*r*]
$\qquad$ SUBJ_i OBJ_i[REFL]

c. (sich) öffnen ⟨$theme_i$⟩ $\_{i}$ ('open', anticausative)
$\qquad$ [−*r*] [−*r*]
$\qquad$ SUBJ_i OBJ_i[REFL]

d. (sich) schämen ⟨$theme_i$⟩ $\_{i}$ ('be ashamed', inherent reflexive)
$\qquad$ [−*r*] [−*r*]
$\qquad$ SUBJ_i OBJ_i[REFL]

In the last two cases, the reflexive is expletive.[42]

[41] The reflexive itself carries a constraining equation checking this feature F to make sure it is only combined with predicates that have undergone reflexivization. It is not spelled out in the article how the presence of the reflexive would be enforced technically, however. One solution is to let the verb and the reflexive be co-heads that check for the presence of each other using constraining equations.

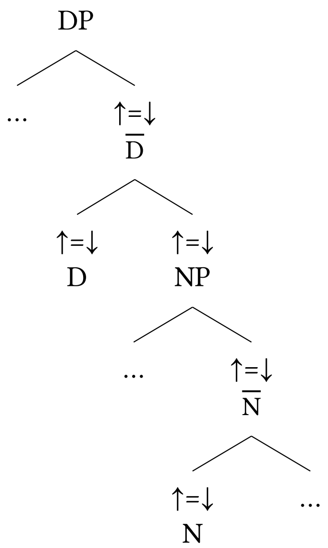
[42] The notation with indices to indicate reflexivity is taken from the paper. Note that, as they are also on expletives, these indices should not be interpreted as semantic co-reference.

3 LFG analyses in the nominal domain

3.1 Overall shape of nominal constituents

The CWG nominal domain has received a lot less attention than the clausal domain in LFG. The authors that have concerned themselves with the nominal domain in more detail all assume a DP analysis (Berman & Frank 1996, Part I, Chapter 3; Dipper 2003, Chapter 7; Strunk, 2004; 2005). The general shape of the nominal constituent is characterized by the familiar (68).

(68)



In Spec-DP, elements like pre-determiners (Dutch: ***al*** *die mensen* 'all those people'), prenominal genitives (that is, non-pronominal possessives; German: ***Karls*** *Auto* 'Karl's car'), and non-genitive prenominal possessors (Low Saxon: ***de'n Jung*** *sien Vadder*, lit. 'the boy his father') can occur. We refer to the discussion of possessives in Section 3.2.1, below, for more elaborate examples. D holds determiners and pronouns, although Dipper (2003) assigns pronouns proper (in contrast to determiner-like pronouns) to a category Pron, which is the single daughter of $\overline{\text{D}}$.

The NP contains mostly lexical material. In Dipper (2003: Chapter 7, the theoretical discussion), the class of adjectival quantifiers (see Section 3.2.2, below) appears in Spec-NP, whereas other attributive adjectives appear as sisters to N. Berman & Frank (1996), however, assume that attributive adjectives are left-adjoined to NP (not shown in the schematic tree), whereas oblique and clausal complements are in Comp-NP, on the right.

This overall picture is slightly different in Dipper (2003: Chapter 8, the computational implementation) and the derived Dipper (2005), which describe a flat DP/$\overline{\text{D}}$, under which predeterminers, determiners, prenominal genitives and adjectival quantifiers (that is, the material in Spec-DP, D, and Spec-NP in 68 above) all appear as sisters of NP.[43]

3.2 Topics in noun phrase syntax

3.2.1 Possessives in Low Saxon

Low Saxon has, amongst others, the range of possessive constructions illustrated in (69–71).

(69) Low Saxon (Strunk 2004: §2.2, examples 2.52 and 2.84)

a. sienen Weg
his.M.SG.ACC way
'his way'

b. jeedeen Oort kreeg [sienen], [...].
every kind got his.M.SG.ACC
'Every kind got its own, [...].'

(70) Low Saxon (Strunk 2004: §2.3, examples 2.112 and 2.177)

a. de'n Jung sien Vadder
the.M.SG.ACC boy his.M.SG.NOM father
'the boy's father'

b. Korl sien
Korl his.N.SG.NOM
'Korl's'

(71) (Strunk 2005: §6, example 61)
[Hinnerk=s Huss] iss groote den [Antje=s].
Hinnerk=POSS house is bigger than Antje=POSS
Hinnerk's house is bigger than Antje's.

Example (69a) contains a possessive pronoun in combination with a noun expressing the possessum. The possessor is anaphorically given as the referent of

[43] This analysis has the explicit goal of 'serv[ing] as the base of a robust and efficient implementation' (Dipper 2005: p101), but its status as a theoretical claim remains a bit unclear. The question of whether a deeper/different analysis would have been preferred in a more theoretically oriented analysis and whether this flat structure should mostly be seen as an operationalization of a deeper structure, is unfortunately not discussed.

the possessive pronoun. Example (70a), is a case of *possessive doubling*: as before we have a possessive pronoun and a noun, but now the possessive pronoun is directly preceded by a nominal in the accusative, which explicitly supplies the possessor. Finally, the first possessive expression in (71) is an instance of an -*s* marked nominal realizing the possessor, followed by the unmarked possessum. Examples (69b), (70b), and the second possessive in (71) show that the possessum can be elided in each of these constructions.

Strunk (2005)[44] models the three constructions in a unified way, crucially relying on optionally specified PRED 'pro' values to capture the differing amounts of explicitly realized referential information. He assumes entries for possessive pronouns along the lines of (72a) and the possessive clitic in (72b), as well as the top level rule for the DP in (72c).

(72) a. *sien* D ((↑ PRED) = 'pro-of⟨POSS⟩')
(↑ AGR) = M.SG
(↑ CASE) = NOM
((↑ POSS PRED) = 'pro')
(↑ POSS MARKING) = +
(↑ POSS AGR) = M.3SG
(↑ POSS CASE) = ACC

b. =*s* D ((↑ PRED)= 'pro-of⟨POSS⟩')
(↑ POSS MARKING) = +

c. DP ⟶ DP $\overline{\text{D}}$
(↑ POSS)=↓ ↑=↓
(↑ POSS MARKING) $=_c$ +

The entry for a possessive pronoun constrains two "regions", the first constraining the f-structure ↑ for the DP it heads – the possessum – and the second constraining the f-structure (↑ POSS) for the possessor. Both regions have agreement constraints, ↑ from inflectional morphology, (↑ POSS) from the choice of the pronominal root. The two regions also each have an optional equation defining PRED to be a pro-form. The four ways to satisfy these constraints correspond to the four cases in (69) and (70). Finally, Strunk (2005) analyses possessive -*s* as a clitic which also sits in D. Like the possessive pronoun this clitic can be realized with or without a possessum in NP. Unlike the pronoun it must be preceded by a DP that supplies a possessor, which explains the absence of an optional (↑ POSS PRED)='pro' in this entry.

[44] Strunk (2004) is an earlier version of this work, which contains a wealth of material on Low Saxon possessives.

Berman & Frank (1996) and Dipper (2003) discuss the standard German prenominal genitive possessive construction, exemplified in (73a). In contrast to the clitic-in-D analysis given above for the (perhaps only superficially) similar Low Saxon possessive *-s*, these authors put the prenominal material completely in Spec-DP. Berman & Frank (1996) also treat the colloquial German possessive doubling construction, which involves a prenominal dative,[45] found in (73b). As in the analysis proposed for the Lower Saxon counterpart above, the (now dative-marked) possessor is located in Spec-DP, and the possessive pronoun in D.

(73) German (Berman & Frank 1996: §3.1.2, example 136, 141)

a. Peters (*das) Haus
   Peter's the house
   'Peter's house'

b. der Frau *(ihr) Haus
   the.F.SG.ACC woman her house
   'the woman's house'

In the analysis put forward by Berman & Frank, the main *structural* difference between the the German prenominal genitive and prenominal dative is that the former requires D to be empty (73a), whereas the latter requires D to be filled (73b). A binary feature on head realization is used to control this.

3.2.2 Declension and the status of quantifiers

Inside the German DP, determiners, adjectives and nouns show agreement with respect to gender, number and case. *Declension* is another agreement dimension, found between determiners and adjectives. Determiners have inherent declension – they can be categorized as inflected (strong declension), uninflected or mixed. In the latter case some cells are inflected/strong and others are not. Inflected adjectives, on the other hand, have strong (more distinctive morphology) and weak (less distinctive morphology) declension paradigms. Adjective declension agrees with the inherent declension of the determiner in the following way:

(74) inflected (strong) determiner: weak adjective
uninflected or no determiner: strong adjective

[45] Kasper (2014: 58–59) calls the prenominal dative possessive a "non-standard German [construction] that is completely absent from the standard but can be found in almost all regional varieties/ dialects". Berman & Frank (1996) discuss the prenominal dative together with the prenominal genitive, and note that the former "allerdings eher in der gesprochenen Sprache auftritt" [is however more likely to occur in spoken language] (p59).

This phenomenon is illustrated in (75–76). For reasons of exposition, the inflection is made explicit and we use a zero morpheme to mark the lack of inflection. Note that *ein* is a member of the mixed declension class and appears both inflected (75b) and uninflected (76b).

(75) German (data from Dipper 2005: §3.2, presentation/glosses changed)

a. d-er süß-e rot-e Wein
   the-M.SG.NOM sweet-WEAK.SG.NOM red-WEAK.SG.NOM wine(M)

b. ein-em süß-en Wein
   a-M.SG.DAT sweet-WEAK.SG.DAT wine(M)

(76) a. süß-er rot-er Wein
   sweet-STRONG.M.SG.NOM red-STRONG.M.SG.NOM wine(M)

b. ein-∅ süß-er Wein
   a-M.SG.NOM sweet-STRONG.M.SG.NOM wine(M)

Determiners that do not inflect at all (for instance, *allerlei* 'every kind', *solcherlei* 'such') are not of the strong declension, and adjectives that do not inflect at all (*lila* 'purple', *rosa* 'pink') are ambiguous between strong/weak declension.

Dipper (2005) models the facts about declension in the following way: The f-structure projected from the DP has a feature DECL, which strong determiners define to be ST-DET and strong adjectives, ST-ADJ. This captures the fact that these two are never seen together. Weak adjectives constrain their containing DP's f-structure by DECL $=_c$ ST-DET, and therefore only co-occur with strong determiners. Uninflected adjectives and determiners do not constrain the DECL feature at all.

As seen in (75a) and (76a), when the DP/NP contains multiple adjectives, they show identical declension. Dipper (2005) uses this fact to address the issue of the categorial status of quantifiers like *alle* 'all' and *mehrere* 'multiple', for which it is difficult to decide whether they are determiners or adjectives. By inspecting the declension of adjectival material in the presence of a quantifier, Dipper is able to clearly distinguish determiner-like and adjective-like quantifiers.

3.2.3 Preposition-determiner contractions

German has a number of lexical preposition-determiner (P-D) contractions, such as *zum* 'to the', and *vom* 'of the', shown in (77).

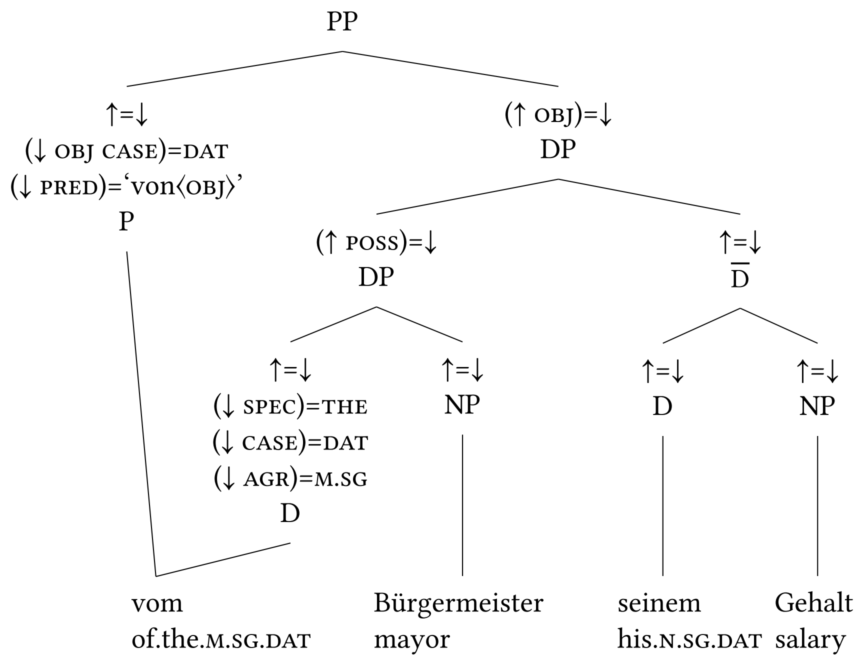
(77) a. zum König
   to.the.M.SG.DAT king
   'to the king'

b. vom Bürgermeister seinem Gehalt
of.the.M.SG.DAT mayor his.N.SG.DAT salary
'of the mayor's salary'

In Berman & Frank (1996), P-D contractions are treated as prepositions that not only constrain their object DP in a P-like manner – it must have a given case – but also in a D-like manner – it is marked definite and has certain agreement features. Most importantly, the object DP may not itself realize its own D. This is enforced using the binary head realization feature also used in the analysis of possessives, sketched in Section 3.2.1 above.

The example in (77b) shows that this picture is too simplistic: here, the D-like properties do not constrain the object DP, but the prenominal dative of this DP. It is inside this prenominal dative that D is left unrealized, and not in the object DP itself, which has *seinem* in D. The correct generalization about P-D contractions must therefore include that the D inherent in the contraction corresponds to a D leftmost in the object DP, and need not be the object DP's head. Wescoat (2007) gives an analysis in terms of lexical sharing that addresses exactly these points. In Wescoat's lexical sharing model, one lexical terminal can correspond to multiple adjacent preterminals. A slightly simplified analysis of (77b) is given in (78).

(78)



The fact that the preterminals projected from *vom* need to be adjacent solves the problem noted above that the generalization about P-D contractions needs to include reference to the left edge of the object DP. In the paper, Wescoat describes further constraints on the function of the projected D inside the object DP.

3.2.4 Indeterminacy of case

The German nominal inflection paradigms show pervasive syncretism. These syncretic forms can either be ambiguous or indeterminate. Ambiguous forms can be used in different contexts, but they can only be in one paradigmatic cell at a time. So *sie* in (79), which in isolation is ambiguous between plural 'they' and feminine singular 'she', can be used in either way (79a,b), but not as both at once (79c). Indeterminate forms *can* function as if they are in different cells simultaneously. For instance, *Papageien* 'parrots', which is indeterminate for case, can at the same time be selected as an accusative object and a dative object (80).

(79) a. Sie hilft Papageien.
she helps parrots

b. Sie helfen Papageien.
they help parrots

c. * Sie hilft und helfen Papageien.

(80) a. Sie hilft Papageien.
she helps parrots.DAT

b. Sie findet Papageien.
she finds parrots.ACC

c. Sie findet und hilft Papageien.

Although a simple disjunctive defining equation for a feature suffices for the ambiguous cases, this is not enough to achieve indeterminacy, since a disjunction does not change the fact that a feature can only have one value at a time. Dalrymple et al. (2009) represent indeterminate features as bundles of binary features, one for each of the values in the paradigmatic dimension. Compatibility with values is given as a disjunction of *positive* specifications, incompatibility as *negative* specifications. Two example lexical specifications are given in (81).

(81) a. *Papageien* (↑ CASE {NOM|GEN|DAT|ACC}) = +

b. *Männer* (↑ CASE {NOM|GEN|ACC}) = +
(↑ CASE DAT) = −

These specifications state that *Papageien* is completely indeterminate with respect to case (81a), whereas *Männer* is non-dative, but otherwise indeterminate with respect to case (81b).

A selecting element then expresses its case requirements in positive terms only. The entries in (82) illustrate this.

(82) a. *hilft*[46] (↑ OBJ CASE DAT) = +
b. *findet* (↑ OBJ CASE ACC) = +

Since the case feature bundles for *Papageien* defined in (81a) can satisfy both these requirements at the same time, we can capture the coordination of (80c). Dalrymple et al. show that this approach can also deal with additional material in the DP like adjectives, which further constrain the case value, and with verbs which themselves are indeterminate about their case requirements on selected arguments.

4 Concluding remarks

This chapter has presented an overview of Lexical-Functional Grammar studies of Continental West Germanic languages. The majority of the work discussed here has dealt with German clausal syntax, followed by discussions of Dutch clausal syntax. This reflects the status of the LFG field as a whole – the nominal domain has received a lot less attention than the clausal/verbal domain, see Börjars & Lowe forthcoming [this volume] – but it also reflects the fact that the other CWG languages – possibly, but not only, minority, regional, and/or non-standardized languages – do not feature prominently in the LFG literature. I hope that the discussion of existing work on the syntax of the two "big" CWG languages in the current chapter may inspire further application of LFG to the other members of the family.

Obviously, not every LFG study that touches upon CWG has been mentioned in this chapter. There are some larger blind-spots that I wish to mention here.

- Bögel (2015) develops an LFG model of the prosody-syntax interface. Recent papers contain applications to Swabian (Bögel & Raach 2020; Bögel 2021) and Standard German (Bögel 2020). See also Bögel forthcoming [this volume].

[46] We follow here the presentation in the paper and gloss over the fact that *helfen* 'help' might be better analysed as taking an OBJ_θ rather than an OBJ, which would complicate modelling the coordination.

- A number of authors have used OT in combination with LFG, especially in the domain of word order variation and information structure. Examples are Choi (1999, 2001), Cook (2001), Cook & Payne (2006),Van der Beek (2005), and Seiler (2007). These have been mentioned in the text, but were not discussed in any detail. OT-LFG is dealt with in Kuhn (forthcoming), and information structure is treated in Zaenen forthcoming [this volume].

- German is blessed with a wide-coverage LFG grammar, implemented in the context of the ParGram project. This grammar can be queried in the interactive XLE-WEB interface at https://clarino.uib.no/iness/xle-web.[47] The project page for the ParGram project in Germany, https://www.ims.uni-stuttgart.de/en/research/projects/pargram/,[48] contains older references. The research activities in and around this project have resulted in a long list of publications. Some of that work has already been discussed above. I will here list a small selection of further papers that also have direct relevance for theoretical debates: Forst & Rohrer (2009) and Kuhn et al. (2010) discuss problems in the analysis of German VP coordination; Rehbein & van Genabith (2006) and Forst et al. (2010) deal with the implementation of particle verbs; Forst (2006) is a "grammar writer's" contribution to the COMP-debate. The desire for parallel structures in the context of ParGram is one of the forces behind the auxiliaries-as-features style of syntactic analysis in LFG. An early contribution and implementation can be found in Butt et al. (1996). Computational work on LFG is the topic of several chapters in Part V of this volume.

Omitting these studies from the main text was a conscious choice, intended to keep the chapter accessible by not introducing too much conceptual machinery and too many problem domains. I made this choice with the knowledge that their topics would be touched upon in other chapters. At the same time, I wish to underline their importance, because exactly the fact that they span multiple domains and methods means that they are excellent demonstrations of the flexibility and precision that LFG offers.

Acknowledgements

I thank Bozhil Hristov and two anonymous reviewers for the detailed and valuable comments that greatly helped to improve this chapter. Moreover, I am grate-

[47]Consulted July 2022
[48]Consulted July 2022

ful to Mary Dalrymple for editing this volume, inviting me to contribute, and her extraordinary patience during the writing process.

Abbreviations

Besides the abbreviations from the Leipzig Glossing Conventions, this chapter uses the following abbreviations.

CWG	Continental West Germanic	OT	Optimality Theory
EXPL	expletive	rb	right bracket
IPP	infinitivus participio	TEINF	(Dutch) infinitive with marker *te*
lb	left bracket	Vf	Vorfeld
Mf	Mittelfeld	ZUINF	(German) infinitive with marker *zu*
Nf	Nachfeld		

References

Alencar, Leonel F. de & Carmen Kelling. 2005. Are reflexive constructions transitive or intransitive? Evidence from German and Romance. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '05 conference*, 1–20. Stanford: CSLI Publications.

Andersen, Christiane. 2016. Syntax in contact. Word order in a contact variety of German spoken in Eastern Siberia. *Journal of Language Contact* 9(2). 264–292. DOI: 10.1163/19552629-00902003.

Bech, Gunnar. 1955/7. *Studien über das deutsche verbum infinitum*. Copenhagen: Det Kongelige Danske Videnskabernes Selskab.

Berman, Judith. 2003. *Clausal syntax of German*. Stanford: CSLI Publications.

Berman, Judith, Stefanie Dipper, Christian Fortmann & Jonas Kuhn. 1998. Argument clauses and correlative 'es' in German – Deriving discourse properties in a unification analysis. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '98 conference*. Stanford: CSLI Publications.

Berman, Judith & Anette Frank. 1996. *Deutsche und Französische Syntax im Formalismus der LFG* (Linguistische Arbeiten 344). Tübingen: Niemeyer. DOI: 10.1515/9783110955354.

Biberauer, Theresa. 2009. How "well-behaved" is Afrikaans? V2 in Modern Spoken Afrikaans. In Hans den Besten, Frans Hinskens & Jerzy Koch (eds.), *Afrikaans: Een drieluik*, 176–208. Amsterdam/ Münster: Stichting Neerlandistiek VU/Nodus Publikationen.

Bögel, Tina. 2015. *The syntax-prosody interface in Lexical Functional Grammar*. Konstanz: University of Konstanz. (Doctoral dissertation).

Bögel, Tina. 2020. German case ambiguities at the interface: Production and comprehension. In Gerrit Kentner & Joost Kremers (eds.), *Prosody in syntactic encoding* (Linguistische Arbeiten 573), 51–84. Berlin: De Gruyter.

Bögel, Tina. 2021. Function words at the interface: A two-tier approach. *Languages* 6(4). DOI: 10.3390/languages6040197.

Bögel, Tina. Forthcoming. Prosody and its interfaces. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Bögel, Tina & Lea Raach. 2020. Swabian *ed* and *edda*: Negation at the interfaces. In Miriam Butt & Ida Toivonen (eds.), *Proceedings of the LFG '20 conference*, 47–67. Stanford: CSLI Publications.

Börjars, Kersti & John J. Lowe. Forthcoming. Noun phrases in LFG. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Bouma, Gerlof. Forthcoming. LFG and Continental West Germanic languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Bresnan, Joan. 1998. Morphology competes with syntax: Explaining typological variation in weak crossover effects. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky (eds.), *Is the best good enough? Optimality and competition in syntax*, 59–92. Cambridge, MA: The MIT Press & MITWPL.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-Functional Syntax*. 2nd edn. (Blackwell Textbooks in Linguistics 16). Malden, MA: Wiley-Blackwell.

Bresnan, Joan, Ronald M. Kaplan, Stanley Peters & Annie Zaenen. 1982. Cross-serial dependencies in Dutch. *Linguistic Inquiry* 13(4). 613–635. https://www.jstor.org/stable/4178298. Reprinted in Savitch, Bach, Marsh & Safran-Naveh (1987: 286-319).

Butt, Miriam, Christian Fortmann & Christian Rohrer. 1996. Syntactic analyses for parallel grammars: Auxiliaries and genitive NPs. In *COLING '96: Proceedings of the 16th Conference on Computational Linguistics*, 182–187. Copenhagen: International Committee on Computational Linguistics. DOI: 10.3115/992628.992662.

Choi, Hye-Won. 1995. Weak crossover in scrambling languages: Precedence, rank and discourse. Paper delivered at the 69th Annual Meeting of the Linguistic Society of America, New Orleans, LA. https://www.linguisticsociety.org/sites/default/files/1995_searchable.pdf.

Choi, Hye-Won. 1999. *Optimizing structure in context: Scrambling and information structure*. Stanford: CSLI Publications.

Choi, Hye-Won. 2001. Phrase structure, information structure and resolution of mismatches. In Peter Sells (ed.), *Formal and empirical issues in Optimality Theoretic Syntax*, 17–62. Stanford: CSLI Publications.

Clément, Lionel, Kim Gerdes & Sylvain Kahane. 2002. An LFG-type grammar for German based on the topological model. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '02 conference*. Stanford: CSLI Publications.

Cook, Philippa. 2001. *Coherence in German: An information structure approach*. (Revised version). University of Manchester. (Doctoral dissertation).

Cook, Philippa. 2006. The datives that aren't born equal: Beneficiaries and the dative passive. In Daniel Hole, André Meinunger & Werner Abraham (eds.), *Between argument structure and event structure* (Studies in Language Companion Series 75), 141–184. Amsterdam/Philadelphia: John Benjamins. DOI: 10.1075/slcs.75.08coo.

Cook, Philippa & John R. Payne. 2006. Information structure and scope in German. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '06 conference*, 124–144. Stanford: CSLI Publications.

Coussé, Evie & Gerlof Bouma. 2022. Semantic scope restrictions in complex verb constructions in Dutch. *Linguistics* 60(1). 123–176. DOI: 10.1515/ling-2021-0172.

Dalrymple, Mary. 2001. *Lexical Functional Grammar* (Syntax and Semantics 34). New York: Academic Press. DOI: 10.1163/9781849500104.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III & Annie Zaenen (eds.). 1995. *Formal issues in Lexical-Functional Grammar*. Stanford: CSLI Publications.

Dalrymple, Mary, Ronald R. Kaplan & Tracy Holloway King. 2001. Weak crossover and the absence of traces. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '01 conference*. Stanford: CSLI Publications.

Dalrymple, Mary, Tracy Holloway King & Louisa Sadler. 2009. Indeterminacy by underspecification. *Journal of Linguistics* 45. 31–68. DOI: 10.1017/s0022226708005513.

Diesing, Molly. 1997. Yiddish VP order and the typology of object movement in Germanic. *Natural Language & Linguistic Theory* 15(2). 369–427. DOI: 10.1023/A:1005778326537.

Dipper, Stefanie. 2003. Implementing and documenting large-scale grammars – German LFG. *Arbeitspapiere des Instituts für Maschinelle Sprachverarbeitung (AIMS)* 9(1).

Dipper, Stefanie. 2005. German quantifiers: Determiners or adjectives? In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '05 conference*, 100–115. Stanford: CSLI Publications.

Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67. 547–619. DOI: 10.2307/415037.

Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2019. *Ethnologue: Languages of the world*. 22nd edn. Dallas: SIL International.

Evers, Arnold. 1975. *The transformational cycle in Dutch and German*. Rijksuniversiteit Utrecht. (Doctoral dissertation).

Eythórsson, Thórhallur & Jóhanna Barðdal. 2005. Oblique subjects: A common Germanic inheritance. *Language* 81(4). 824–881.

Findlay, Jamie Y. & Anna Kibort. Forthcoming. Argument structure and mapping theory. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Forst, Martin. 2006. COMP in (parallel) grammar writing. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '06 conference*, 222–239. Stanford: CSLI Publications.

Forst, Martin, Tracy Holloway King & Tibor Laczkó. 2010. Particle verbs in computational LFGs: Issues from English, German, and Hungarian. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '10 conference*, 228–248. Stanford: CSLI Publications.

Forst, Martin & Christian Rohrer. 2009. Problems of German VP coordination. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '09 conference*. Stanford: CSLI Publications.

Fortmann, Christian. 2006. The complement of verba dicendi parentheticals. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '06 conference*. Stanford: CSLI Publications.

Frank, Anette. 2006. A (discourse-)functional analysis of asymmetric coordination. In Miriam Butt, Mary Dalrymple & Tracy Holloway King (eds.), *Intelligent linguistic architectures: Variations on themes by Ronald M. Kaplan*, 259–286. Stanford: CSLI Publications.

Haeseryn, Walter (ed.). 1997. *Algemene nederlandse spraakkunst*. 2nd edn. Groningen: Nijhoff.

Haider, Hubert. 2010. *The syntax of German* (Cambridge Syntax Guides). Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9780511845314.

Haider, Hubert & Inger Rosengren. 2003. Scrambling: Nontriggered chain formation in OV languages. *Journal of Germanic Linguistics* 15(3). 203–267. DOI: 10.1017/S1470542703000291.

Hammarström, Harald, Robert Forkel & Martin Haspelmath (eds.). 2019. *Glottolog 4.1*. Jena: Max Planck Institute for the Science of Human History. DOI: 10.5281/zenodo.3554959. https://glottolog.org.

Jones, Stephen. 2020. Multifunctional Dutch 'er'. In Miriam Butt & Ida Toivonen (eds.), *Proceedings of the LFG '20 conference*. Stanford: CSLI Publications.

Kaplan, Ronald M. & Annie Zaenen. 2003. West-Germanic verb clusters in LFG. In Pieter A. M. Seuren & Gerard Kempen (eds.), *Verb constructions in German and Dutch*, 127–150. Amsterdam: John Benjamins. DOI: 10.1075/cilt.242.05kap.

Kaplan, Ronald M. & Annie Zaenen. Forthcoming. Long-distance dependencies. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Kasper, Simon. 2014. Linking syntax and semantics of adnominal possession in the history of German. In Chiara Gianollo, Agnes Jäger & Doris Penka (eds.), *Language change at the syntax-semantics interface*, 57–100. Berlin, München, Boston: De Gruyter Mouton. DOI: 10.1515/9783110352306.57.

Kordoni, Valia. 2003. Valence alternations in German: An LMT analysis. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '03 conference*, 250–268. Stanford: CSLI Publications.

Kuhn, Jonas. 2001. Resource sensitivity in the syntax-semantics interface: Evidence from the German split NP construction. In W. Detmar Meurers & Tibor Kiss (eds.), *Constraint-based approaches to Germanic syntax* (Studies in Constraint-Based Lexicalism), 177–216. Stanford: CSLI Publications.

Kuhn, Jonas. Forthcoming. LFG, Optimality Theory and learnability of languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Kuhn, Jonas, Christian Rohrer & Sina Zarrieß. 2010. Right node raising in parsing and generation. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '10 conference*. Stanford: CSLI Publications.

Lee-Schoenfeld, Vera. 2007. *Beyond coherence: The syntax of opacity in German*. Vol. 114 (Linguistik Aktuell). Amsterdam: John Benjamins. DOI: 10.1075/la.114.

Maling, Joan & Annie Zaenen (eds.). 1990. *Modern Icelandic syntax* (Syntax and Semantics 24). San Diego, CA: Academic Press. DOI: 10.1163/9789004373235.

Meinunger, André. 2007. About object 'es' in the German *Vorfeld*. *Linguistic Inquiry* 38(3). 553–563. DOI: 10.1162/ling.2007.38.3.553.

Müller, Stefan, Felix Bildhauer & Philippa Cook. 2012. Beschränkungen für die scheinbar mehrfache Vorfeldbesetzung im Deutschen. In Colette Cortès (ed.), *Satzeröffnung. Formen, Funktionen, Strategien*, 113–128. Tübingen: Stauffenburg.

Netter, Klaus. 1988. Non-local dependencies and infinitival constructions in German. In Uwe Reyle & Christian Rohrer (eds.), *Natural language parsing and linguistic theories*, vol. 35 (Studies in linguistics and philosophy), 356–411. Dordrecht: Reidel. DOI: 10.1007/978-94-009-1337-0_13.

Poortvliet, Marjolein. 2015. An LFG approach to nested dependencies in Dutch. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '15 conference*. Stanford: CSLI Publications.

Pullum, Geoffrey K. & Gerald Gazdar. 1982. Natural languages and context-free languages. *Linguistics and Philosophy*. 471–504. DOI: 10.1007/bf00360802.

Putnam, Michael (ed.). 2011. *Studies on German-language islands* (Studies in Language Companion 123). Amsterdam: John Benjamins. DOI: 10.1075/slcs.123.

Rehbein, Ines & Josef van Genabith. 2006. German particle verbs and pleonastic prepositions. In *Proceedings of the third ACL-SIGSEM workshop on prepositions*. DOI: 10.3115/1621431.1621440.

Reis, Marga. 1996. Extraction from verb-second clauses in German? In Uli Lutz & Jürgen Pafel (eds.), *On extraction and extraposition in German*, 45–88. Amsterdam: John Benjamins. DOI: 10.1075/la.11.03rei.

Rohrer, Christian. 1996. Fakultativ kohärente Infinitivkonstruktionen im Deutschen und deren Behandlung in der Lexikalisch Funktionalen Grammatik. In Gisela Harra & Manfred Bierwisch (eds.), *Wenn die Semantik arbeitet*, 89–108. Mouton de Gruyter. DOI: 10.1515/9783110919899.89.

Santorini, Beatrice. 1992. Variation and change in Yiddish subordinate clause word order. *Natural Language & Linguistic Theory* 10(4). 595–640. DOI: 10.1007/bf00133331.

Savitch, Walter J., Emmon Bach, William Marsh & Gila Safran-Naveh (eds.). 1987. *The formal complexity of natural language* (Studies in Linguistics and Philosophy). Dordrecht: Springer.

Seiler, Guido. 2007. Microvariation in LFG and OT. In Annie Zaenen, Jane Simpson, Tracy Holloway King, Jane Grimshaw, Joan Maling & Chris Manning (eds.), *Architectures, rules, and preferences: Variations on themes by Joan W. Bresnan*, 529–547. Stanford: CSLI Publications.

Sells, Peter, Annie Zaenen & Draga Zec. 1987. Reflexivization variation: Relations between syntax, semantics, and lexical structure. In Masayo Iida, Stephen Wechsler & Draga Zec (eds.), *Working papers in grammatical theory and discourse structure, volume 1: Interactions of morphology, syntax, and discourse* (CSLI Lecture Notes 11), 169–238. Stanford: CSLI Publications.

Shieber, Stuart M. 1985. Evidence against the context-freeness of natural language. *Linguistics and Philosophy* 8(3). 333–343. DOI: 10.1007/978-94-009-3401-6_12.

Strunk, Jan. 2004. *Possessive constructions in Modern Low Saxon*. Stanford: Stanford University. (MA thesis).

Strunk, Jan. 2005. Pro-drop in nominal possessive constructions. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '05 conference*, 447–467. Stanford: CSLI Publications.

Theiler, Nadine & Gerlof Bouma. 2012. Two for the price of one: An LFG treatment of sentence initial object es in German. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '12 conference*. Stanford: CSLI Publications.

van der Beek, Leonoor. 2005. *Topics in corpus-based Dutch syntax*. University of Groningen. (Doctoral dissertation). https://research.rug.nl/nl/publications/topics-in-corpus-based-dutch-syntax.

Wescoat, Michael T. 2007. Preposition-determiner contractions: An analysis in Optimality-Theoretic Lexical-Functional Grammar with lexical sharing. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '07 conference*, 439–459. Stanford: CSLI Publications.

Wurmbrand, Susi. 2004. West Germanic verb clusters: The empirical domain. In Katalin É. Kiss & Henk van Riemsdijk (eds.), *Verb clusters: A study of Hungarian, German, and Dutch* (Linguistics Today 69), 43–85. Amsterdam: John Benjamins. DOI: 10.1075/la.69.05wur.

Zaenen, Annie. 1989. *The place of bevallen (please) in the syntax of Dutch*. Tech. rep. SSL-89-17. System Sciences Laboratory, Palo Alto Research Center. https://web.stanford.edu/~azaenen/publications/Bevallen.pdf.

Zaenen, Annie. 1993. Unaccusativity in Dutch: Integrating syntax and lexical semantics. In James Pustejovsky (ed.), *Semantics and the lexicon*, 129–161. Dordrecht: Kluwer Academic Publishers. DOI: 10.1007/978-94-011-1972-6_9.

Zaenen, Annie. 1996. Contrastive dislocation in Dutch and Icelandic. In Elena Anagnostpoulou, Henk van Riemsdijk & Frans Zwarts (eds.), *Materials on left dislocation*. Amsterdam: John Benjamins. DOI: 10.1075/la.14.09zae.

Zaenen, Annie. Forthcoming. Information structure. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Zaenen, Annie & Ronald M. Kaplan. 1995. Formal devices for linguistic generalizations: West Germanic word order in LFG. In Jennifer S. Cole, Georgia M. Green & Jerry L. Morgan (eds.), *Linguistics and computation*, 3–27. Stanford: CSLI Publications. Reprinted in Dalrymple, Kaplan, Maxwell & Zaenen (1995: 215–240).

Zaenen, Annie & Ronald M. Kaplan. 2002. Subsumption and equality: German partial fronting in LFG. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '02 conference*, 409–426. Stanford: CSLI Publications.

Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language & Linguistic Theory* 3(4). 441–483. DOI: 10.1007/bf00133285. Reprinted in Maling & Zaenen (1990: 95–136).

Zifonun, Gisela, Ludger Hoffmann & Bruno Strecker. 1997. *Grammatik der deutschen Sprache*. Vol. 7 (Schriften des Instituts für deutsche Sprache). Berlin: Walter de Gruyter. DOI: 10.1515/9783110872163.

Zwart, Jan-Wouter. 2008. Continental West-Germanic languages. In Guglielmo Cinque & Richard S. Kayne (eds.), *The Oxford handbook of comparative syntax*. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780195136517.013.0021.

Chapter 5

LFG and Sinitic languages

Olivia S.-C. Lam
University of Hong Kong

One-Soon Her
Tunghai University & National Chengchi University

Jing Chen
The Hong Kong Polytechnic University

Sophia Y.-M. Lee
The Hong Kong Polytechnic University

The assumptions of LFG have been applied to the research on a number of grammatical phenomena in Chinese languages. In this chapter, we present an overview of some of the studies devoted to investigating the syntactic patterns of two varieties of Chinese: Mandarin and Cantonese. This chapter includes a discussion on the expression and identification of grammatical functions, *ba*, *bei* and related constructions, the dative alternation, compounds (VO compounds and resultative compounds), the locative inversion, and classifiers and measure words. The chapter concludes with a brief overview of the applications of LFG in Chinese language processing.

1 Introduction: Chinese or Sinitic Languages

LFG is a lexicon-driven, unification-based linguistic theory aiming to account for both variations and universals found in human languages. The well-known parsimony of morpho-syntactic markings in Chinese poses an interesting challenge to the theory, but at the same time provides an opportunity to showcase

Olivia S.-C. Lam, One-Soon Her, Jing Chen & Sophia Y.-M. Lee. Forthcoming. LFG and Sinitic languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

the explanatory adequacy of LFG. The term 'Chinese' is commonly replaced by 'Sinitic languages' or 'Chinese languages' in the linguistics literature. These two terms refer to a family of varieties which are genetically related but are, very often, not mutually intelligible (Handel 2015; Huang & Shi 2016). Wurm & Liu (1987) list 10 varieties under 'Chinese' in the *Language Atlas of China*, while the *Ethnologue* lists 16 (Eberhard et al. 2020). The more prominent varieties are traditionally known as *fangyan* 方言 (literally 'regional speech' or 'dialect'), and are classified into 7 groups: Mandarin, Xiang, Gan, Wu, Yue, Hakka and Min. Drawing data from both Mandarin and Cantonese (a Yue dialect), we will be using the term 'Chinese' to loosely refer to the Sinitic family, and reserve the terms 'Mandarin' and 'Cantonese' for these two individual varieties.

LFG has been adopted to study Chinese since 1985. Earlier studies, such as Huang (1985, 1986, 1987, 1988, 1989b,a, 1990) and Huang & Mangione (1985), present LFG accounts of a wide range of grammatical structures in Mandarin Chinese, including the internal structure of NPs, the subcategorized topic, and lexical discontinuity. Her (1990) investigates the grammatical functions in Mandarin, while Tan (1991) focuses on the subject in Mandarin. Bodomo & Luke (2003), the monograph resulting from the first LFG Workshop dedicated to the analysis of Chinese languages in 2001, contains studies on Mandarin, Cantonese, and other Sinitic languages.

It is important to note that, although this chapter focuses on Cantonese and Mandarin, LFG has in fact been successfully applied to a wide range of varieties in China. For instance, Huang (1991) provides an account of adjectival reduplication in Taiwan Southern Min. Studies on Zhuang, a Tai-Kadai language spoken in southern China, include Pan (2010), Bodomo (2011), and Burusphat & Qin (2012).

There is also a well-established collection of LFG literature written in Chinese, with most of them providing an introduction to the framework. These include Huang (1988, 1989b), Fu (1990a,b), Fu (1993), Feng (2004), Gao & Li (2009) and Wei (2014).

In the following sections, we first outline the prominent grammatical properties of Chinese from an LFG perspective (Section 2). Section 3 discusses the encoding of grammatical functions in Chinese, while Sections 4–10 provide an overview of the major grammatical phenomena which have been analyzed in LFG. Section 11 concludes the chapter by highlighting LFG analyses which have contributed to the understanding of Sinitic languages, and how the studies on Sinitic languages have contributed to the development of LFG.

2 Grammatical properties: An LFG Perspective

This section introduces important grammatical features of Chinese from an LFG perspective, including the morpho-syntactic encoding of grammatical functions (Section 2.1); the classifier system (Section 2.2); and the canonical word order and the role of information structure (Section 2.3). For more in-depth and recent discussions on issues in Chinese linguistics, see Huang et al. (2009), Wang & Sun (2015), Huang & Shi (2016), and Huang et al. (2022), among others.

2.1 Morpho-syntactic encoding

Chinese has been described in the literature as being 'morphologically impoverished' (e.g. Packard 2000; Hsieh et al. 2022). This, however, does not mean that there is no morpho-syntactic encoding. In (1a), TENSE is not encoded on the verb, but in (2), aspect is.[1,2]

(1) a. Cantonese
Zoengsaam kam jat/ gam jat/ ting jat faangung.
Zoengsaam yesterday/ today/ tomorrow work
'Zoengsaam went to work yesterday/ is going to work today/ will go to work tomorrow.'

b. Mandarin
Zhangsan zuotian/ jintian/ mingtian shangban.
Zhangsan yesterday/ today/ tomorrow work.
'Zhangsan went to work yesterday/ goes to work today/will go to work tomorrow.'

(2) a. Cantonese
Zoengsaam tai-zo/ -gan/ -gwo bun syu.
Zoengsaam read-PFV/ -PROG/ -EXP CLF book
'Zoengsaam has read/is reading/read the book.'

b. Mandarin
Zhangsan du-le/ zhengzai du/ du-guo (yi) ben shu.
Zhangsan du-PRF/ ZAI read/ read-EXP (NUM) CLF book.
'Zhangsan has read/is reading/read a book.'[3]

[1] Tones are omitted unless they are relevant to the discussion.
[2] Examples in Cantonese are romanized using the scheme developed by LSHK (2002).
[3] The marker *-gwo*, and the Mandarin equivalent *-guo*, express the 'experiential aspect' in Chinese.

There is no person, number or gender agreement between a verb and its arguments.

(3) Cantonese

a. Zoengsaam gin-dou keoidei.
   Zoengsaam see-DOU they
   'Zoengsaam saw them.'

b. ngo gin-dou Zoengsaam.
   I see-DOU Zoengsaam
   'I saw Zoengsaam.'

Note that the changes in person and number do not affect the verb forms in (3). Note also that *-dou* (到; *-dao* in Mandarin) is not a tense marker – it marks accomplishment and is part of a verb-result compound.

There is no case-marking in Chinese. Pronouns are not case-marked, either:

(4) Cantonese

a. ngo gin-dou keoi.
   I see-DOU s/he
   'I saw him/her.'

b. keoi gin-dou ngo.
   s/he see-DOU I
   'S/he saw me.'

2.2 Number-marking, classifiers and the expression of quantities

Most nouns are not number-marked. The only marker which codes number in Mandarin is the plural marker *-men* (Hsieh et al. 2022). Yet, even for human nouns, a bare noun is unspecified for number, allowing both a singular and a plural reading, as exemplified in (5).

(5) Mandarin
   Gebi de xuesheng hen chao.
   next.door DE student very noisy
   'The student(s) next door is/are very noisy.'

Classifiers are a significant feature of the Chinese languages. As number is not explicitly encoded in Chinese, nouns can only be enumerated when they are individuated by classifiers in the [NUM CLF N] structure. Some scholars believe that classifiers 'serve to profile an essential or inherent feature of the head noun...and contribute no additional meaning to the head noun' (Her 2012a; see also Cheng & Sybesma 1999). Others (e.g. Huang & Ahrens 2003; Chen et al. 2022), however, argue that classifiers make a crucial contribution to the meaning through coercion.

(6) Mandarin

a. san ben shu
three CLF book
'three (volumes/copies of) books'

b. san xiang shu
three CLF book
'3 boxes of books'

Cantonese, among other varieties of Chinese and unlike Mandarin, allows the omission of the numeral one. Whether 'one' is expressed depends on the information structure and the grammatical function of the noun. The structure [CLF N] receives a definite, or contextually retrievable, interpretation when it serves as the SUBJ, but when it is an OBJ, either a definite or an indefinite reading is possible:

(7) Cantonese

a. [CLF N] as SUBJ
(Context: What happened to the book?)
[bun syu] laan-zo.
CLF book damage-PFV
'The book is damaged.'

b. [CLF N] as OBJ

i. With a definite reading
(Context: Where is the book?)
ngo m gin-zo [bun syu].
I not see-PFV CLF book
'I have lost the book.'

ii. With an indefinite reading
ngo kam jat maai-zo [bun syu].
I yesterday buy-PFV CLF book
'I bought a book yesterday.'

2.3 Canonical word order

Different views can be found in the literature regarding the canonical word order in Chinese languages. While there is a long tradition of analyzing Chinese as having a canonical SVO word order (e.g. Light 1979; Mei 1980; Sun & Givón 1985; Dryer 2005), there are also arguments for treating the SOV order as the canonical word order (see, for instance, Tai 1973; Li & Thompson 1974). The empirical and theoretical arguments for both the SVO and SOV accounts can be found in Liu (2022) and Xu & Dong (2022) respectively. In some Wu varieties, it has also been observed that the SOV or OSV orders occur more frequently than the SVO order, especially in cases where OBJ expresses the patient role (Yue 2003).

Despite the ongoing debate on the canonical word order, it is generally accepted that word order variations in Chinese can be accounted for in terms of information structure (Shyu 2016). Chinese has been well-established as a topic-prominent language since Li & Thompson (1976). Constituents bearing almost any grammatical function can be easily placed in the sentence-initial position as long as they are topics. Kroeger (2004) provides a clear overview on the grammatical functions which can be topicalized in Chinese, including the possessor (Xu & Langendoen 1985). Identifying grammatical functions in Chinese is thus far from being straight-forward – grammatical functions may be expressed in various syntactic positions depending on the discourse context, and they are not morphologically encoded. The OBJ *pingguo* can appear in the canonical object position (8a), sentence-initially if it is topical (8b), and between the SUBJ and the V, where the marker *ba* is optional:[4]

(8) Mandarin

a. ta chi le [pingguo].
s/he eat PFV apple
'S/he ate the apple/apples.'

b. [pingguo] ta chi le.
apple s/he eat PFV
'S/he ate the apple/apples.'

[4]Whether the marker *ba* is required depends on the semantic features of the displaced NP. A displaced human NP must be marked:

(i) Ta *(ba) laoshi tuidao le.
3SG.M BA teacher push.over PFV
'He pushed over the teacher.' (Yang & van Bergen 2007: 1622)

c. ta (ba) [pingguo] chi le.
s/he BA apple eat PFV
'He ate the apple/apples.' (Yang & van Bergen 2007: 1622)

Other word order variations are found in Chinese. These will be discussed in Section 4.

Chinese is also well-known for having 'Chinese-style topics' (Chafe 1976), or 'dangling topics'. These topics are unique in that they are not subcategorized for by the predicate in the comment (Pan & Hu 2008). In (9), the predicate in the comment is *lai* 'come', which is intransitive and only subcategorizes for a subject, *xiaofangdui* 'fire-brigade'. The topic [*nei chang huo*] 'that fire' is not related to the predicate-argument structure of *lai* 'come', and is thus considered a 'dangling' topic.

(9) Mandarin
[nei chang huo], xingkui xiaofangdui lai de kuai.
that CLF fire fortunately fire-brigade come DE quick
'As for that fire, fortunately the fire-brigade came quickly.'
(Li & Thompson 1976)

It is also possible and entirely natural to have more than one topic at the beginning of a sentence in Chinese, i.e. 'topic-chain constructions':

(10) Mandarin
[zhei jian shi], (Zhangsan), ta mei you cuo.
this CLF matter Zhangsan 3SG not have fault
'Regarding this matter, Zhangsan is not at fault.'
(Her 1990; glosses modified)

We provide a more detailed discussion on the TOPIC as a grammatical function in Section 3.

3 Grammatical functions and word order variations in Chinese

We provide a synopsis of the state-of-the-art LFG research on Chinese in this section and Sections 4–10. We begin with the fundamental issue of encoding grammatical functions in Chinese.

Identifying grammatical functions in Chinese can be challenging due to the lack of morphological encoding of grammatical functions, and to the fact that Chinese has relatively free word order. We offer an overview of the grammatical functions in Mandarin (Section 3.1), and in Cantonese (Section 3.2).

3.1 Mandarin

Almost all early LFG studies on Chinese have included a classification of grammatical functions. Interestingly, although there are no obligatory morphological encodings of GFs, there is general consensus as to the grammatical functions which can be identified for Chinese. Huang (1989b, 1993a), adopting the assumptions of classical LFG (Bresnan 1982b; Bresnan & Kanerva 1989), shows that GFs in Mandarin can be identified by their unambiguous syntactic positions at the surface level, and can be classified into four types based on two features: [±restricted] and [±objective]. Her (1990, 2008) presents an expanded set of GFs in Mandarin, and recognizes SUBJ, OBJ, OBJ2, OBL_θ (oblique function which includes subtypes $\text{OBL}_{\text{THEME}}$ (theme), OBL_{GOAL} (goal), OBL_{BEN} (beneficiary), OBL_{LOC} (location)), and COMP (complement function that includes subtypes XCOMP, SCOMP, and NCOMP) as subcategorizable GFs, while TOPIC, ADJunct (adjunct function that has two subtypes ADJ and XADJ), and POSS are identified as non-subcategorizable, as shown in Figure 1. It should be noted that, in the current LFG literature, the restricted object function OBJ_θ has replaced OBJ2, while grammatical function labels such as SCOMP and NCOMP, which make reference to c-structure categories, are no longer adopted.

The syntactic encoding of GFs is via both the c-structure and the predicate argument structure (AS). Take the lexical verb *da* ‘hit’, for example: it has a predicate argument structure of ⟨AGENT, THEME⟩, and subcategorizes for ⟨SUBJ, OBJ⟩, where the linking between the argument roles and the grammatical functions is constrained by the Lexical Mapping Theory (LMT; Bresnan & Kanerva 1989).

(11) Mandarin
Lisi da Zhangsan.
Lee hit John
‘Lee hit John.’

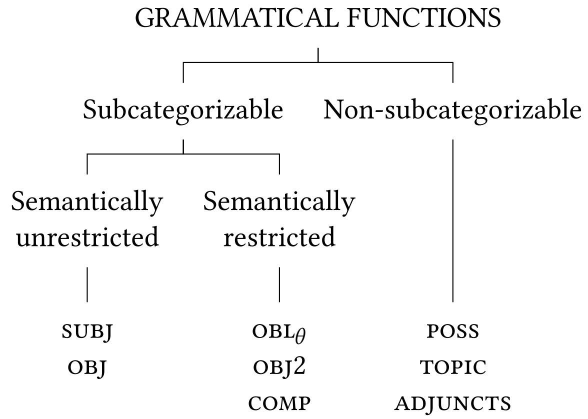
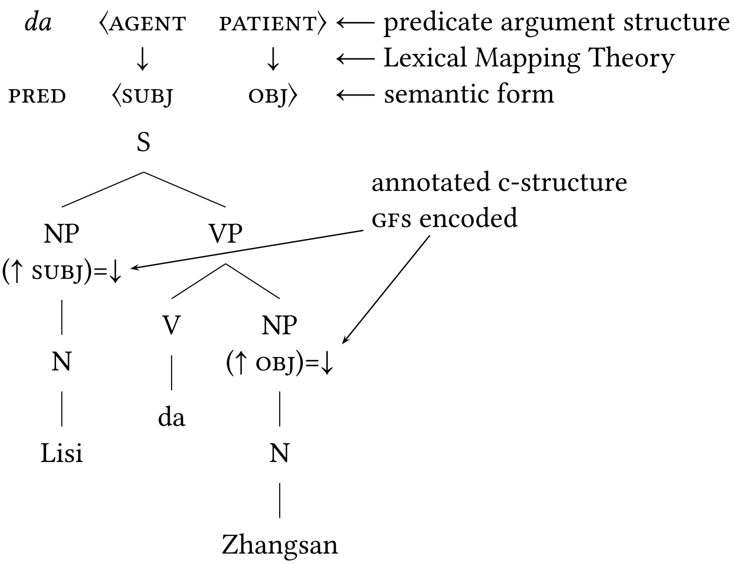


Figure 1: Classification of grammatical functions in Mandarin (Her 1990, 2008)



The treatment of TOPIC above touches on a fundamental issue related to the universal properties of GFs. Recall that Chinese is a topic-prominent language (see, for instance, Tsai (2022), for a discussion on the syntactic approaches to the phenomenon, and Tao (2022), among others, for a discussion on the functional approaches). Thematic TOPICs may be 'preposed', while non-thematic TOPICs may remain *in situ*. A set of frequently used constructions known as 'Pseudo-transitive constructions' (Chang et al. 1988) pose challenges to the grammatical status of TOPIC, and this has been treated in detail in Huang (1989a). In these constructions, an NP which is clearly an argument of the verb may only occur in the pre-verbal TOPIC position or some OBL positions, but never in the postverbal OBJ position. The following two examples are from Huang (1989a).

(12) Mandarin
  a. zeijian shi, ni zuozhu.
     this matter you make.master
     'You'll take charge of this matter.'
  b. *ni zuozhu zeijian shi.
     you make.master this matter

(13) Mandarin
  a. yuyanxue, ta nashou.
     linguistics s/he take.hand
     'S/he is good at linguistics.'
  b. *ta nashou yuyanxue.
     s/he take.hand linguistics

Huang (1989a) has made the following observations: (i) the topical NPs have clearly subcategorizable semantic roles; (ii) these constructions involve a large set of compound verbs, including some VN compounds which are practically all disyllabic in Chinese, and all of the quadrisyllabic compounds, and (iii) TOPICS can be regarded as being subcategorized (Bresnan 1982a). Based on these three observations, Huang shows that the most efficient account is to treat the topical NPs as subcategorized TOPICS. Mo (1990) has proposed a new grammatical function STOPIC ("s" for 'subcategorized') to differentiate them from the non-thematic TOPICS.

According to Huang (1989a), the subcategorizable TOPIC achieves parsimony in terms of lexical encoding and mapping to c-structure, but this would introduce complexities to the LMT. Her (1991, 2010), based on the same LMT considerations, argues that TOPICS should be regarded as strictly non-subcategorizable. To deal with the fact that pseudo-transitive verbs do not allow the stipulated OBJS to be realized in the canonical OBJ position, a feature-value pair [FRAME +] is assigned to those verbs. The [FRAME +] feature can only be obtained by way of unification with the TOPIC. The annotated PSR in (14b) specifies that TOPIC receives the feature [FRAME +] and it must be associated with some GF in the f-structure to fulfill the Extended Coherence Condition.

(14) a. *nashou* V (↑ PRED) = 'BE-GOOD-AT⟨SUBJ OBJ⟩' (Her 2010)
(↑ OBJ FRAME) $=_c$ +

b. S′ ⟶ NP S
(↑ TOPIC)=↓ ↑=↓
(↑ FRAME)=+

It is important to note that neither account explicates how it will account for the NPs occurring in other non-OBJ positions, such as in (15).

(15) Mandarin
[[Mali zui nashou] de kemu] shi shuxue.
Mary most take.hand DE subject be math
'The subject that Mary is best at is math.'

In (15), a gap in the relative clause is linked to the head noun, and is then linked to the complement of the verb *shi* 'be'. In Huang's (1989a) account, the subcategorized SUBJ will have to be linked to other GFs following the same mechanisms for control and complementation (Bresnan 1982a). See Her (2010) for a different account. In both cases, however, there is neither a clear solution to the entailed complexities for LMT, nor an answer to the question of why such a high-level solution is needed for what seems to be a parochial fact limited to a set of predicates in a specific language.

In sum, the pseudo-transitive verbs in Mandarin, where the OBJ-like arguments can only occur in the TOPIC position, pose a great challenge to the theory of GF-encoding in LFG. The two existing proposals (Huang 1989a and Her 2010) both have their strengths and weaknesses. The fact that the set of verbs involved are some of the verbs currently undergoing changes in transitivity (Jiang & Huang 2018) suggests that the ultimate solution may involve a theory which takes historical changes involving GFs into consideration.

3.2 Cantonese

In contrast to the issue-drive discussion on GFs in Mandarin in the last section, this section will provide a survey on SUBJ (Lee 2003), OBJ (Lam 2008), and the complement (Bodomo & Lee 2003; Lee 2002) in Cantonese.

3.2.1 Subject in Cantonese

Lee (2003) shows that two syntactic properties are of particular relevance in the identification of the subject in Cantonese. The first is the binding of the reflexive pronoun *zigei* 'self' to the subject *Mary* within the same clause, or to the subject *John* in the clause containing the local clause. This, following Tan (1991), clearly distinguishes the subject from the topic, both of which can be found preverbally.

(16) Cantonese

John_i	zi1	Mary_j	sik6-zo2	$\text{keoi5-zi6gei2}_{i/j}$	haap6	faan6.
John	know	Mary	eat-PFV	3.SG-self	CLF	rice

'John knows that Mary ate his/her lunch box.' (Lee 2003: 30)

The second distinctive property of the subject is that the possessor of the subject can be easily relativized with the gap strategy (17a), but the possessor of the object cannot be relativized in the same way (17b):

(17) Cantonese

a.

[_ sing4zik1]	ji5ging1	gung1bou3-zo2	ge3	hok6saang1
grades	already	announce-PFV	REL	students

'the students whose grades have been announced.' (Lee 2003: 37)

b.

*hok6haau6	ji5ging1	gung1bou3-zo2	[_ sing4zik1]	ge3	hok6saang1
school	already	announce-PFV	grades	REL	students

(Intended meaning: 'the student whose grades have been announced by the school') (Lee 2003: 38)

Luke et al. (2001) discuss the Subject Condition in Cantonese. As with Sinitic languages in general, Cantonese allows pro-drop even without agreement morphology or case-marking. This poses a challenge to the identification of grammatical functions at f-structure. Luke et al. (2001) show that apparently 'subjectless' sentences, in fact, do have a subject, but discourse-pragmatic criteria, such as the speech context, must be taken into consideration in order to retrieve the subject. See also Liao (2010) for a discussion on the pro-drop patterns in Mandarin Chinese, and for an analysis within LFG.

3.2.2 Object in Cantonese

Lam (2008) investigates the syntax of objects in Cantonese, in particular, their syntactic behaviours in double object constructions (DOCs). Without morphological marking, the structural position of each object becomes an important clue in the identification of the different object functions – in (18), the recipient-object is found immediately postverbally, with the theme-object following it:

(18) Cantonese

a. Recipient-NP < Theme-NP
ngo gaau siupangjau zungman.
1sg teach children Chinese
'I teach children Chinese.'

b. *Theme-NP < Recipient-NP
*ngo gaau zungman siupangjau.
1sg teach Chinese children
'I teach children Chinese.'

This, however, is not the canonical order of objects for the verb GIVE – it is the theme-object that must be immediately postverbal.

(19) Cantonese

a. Theme-NP < Recipient-NP
ngo bei-zo bun syu ngo gaaze.
1sg give-pfv clf book 1sg elder.sister
'I gave the book to my elder sister.'

b. *Recipient-NP < Theme-NP
*ngo bei-zo ngo5 gaaze bun syu.
1sg give-pfv 1sg elder.sister clf book

A related question is – which one of these objects is the unrestricted object obj, and which is the restricted one obj_θ? In LFG, the object in a DOC which grammatically patterns with the monotransitive object is obj, while the one which does not is obj_θ. Passivization is often seen as *the* diagnostic for unrestricted objecthood, but in Cantonese, as in Mandarin Chinese, passivization is often constrained - the passive is associated with a meaning of adversity. As a result, not all verbs, even monotransitive ones, can be involved in passivization (20). It is therefore not a very helpful test for the unrestricted object. We shall return to a discussion of passivization in Section 4.

(20) Mandarin

a. Zhangsan gei ren du-si le.
Zhangsan give people poison-die part
'Zhangsan was poisoned to death by people.'

b. *Zhangsan gei ren yi-hao le.
Zhangsan give people cure part
'Zhangsan was cured by people.' (Lefebvre 2011: 257)

Patterns of relativization and pro-drop show that it is the theme-object which behaves like the monotransitive object. Lam (2008) thus concludes that the theme-object is the unrestricted object in Cantonese, while the recipient-object is the restricted object.

3.2.3 Complement in Cantonese

Lee (2002) and Bodomo & Lee (2003) show that Cantonese verbs such as *zidou* 'think' may take either a COMP (21a) or an OBJ (21b), while other verbs subcategorize for only a COMP (22a) but not an OBJ (22b):

(21) Cantonese

a. ngo zi dou $_S$[keoi hai hou jan].
   1SG know 3SG be good person
   'I know that s/he is a good person.'

b. ngo zi dou $_{DP}$[li gin si].
   1SG know this CLF matter
   'I know (about) this.'

(22) Cantonese

a. ngo hei mong $_S$[keoi hai hou jan].
   1SG hope 3SG be good person
   'I hope that s/he is a good person.'

b. *ngo hei mong $_{DP}$[keoi].
   1SG hope 3SG
   ['I hope him/her.']

They therefore argue that Cantonese is a 'mixed language', along the lines of Dalrymple & Lødrup (2000).

4 *Ba*, *Bei*, and Related Constructions

4.1 Mandarin

The Mandarin *bei* construction is considered to be the equivalent of the English *by* passive in the literature. The discussion of the *bei* passive is frequently compared to the *ba* construction, as they share almost identical surface structures. Note that in (23), the agent *gemi* 'fans' is optional, much like the *by*-agent phrase in English. A *bei* construction with the agent phrase is known as the 'long' passive, while a *bei* construction with the agent phrase omitted is the 'short' passive (Huang et al. 2009; Huang & Shi 2016).

(23) Mandarin
Amei bei (gemi) weizhu.
Amei BEI fans encircle
'Amei was encircled (by the fans).'

(24) Mandarin
gemi ba *(Amei) weizhu.
fans BA Amei encircle
'The fans encircled Amei.'

Several important and controversial issues have been raised over the passive analysis of the *bei* construction. The first is whether *bei* is a preposition like the English *by* (Huang 1982; Li 1990; Li & Thompson 1981; Lü 1980; McCawley 1992; Tsao 1996) or a verb (Bender 2000; Feng 1995; Her 1989, 2009; Hsueh 1989; Huang 1999). The current dominant view of *bei* as higher verb is heralded by Huang & Mangione's (1985) formal semantic account, and was first adopted in LFG syntactic studies (e.g. Huang & Mangione 1985; Bender 2000).

The second issue is whether there is one or two passive constructions. The dominant GB analysis treats the passive in Mandarin as having 'split' into two different constructions: the agentless short passive versus the long passive with an overt agent. This is motivated by the observation that the long passive allows a much wider range of syntactic behaviours than the short passive. Yet Her (2009) shows, with corpus data from Sinica Corpus (Chen et al. 1996), that the short passive in fact exhibits the same range of syntactic behaviours, and argues that the two should receive exactly the same analysis, with the only difference being whether the agent is overt or covert. The evidence is presented below. First, Her (2009) shows that short passives (26), just like long passives (25), allows long-distance gaps:

(25) Mandarin
bei ta qitu nuyi de ziyou renmin.
BEI 3SG attempt enslave DE free people
'The free people who were "attempted-to-enslave" by him.'

(26) Mandarin

a. gongsi-de wanglu bei qitu ruqin.
company-POSS network BEI attempt hack
'The company network has been "attempted-to-hack".'

b. ziliao bei shefa kaobei le.
document BEI manage copy PRF
'The documents have been "managed-to-copy".'

Second, the claims in the literature that a long passive, but not a short passive, allows a resumptive pronoun to fill a gap are also incorrect, as in (27) and (28).

(27) Mandarin
Zhangsan$_i$ bei wo piping-le ta$_i$ yidun.
John BEI 1SG criticize-PFV 3SG once
'John was criticized once by me.'

(28) Mandarin
ta ba$_i$ pa bei renwei ta$_i$ wufa guanjiao haizi.
3SG father afraid BEI consider 3SG fail discipline children
'His father was afraid to be considered that he failed to discipline his children.'

Third, the split view claims that the pronominal particle *suo* is allowed in the long passive only, as in (29), and not the short passive. The corpus example in (30) shows that *suo* can be found in the short passive as well:

(29) Mandarin
ni hui bei ren suo chixiao.
2SG will BEI person SUO sneer
'I'm afraid your recent behavior toward him will be sneered at.'

(30) Mandarin
ni nanmian bu bei suo pian.
2SG unavoidably not BEI SUO trick
'Unavoidably you would be tricked.'

Finally, the split view claims that only the long passive allows an adverbial PP, as in (31), but not in the short passive. This is again shown to be wrong by the corpus example in (32).

(31) Mandarin
Zhangsan bei Lisi zai xuexiao pian-zou le.
John BEI Lee at school abduct PFV
'John was abducted at school by Lee.'

(32) Mandarin
yi zhi laoshu bei zai jiujing zhong jinpao-le yi nian.
one CLF mouse BEI at alcohol inside soak-PFV one year
'A mouse has been soaked in alcohol for a year.'

The analysis proposed by Her (2009) has *bei* as a three-place predicate requiring three theta roles, which are mapped to SUBJ, OBJ, and XCOMP, with a meaning that approximates (33). The lexical entry, including its lexical category, lexical form, and the control relations, is shown in (34). Note that the operation that links theta roles with GFs is γ; thus $\gamma(\widehat{\Theta})$ in (34) refers to the GF linked to the logical subject. ($\uparrow$ OBJ) = ($\uparrow$ XCOMP $\gamma(\widehat{\Theta})$) thus means that OBJ controls the GF in XCOMP that is linked to the $\widehat{\Theta}$. The f-structure of a typical *bei* sentence is illustrated in (35).

(33) *bei* ⟨*x y z*⟩: *x* is (adversely) affected by *y* in a way that *z* describes

(34) *bei* V ($\uparrow$ PRED) = 'BEI⟨SUBJ OBJ XCOMP⟩'
($\uparrow$ SUBJ) = ($\uparrow$ XCOMP TOPIC)
($\uparrow$ OBJ) = ($\uparrow$ XCOMP $\gamma(\widehat{\Theta})$)
$\neg$($\uparrow$ OBJ) $\Rightarrow$ ($\uparrow$ OBJ PRED) = 'PRO'

(35) Mandarin
na jian fangzi bei (Lisi) chai-le.
that CLF house BEI Lee demolish-PFV
'That house got demolished (by Lee).'

$$\left[\begin{array}{ll} \text{PRED} & \text{'BEI⟨SUBJ OBJ XCOMP⟩'} \\ \text{SUBJ} & \left[\text{PRED 'HOUSE'}\right] \\ \text{OBJ} & \left[\text{PRED 'LEE'}\right] / \left[\text{PRED 'PRO'}\right] \\ \text{XCOMP} & \left[\begin{array}{ll} \text{TOPIC} & \\ \text{PRED} & \text{'DEMOLISHED⟨SUBJ⟩'} \\ \text{SUBJ} & \end{array}\right] \end{array}\right]$$

In (35), ($\uparrow$ OBJ), which is either an overt agent *Lee* or a covert pronoun, is responsible for adversely affecting ($\uparrow$ SUBJ), the house, in a way described by

(↑ XCOMP), i.e., the house is demolished. Note that (↑ SUBJ) controls the TOPIC in XCOMP, which is anaphorically linked to SUBJ, indicated by the dotted line. The matrix SUBJ, the house, is also the SUBJ of the embedded clause, which is passive in nature. A non-canonical example is given in (36), with both c-structure and f-structure illustrated.

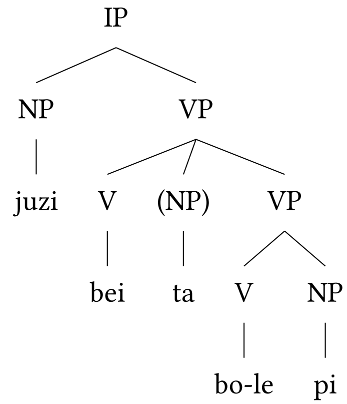
(36) Mandarin

juzi bei (ta) bo-le pi.

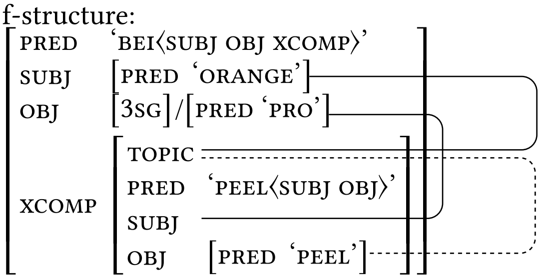
orange BEI 3SG peel-PFV peel

'The orange has its peel peeled off (by him).'

a. c-structure:



b. f-structure:



In (36), (↑ OBJ), which is again either overt or covert, is responsible for adversely affecting (↑ SUBJ), the orange, in a way described by (↑ XCOMP), i.e., the orange has its peel peeled off. Note that (↑ SUBJ) controls the TOPIC in XCOMP, and (↑ OBJ) controls the SUBJ in XCOMP. Within the XCOMP, TOPIC is anaphorically linked to OBJ.

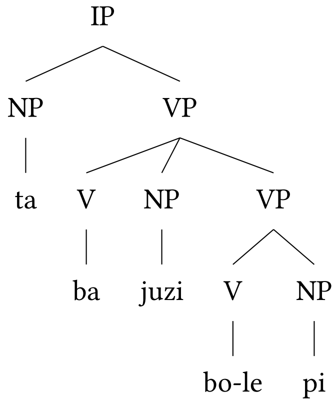
Based on this account, Her (2009) contends that the *bei* construction is the passive counterpart of the *ba* construction, not the canonical active sentence. Thus, *ba* is likewise a three-place predicate, as in (37), and its lexical entry is shown in (38). The example in (39) is therefore the active counterpart of the passive (36). See also Bender (2000) for an LFG analysis of the *ba* construction in Mandarin.

(37) *ba* ⟨*x y z*⟩: *x* affected *y* in a way that *z* describes

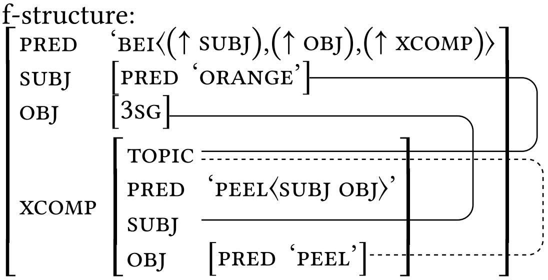
(38) *ba* V (↑ PRED) = ‘BA⟨SUBJ OBJ XCOMP⟩’
(↑ OBJ) = (↑ XCOMP TOPIC)
(↑ SUBJ) = (↑ XCOMP $\gamma(\widehat{\Theta})$)

(39) Mandarin
ta ba juzi bo-le pi.
3SG BA orange peel-PFV peel
‘He peeled the peel off the orange.’

a. c-structure:



b. f-structure:



In (39), (↑ SUBJ), *he*, is responsible for affecting (↑ OBJ), the orange, in a way described by (↑ XCOMP), i.e., *he peeled the peel off the orange*. Note that (↑ SUBJ) controls the SUBJ in XCOMP, and (↑ OBJ) controls the TOPIC in XCOMP, which is in turn anaphorically linked to OBJ.

In summary, *ba* and *bei* are both treated as three-place predicates. While the former involves a causer as SUBJ, an affectee as OBJ, and an active proposition describing the caused event as XCOMP, the latter involves an affectee as SUBJ, a causer as OBJ, and a passive proposition describing the caused event as XCOMP. Thus, in this sense the *bei* construction is the passive counterpart of the *ba* constructive.

See also Yang (2020) for a discussion of the impersonal BEI-passive in Mandarin.

4.2 Cantonese

A discussion on aspects of the passive structure in Cantonese is offered in Chow (2019). While Cantonese shares a phonologically similar passive morpheme *bei* with Mandarin, the two counterparts differing only in tones, one clear morphosyntactic difference is that the NP following *bei* in Mandarin is optional (40a), while that in Cantonese is obligatory (40b). In other words, the 'short' passive discussed in the previous section is not allowed in Cantonese. Even in agentless passives, the NP *jan* 'person' must follow *bei*.

(40) a. Mandarin
Zhangsan bei (Lisi) daa-le.
Zhangsan BEI Lisi hit-PFV
'Zhangsan has been hit (by Lisi).'

b. Cantonese
Siuming bei *(jan) daa.
Siuming PASS people hit
'Siu Ming was beaten up.'

Based on this, Chow (2019) argues that passivization in Cantonese involves the subject being linked to an oblique object, a non-core argument (Bresnan 1982c; Chow 2019: 232). It is also shown that, unlike Kit's (1998) and Her's (2009) analyses for the Mandarin *bei*, the Cantonese *bei* is a 'non-argument taking and a non-predicative' coverb (Chow 2019: 186), which contributes a (↑ VOICE)=PASS feature to f-structure.

Similar to Her (2009), Chow (2019) acknowledges that the matrix subject in a passive structure is linked to the topic role. Indeed, the same propositional content may be expressed by an active, a 'direct' or canonical passive (41a), or a 'indirect' passive (42b) structure, depending on the information structure to be expressed. In an canonical passive structure, the entire theme-NP is topical

– it is expressed as the subject. In an 'indirect' passive structure, however, it is the possessor of the theme-NP which is topical – the possessor is linked to the subject.

(41) Cantonese

a. The 'direct' or canonical passive

[Can saang gaa ce] bei tungsi zong-laan zo.
Mr. Chan CLF car PASS colleague crash-broken PFV
'Mr. Chan's car has been crashed by his colleague.'

b. The 'indirect' passive

[Can saang] bei tungsi zong-laan zo [gaa ce].
Mr. Chan PASS colleague crash-broken PFV CLF car
'Mr. Chan had his car crashed by his colleague.'

Semantically, the subject must be adversely affected in order for an indirect passive to be acceptable. Chow (2018, 2019) proposes that, for the indirect passive structure [NP1 BEI2 NP2 V NP3] to be licensed, an additional malefactive role, which must be topical, is introduced into the structure. Due to the limits of space, we shall leave the discussion here and ask interested readers to refer to these studies.

5 Dative alternation

Dative alternations, as well as ditransitive constructions, have been extensively discussed in the Chinese linguistics literature. In addition to the word order variations and the introduction of an applied object common in other languages (e.g. Bresnan et al. 2007), the challenges in analyzing the Mandarin dative alternative involve the position and the grammatical status of the lexical form *gei* 'give' (e.g. Chao 1968; Zhu 1982). The discussion in this section focuses on Mandarin only, as the dative alternation is not attested in Cantonese (Lam 2008).

(42) Ditransitive constructions with *gei* in Mandarin (42a, 42c & 42d are from Huang & Ahrens 1999)

a. SUBJ *gei* IO V DO

b. SUBJ DO V (*gei*) IO

c. SUBJ V (*gei*) IO DO

d. SUBJ V DO *gei* IO

The pattern in (42d) will be treated as the semantically most transparent word order for *gei* 'give.to'. It should, however, be noted that it is not clear whether a clearly favoured canonical word order is available (Yao & Liu 2010). In a ditransitive construction, *gei* introduces the IO as the goal towards which a theme DO moves. In the literature, there are several different views regarding the grammatical status of *gei*: *gei* is (i) a verb, producing a serial verb construction with the other verb in the construction; (ii) a co-verb/preposition, marking the IO in the construction. The verb/preposition debate is familiar in the Chinese linguistics literature, and has been applied to several other lexical items with similar distributions Either account is generally adequate in describing the patterns in (42a) and (42d). The patterns in (42b) and (42c), where *gei* is optional and the optionality depends on the V, has generated interesting debates specific to the ditransitive construction (e.g. Chao 1968; Li & Thompson 1981; Cheng & Huang 1988). Huang & Ahrens (1999) observe that verbs without an inherent meaning of transfer (e.g. *ti* 'to kick', *bian* 'to knit'), typically require the presence of *gei*, while *gei* in structures with verbs with an inherent meaning of transfer may be optional (e.g. *song* 'to give as a gift, to send', *zhu* 'to lend', *mai* 'to sell,' and *gei* 'to give'). This suggests that the *gei* immediately after the verb is a stem that introduces an applicative goal role to the argument structure of the verb. This account has been incorporated into Huang's (1993a) LMT of Mandarin. The compounding account has also been adopted by several Construction Grammar-based accounts (e.g. Ahrens 1995; Zhang 1999; Liu 2006). Huang (1993a) argues that the postverbal *gei* is a part of the complex predicate which involves a morpholexical rule introducing an additional goal role into the argument structure. The study has also observed that there is a significant contrast between the English and Chinese dative constructions - the theme can become SUBJ in a passive construction in Mandarin, but not the goal.

(43) a. Mary gave John a book.
b. John was given a book by Mary.
c. A book was given to Mary by John.

(44) Mandarin (adapted from Huang (1993a: example 22))
a. Zhangsan ti-*(gei) Lisi yi ge qiu.
Zhangsan kick-GEI Lisi one CLF ball
'Zhangsan kicked a ball to Lisi.'
b. *Lisi (bei Zhangsan) ti-gei-le yi ge qiu.
Lisi BEI Zhangsan kick-GEI-PFV one CLF ball

c. *nei ge qiu (bei Zhangsan) ti-gei-le Lisi.
that CLF ball BEI Zhangsan kick-GEI-PFV Lisi
'That ball was kicked to Lisi (by Zhangsan).'

In sum, this account of V-*gei* compounding, adding an applicative GOAL role, illustrates the lower accessibility of the goal role on the Thematic Hierarchy, and predicts that the goal role cannot be linked to SUBJ in a Mandarin passive structure.

See also Her (2006a) for an alternative analysis of the Mandarin dative alternation.

6 Compounds

Compounding is a productive morpholexical process in Chinese (Hsieh et al. 2022). Mandarin is known to have at least the following types of compounds that can introduce new predicate-argument structures: (i) subject-verb (SV) compounds; (ii) verb-object (VO) compounds; (iii) verb-resultative (VR) compounds; and, (iv) verb-verb (VV) compounds. In this section, the LFG treatments of resultative compounds and VO compounds are presented in Sections 6.1-6.2 and Section 6.3, respectively.

6.1 Early LFG studies on Mandarin compounds

Chao (1968) has observed that a number of most distinctive grammatical features of Chinese are related to the prevalence of compounds: (i) V+N compounds tend not to take another object directly; (ii) the noun in the compound is often separable even though it is a sub-lexical unit (called 'ionization' in Chao 1968); (iii) separable compounds allow certain degrees of internal modification, and in some cases, an object may appear in non-canonical positions. The earliest published studies in the LFG literature on Mandarin, Huang (1985, 1986, 1988, 1990) for example, have aimed to account for these separable compounds and their non-canonical object positions.

Huang (1990) provides an account for VO compounds in Mandarin. One example that is of particular interest is the idiom chunk *chi cu* 'be jealous of', consisting of the lexical verb *chi* 'eat' and the noun *cu* 'vinegar'. The chunk is a non-compositional compound, as the overall meaning is only available if both the V and the N are found in the sentence. What is interesting, and yet challenging, is the fact that the V and the N in the compound can be separated, by *de* in the following example:

(45) Mandarin
Sanbai conglai bu chi Yunniang de cu.
Sanbai ever NEG eat Yunniang DE vinegar
'Sanbai is never jealous of Yunniang.'

Huang (1990) proposes an account for separable compounds in terms of lexical discontinuity – both the verb and the separable noun contribute information to the overall interpretation. The subscript marks the use of this form as a component of an idiom. Note that PRED is associated with the noun, and the constraining equation ensures that the non-compositional meaning will only be available if the form *chi* also occurs in the sentence.

(46) a. chi_2 V (↑ VMORF) = CHI
b. cu_2 N (↑ PRED)='BE.JEALOUS⟨SUBJ OBL⟩'
(↑ VMORF) $=_c$ CHI
(↑ CL) = DE

Huang (1990) further shows that this proposal successfully accounts for various constructions in which the compound occurs, including topicalization. This example illustrates how complex structures can be captured with simple lexical rules.

6.2 VO compounds in Cantonese

VO compounds are found in Cantonese, too. As discussed in the previous section and as observed in Bodomo et al. (2017), among others, the challenge with analyzing VO compounds is that they seem to be lexical in that their meanings are often non-compositional and depend on the co-occurrence of a V and some specific N; but, at the same time, they seem to be phrasal in that other constituents can clearly be inserted in between the V and the N:

(47) Cantonese
a. jau-seoi
swim-water
'swim'
b. ngo jau-zo zan seoi.
1SG swim-PFV for.a.while water
'I have swum for a while.' (Bodomo et al. 2017: 389, ex. 18)

Bodomo et al. (2017) treat *seoi* ‘water’ as a syntactic object, whose form is obligatorily required to give the target meaning (hence the FORM feature in OBJ below), but it is not subcategorized for by the PRED, as the VO compound *jau-seoi* ‘swim’ requires only an agent argument at a-structure and *seoi* ‘water’ is athematic in the compound (Bodomo et al. 2017: 389):

(48)
$$\left[\begin{array}{ll} \text{PRED} & \text{‘SWIM}\langle\text{SUBJ OBJ}\rangle\text{’} \\ \text{SUBJ} & \left[\begin{array}{ll}\text{PRED} & \text{‘PRO’}\\ \text{NUM} & \text{SG} \\ \text{PERS} & 1\end{array}\right] \\ \text{ASP} & \text{PFV} \\ \text{OBJ} & \left[\text{FORM}\ \ \text{SEOI}\right] \\ \text{ADJ} & \left\{\left[\text{PRED}\ \ \text{‘ZAN’}\right]\right\}\end{array}\right]$$

Bodomo et al. (2017) apply this analysis to Mandarin VO compounds, too. Che & Bodomo (2018) discuss Mandarin VO compounds, as well as idioms, and adopt a complex predicate analysis for VO compounds.

A complex predicate approach has also been proposed to analyze serial verb constructions, which are common in Chinese. See Bodomo et al. (2003) for a syntactic and semantic account of Cantonese serial verb constructions involving the benefactive role.

6.3 Resultative compounds

Chinese resultative compounds involve the concatenation of two verbs, and the merge of their predicate argument structures. They are called resultative compounds (VR) because the first verb denotes an action, and the second verb typically refers to the result caused. Previous studies have found that both verbs contribute to the argument structure of the compound. Li (1990) proposes a structure-based account that allows most possible predicate-argument structures, but fails to select the correct reading among other possibilities. Huang & Lin (1992) assume that VV compounds in Mandarin represent composite event structures and the complex predicate formation can be resolved with morpholexical mapping based on prototypical argument templates. Li (1995) proposes another account based on the causative hierarchy. Her (2004, 2007) offers an LFG account by incorporating unified mapping principles of LMT.

Her’s (2004; 2007) account focuses on cases where the first V has either one or two arguments, while the second V has only one argument. In addition, it is assumed that the VR compounds have two arguments. Hence, there are cases in which each verb contributes an argument, or the more complicated cases where the argument from the second verb can be merged with either the first or the second argument of a transitive verb, such as *niu* 'to wring' in (50). The two argument merging scenarios are given in (49).

(49) V-V Resultative Compounding

$V_1\langle x\ y\rangle + V_2\langle z\rangle \rightarrow$ (i) $\langle x\ y\text{-}z\rangle$

(ii) $\langle x\text{-}z\ y\rangle$

Given that the resultative compound is transitive, thus a two-place predicate, the single role of V_2 must join one of the two roles of V_1 and form a composite role. Logically, two possibilities are obtained as shown in (49), but three patterns of argument-function linking are observed, as in (50)-(52). Note also that a causative reading is also obtained, except in (51).

(50) Mandarin; causative

Lisi	niu-gan-le	maojin.
Lee	wring-dry-PFV	towel

'Lee wrung the towel dry.'

$\langle x$	$y\text{-}z\rangle$
↓	↓
SUBJ	OBJ
Lee	towel

(51) Mandarin; non-causative

Zhangsan	chi-yan-le	zhe	zhong	dongxi.
John	eat-tired-of-PFV	this	kind	stuff

'John got tired of eating this kind of stuff.'

$\langle x\text{-}z$	$y\rangle$
↓	↓
SUBJ	OBJ
John	stuff

(52) Mandarin; causative

zhe	zhong	dongxi	hui	chi-si	ni.
this	kind	stuff	will	eat-dead	2SG

Eating this kind of stuff will make you dead.'

$\langle x\text{-}z$	$y\rangle$
SUBJ	OBJ
stuff	you

Her's (2007) resultative compound rules are given below in (53).

(53) V-V Resultative Compounding

$V_{caus}\langle x\ y\rangle + V_{res}\langle z\rangle \rightarrow V_{caus}V_{res}\langle \alpha\ \beta\rangle$, where $\langle \alpha\ \beta\rangle^{*}$ =

(i) $\langle x\ y\text{-}\cancel{z}\rangle$

(ii) $\langle x[\text{caus}]\ \cancel{y}\text{-}z[\text{af}]\rangle$

(iii) $\langle x\text{-}\cancel{z}\ y\rangle$

(iv) $\langle \cancel{x}\text{-}z[\text{af}]\ y[\text{caus}]\rangle$

*Unsuppressed *z* and the other unsuppressed role receive [af] and [caus], respectively

With these rules, and a modified version of LMT, all possible interpretations of resultative compounds with V⟨*x y*⟩ and V⟨*x*⟩ combinations can be accounted for. See Her (2007) for details.

7 Lexical Mapping Theory and locative inversion

LFG crucially observes radical lexicalism (Karttunen 1989), and views grammatical operations as the projection and unification of mentally represented lexical information (Bresnan 1982b). Word order variations and alternations are not accounted for by transformational rules, but by the projection and unification of the mental representation of information from conceptual structure to c-structure. See Belyaev forthcoming(b) [this volume] and Belyaev forthcoming(a) [this volume] for a discussion on the architecture of LFG.

The introduction of Lexical Mapping Theory (LMT) to LFG to derive lexicalized argument structures in terms of GFs is crucial in allowing the theory to account for concept-driven lexicalization. It also provides an elegant way to account for word order and other typological variations. LMT formulates rules to capture how conceptualized event structures are lexicalized as argument structures to mediate mapping to functional structures (Bresnan & Kanerva 1989; Alsina 1993). Huang (1993a) proposes an adapted LMT for Mandarin, adopting previous assumptions that the mapping is determined by the thematic hierarchy, and the theory of intrinsic and default classification of grammatical functions. The adaptations are proposed, taking into consideration both the theoretical concerns to incorporate Dowty's (1991) Proto-role properties, and the need to capture several atypical argument realization patterns in Mandarin. These patterns include the NP realization of extent/dimension (54a), and the use of time/location NPs instead of pleonastic pronouns in the subject position in presentative constructions (54b).

(54) Mandarin

a. Ta tile wo **yi jiao**.
s/he kick.PFV I/me one foot
'S/he kicked me once'

b. Qiangshang guale ji fu hua.
wall.top hang.PFV several CLF painting
'There are several paintings on the wall.'

Huang also provides evidence to show that the GOAL role is below the THEME role on the thematic hierarchy in lexicalized compounds, idiom chunks and ditransitive verbs. The thematic hierarchy for Mandarin is thus revised, as shown in (55). The intrinsic and default classification of grammatical functions (57a) are slightly modified to simplify feature assignments, and to accommodate the locative inversion construction in Mandarin.

(55) Thematic hierarchy for Mandarin Chinese (Huang 1993a)
agent > ben/mal > instr > th/pat > exp/goal > loc/dom

Huang & Her (1998) and Her (2010) propose a simplified LMT. This proposal keeps the universal thematic hierarchy, with the assumption that morpholexical operations can replace the Subject Condition. Note that the two proposals take different approaches to accommodate the Mandarin Chinese data. Huang (1993a) has revised the thematic hierarchy, but has kept intrinsic and default classification of grammatical functions, while Her (2010) has kept the thematic hierarchy (56), but has adjusted the criteria for the ±r(estricted) and ±o(bjective) specifications. The different proposals aim to account for several important generalizations in Chinese, some of which will be discussed below.

(56) Thematic Hierarchy for Mandarin Chinese (Her 2010)
ag > ben > go/exp > inst > pt/th > loc

In terms of the classifications, the (−) values, considered less marked than the [+] values, are thus given a higher position on the hierarchy. Her (2010) also assumes that that [−r (unrestricted)] is less marked than [−o] (non-object-like), given that [−r] GFs are not restricted to specific argument roles, Huang (1993a) does not make the same assumption.

(57) Markedness Hierarchy of Grammatical Functions:

a. SUBJ ([−r −o]) > OBJ ([−r +o]) / OBL_θ ([+r −o]) > OBJ_θ ([+r +o]) (Huang 1993b)

b. SUBJ ([−r −o]) > OBJ ([−r +o]) > OBL$_\theta$ ([+r −o]) > OBJ$_\theta$ ([+r +o]) (Her 2010)

See also Fu (1993) and Pan (1997) for introductions to LMT published in Chinese journals.

Locative inversion is heavily influenced by considerations at information structure (Bresnan 1989; Dalrymple 2001: 209). It is also known as presentative or existential construction. Gu (1992, 1997) assumes that most of verbs which may participate in the locative inversion in Mandarin are derived from transitive verbs. Pan (1996, 1997) argues that it is necessary to distinguish two types of locative inversion, based on the presence of the aspectual markers *-le* PFV or *-zhe* DUR on the verb. Huang et al. (1999) shows that the range of different meanings associated with the locative inversion and the presentative sentences can be accounted for by considering the interaction of constructional and lexical meanings. Cui & Yuan (2020) suggest that existential sentences exhibit features of ergativity.

The challenge that the locative inversion presents to LFG, especially to LMT, is how it is possible to map the locative role, ranked low on the thematic hierarchy, to the most prominent grammatical function SUBJ. Bresnan & Kanerva (1989), based on data from Chicheŵa, propose a special default rule for the presentational focus construction. The rule assumes that the locative phrase bears the focus feature and ensures that a locative [−r] argument appears. Bresnan (1994) extends the account to English. Huang & Her (1998), however, shows that the proposal cannot account for the locative inversion in Mandarin, especially in constructions involving three-place predicates, such as *fang* 'put':

(58) Mandarin

a. Lisi fang-qian zai zhuo-shang.
Lisi place-money at table-top
'Lisi placed some money on the table.'

b. qian (Lisi) fang zai zhuo-shang.
money Lisi placed at table-top
'Money was placed on the table by Lisi.'

c. zhuo-shang (Lisi) fang-le qian.
table-top Lisi place-PFV money
'On the table was placed some money.'

Crucially, both (58b) and (58c) are treated as locative inversion structures. There is, however, evidence suggesting that (58b), in fact, involves topicalization, but not locative inversion. First, *qian* 'money' is not a locative phrase. Second, the verb in (58b) does not require the presence of the aspectual markers -*zhe* DUR or -*le* PRF, unlike the verb in well-accepted Mandarin locative inversion structures. The preposed NP in (58b) can therefore be treated as a regular topicalized phrase, without further stipulations. See also Lui (2020) for a discussion of the locative inversion in Cantonese.

8 Classifiers and measure words

Mandarin is a textbook example of a numeral classifier language. As a lexical category, numeral classifiers have two subcategories, namely sortal classifiers (C), *aka* classifiers; and mensural classifiers (M), *aka* measure words (Huang & Shi 2016). See (59) and (60) for examples of Cs and Ms, respectively (Her 2012b).

(59) Mandarin

a. san gen xiangjiao
   3 CLF banana
   '3 bananas'
b. yibai ben shu
   100 CLF book
   '100 books'
c. shi pi ma
   10 CLF horse
   '10 horses'

(60) Mandarin

a. san da xiangjiao
   3 M-dozen banana
   '3 dozens of bananas'
b. yibai xiang shu
   100 M-box book
   '100 boxes of books'
c. shi qun ma
   10 M-herd horse
   'ten herds of horses'

C and M consistently appear after a numeral (Num) and before a noun (N) and are mutually exclusive in this position, as only one C/M can be used. It is a near consensus in the Chinese linguistics literature to assign the same phrasal structure to them. The syntactic position is typically called the classifier position. See Jiang et al. (2022) for a summary of syntactic approaches, and Chen et al. (2022) for a summary of semantic approaches to the Chinese classifier system.

Cs and Ms, however, do exhibit some differences (Chao 1968; Her 2017; see also Huang 2015 for an ontological account). In terms of modification, the adjective, whether it is found before or after a C, modifies the head N. (61a) and (61b) therefore have the same meaning. An adjective in a nominal structure with an M, however, modifies the immediately following element. Thus, in (62a), *da* 'big' modifies *xiang* 'box', yielding the meaning 'one big box of apples', while in (62b), *da* 'big' modifies *pingguo,* yielding the meaning 'one box of big apples' (Her 2012b):

(61) Mandarin sortal classifiers

a. yi da ke pingguo
   1 big CLF apple

b. yi ke da pingguo
   1 CLF big apple
   'one big apple'

(62) Mandarin mensural classifiers

a. yi da xiang pingguo
   1 big M-box apple
   'one big box of apples'

b. yi xiang da pingguo
   1 M-box big apple
   'one box of big apples'

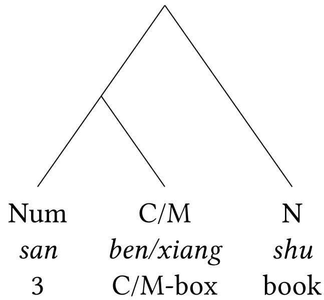
Another difference between Cs and Ms is that the former has the fixed numeral value of precisely *1*, while Ms can be of any value, numerical or non-numerical, except *1*, as shown in (63). In (63), K is a C or M, and k is the mathematical value of K.

(63) C/M distinction in mathematical values
   [Num $\underline{\text{K}}$ N] = [NUM×*k*N], where K=C *iff* $k = 1$, otherwise K=M.

The LFG account offered in Her (2012b) assigns a left-branching c-structure to C/M, as in (64), consistent with the traditional approach but contra the dominant right-branching structure preferred in recent derivational syntax. See Her

(2017) and Her & Tsai (2020) for arguments from typological as well as Mandarin-internal perspectives. Sample lexical entries of N, Num, C, and M are given in (65).

(64) Unified left-branching c-structure of the classifier construction



(65) Sample lexical entries

a. *shu* N (↑ PRED)=‘BOOK’
(↑ PROFILABLE)={BEN本,CE冊}

b. *san* Num (↑ CARD)=3

c. *ben* C/M (↑ PROFILED)=BEN本

d. *xiang* C/M (↑ PRED)=‘BOX’

Cs and Ms are two subcategories of a single lexical category C/M. Their differences are located in f-structure. CLF in (65c) has no PRED, but M in (65d) does; C, however, has a feature PROFILED, whose value is the essential property each CLF serves to profile, or highlight. A noun can only have one or more of its essential features profiled and may thus co-occur with more than one C, though one at a time as a formal requirement. In (65a), for example, *shu* ‘book’ normally takes the CLF *ben*, but *ce* is also an option, accounted for by the feature PROFILABLE, which takes a set, {BEN, CE}, as its value. The relevant annotated phrase structure rules are given in (66).

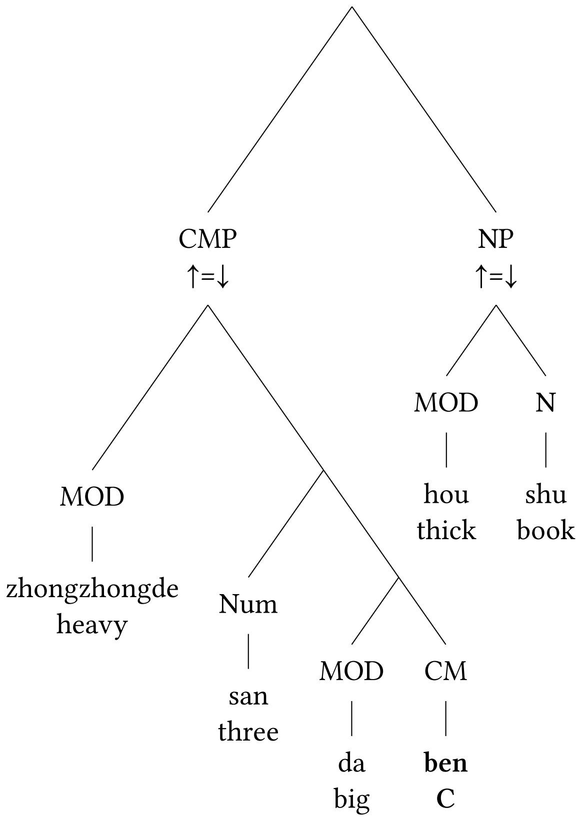
(66) Annotated phrase structure rules for the classifier construction

a. NP ⟶ ... C/MP ... N
(↓ PRED) ⇒ (↑ QUANTIFIER)=↓ ; ↑=↓ (N)
¬(↓ PRED) ⇒ {↑=↓, (↓ PROFILED) $\in_c$ (↑ PROFILABLE)}

b. C/MP ⟶ ... Num ... C/M
Num: ↑=↓
C/M: ↑=↓, ((↑ CARD)=1)

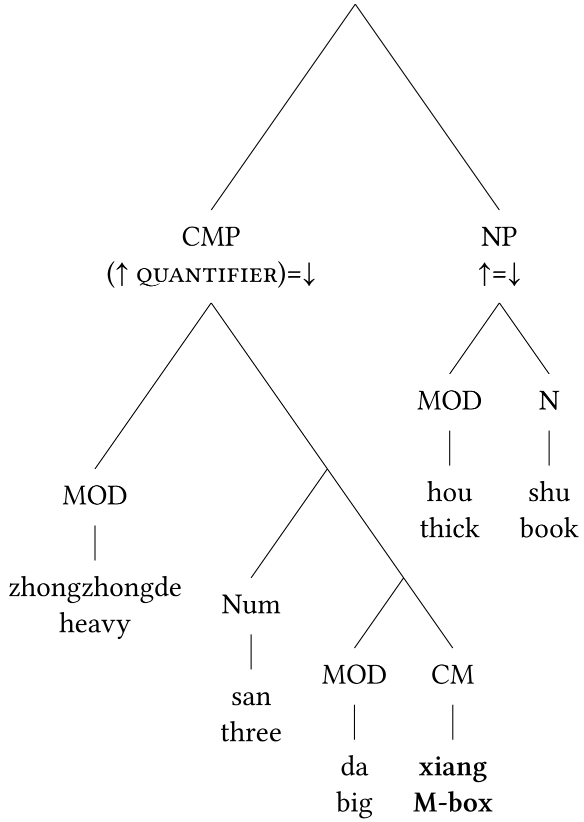
The notation ‘A⇒B’ in (66) means ‘if A, then B’. Thus, in a C/MP, if it has PRED, indicating it is an M, then the information goes in a QUANTIFIER function; if it does not have PRED, indicating it is a C, then it serves as a co-head with N and its PROFILED value must be a member of N’s PROFILABLE set of values. The c-structure and f-structure of two nominal phrases with a CLF and an M are given in (67) and (68), respectively.

(67) Mandarin
zhongzhongde san da ben hou shu
heavy 3 big CLF thick book
‘three heavy big thick books’



$$\left[\begin{array}{ll} \textsc{pred} & \textsc{book} \\ \textbf{\textsc{profiled}} & \textbf{\textsc{ben}}\ \textbf{本} \\ \textsc{profilable} & \left\{\textsc{ben}\ 本,\ \textsc{ce}\ 冊\right\} \\ \textsc{card} & 3 \\ \textsc{adjuncts} & \left\{\begin{array}{l} \left[\text{“}\textsc{heavy}\text{”}\right] \\ \left[\text{“}\textsc{big}\text{”}\right] \\ \left[\text{“}\textsc{thick}\text{”}\right] \end{array}\right\} \end{array}\right]$$

(68) Mandarin
zhongzhong-de san da xiang hou shu
heavy-DE 3 big M-box thick book
'three heavy big boxes of thick books'



$$\left[\begin{array}{ll}
\textsc{pred} & \textsc{book} \\
\textsc{adjuncts} & \left\{\left[\text{“}\textsc{thick}\text{”}\right]\right\} \\
\textsc{profilable} & \left\{\textsc{ben}\ 本,\ \textsc{ce}\ 冊\right\} \\
\textsc{card} & 3 \\
\textsc{quantifier} & \left[\begin{array}{ll}
\textsc{pred} & \text{‘}\textsc{box}\text{’} \\
\textsc{card} & 3 \\
\textsc{adjuncts} & \left\{\begin{array}{l}\left[\text{“}\textsc{heavy}\text{”}\right] \\ \left[\text{“}\textsc{big}\text{”}\right]\end{array}\right\}
\end{array}\right]
\end{array}\right]$$

The parallel architecture of c-structure and f-structure allows Cs and Ms to belong to one syntactic category and (67) and (68) thus share the same c-structure, while their differences are captured in the f-structure, where a CLF serves as a co-head of the nominal construction and an M serves as the head of a quantifier phrase.

See Börjars et al. (2018) for a different proposal for the c- and f-structures of Mandarin noun phrases containing classifiers and measure words, and Huang & Ahrens (2000) for a discussion on kind and event classifiers in Mandarin.

9 Other properties and phenomena

A number of other properties and phenomena are prominent in Chinese as well, and studies of these are available in the very large body of LFG literature on the analysis of Chinese. However, due to constraints of space and scope, we cannot discuss all of these in detail in this chapter. This section will hopefully serve as a pointer to some of these works. The syntax of Mandarin questions has been investigated in Shiu & Huang (1989) and Huang (1993b). Relativization and topicalization phenomena in Mandarin have been studied in Huang (1992), where the author proposes a functional uncertainty analysis (Kaplan & Zaenen 1989). Huang (1988) analyses ‘possessive subjects’ in Mandarin, while Huang (1990) offers an LFG account of possessive-object constructions in Mandarin, showing how these display lexical discontinuity. Chief (1996) explores an LFG account of Mandarin reflexive verbs. Dong (2016) provides an LFG analysis of pronominal binding in Mandarin. Lam (2020) investigates anaphoric and functional control in Mandarin. Che (2014) is a study of particles in Mandarin.

10 NLP applications of LFG in Chinese

LFG has played an important role in the development of Chinese NLP. Joan Bresnan, Ronald Kaplan, Lauri Karttunen and Annie Zaenen visited Taiwan at the dawn of Chinese computational linguistics in 1989 and made lasting impact (Bresnan 1989). One of the immediate outcomes was the Information-based Case Grammar (ICG, Chen & Huang 1990), the first comprehensive grammar of Chinese that incorporated features of both LFG and HPSG. Her et al. (1991) and Her (1995) describe a rule-based commercial machine translation system for English-Chinese, where parsing, transfer and generation are all based on LFG. This system was later acquired by Apptek (https://www.apptek.com/) and expanded to include multiple language pairs and many other NLP applications. Kit (1992, 1993a,b) and Kit & Webster (1992) are also among the earliest studies applying LFG assumptions to parse Chinese. Webster & Kit (1995) describe the use of a 'Chinese-Lexical Functional Grammar (C-LFG)' parser to analyze simple sentences from texts. Sun (2001) outlines the computational implementation of LFG in Chinese. Fang & King (2007) provide an LFG grammar of Mandarin for machine use. Guo et al. (2008) describes LFG-based generation for Chinese, while Burke et al. (2004) and Guo (2009) describe LFG-based Chinese treebanks. Chief et al. (2000) present a corpus-based approach to the analysis of synonyms in Chinese. Jiang et al. (2018) annotate Chinese light verb constructions according to the paradigm of PARSEME, a platform built based on LFG and other theoretical frameworks.

11 Conclusion: LFG and Chinese Linguistics

The assumptions of LFG have been applied to the research on a number of grammatical phenomena in Chinese languages since Huang (1985). A number of LFG-based studies on Chinese have made a significant impact to Chinese linguistics. Huang & Mangione (1985), one of the earliest LFG papers on Chinese, has inspired Huang's (1988) treatment of, and a long debate on, the status of V1 and V2 in the Mandarin resultative verb construction. Interestingly, the V2-as-matrix-verb analysis, initially proposed by Huang & Mangione (1985), is gradually emerging as a possible consensus. Similarly, the functional uncertainty of LFG allows a transparent account of Mandarin long-distance dependencies without abstract levels and movements (Huang 1992). Huang (1993a) first introduced the concept of applicatives to Mandarin, and initiated many interesting discussions in Chinese linguistics in the past 20 years. LFG studies Huang 1989a; Tan 1991; Her 1991) on the TOPIC and SUBJ functions in Chinese have contributed to the ongoing topic/subject debate in Chinese. LFG studies have also provided crucial insights to the understanding of the *ba* and *bei* constructions in Chinese (e.g. Her 1989; Bender 2000; Her 2009), especially in terms of treating *ba* and *bei* as the main predicate. The seeming dilemma of Chinese compounds displaying lexical non-compositionality and phrasal compositionality (e.g. the separable compounds) can be straight-forwardly dealt with by adopting the assumptions of LFG. This is perhaps one of the topics receiving the most attention in the LFG literature on Chinese, including but not limited to Huang (1990), Huang & Lin (1992), Her (1996, 1997), and Bodomo et al. (2017).

Accounts of Chinese languages have contributed to the development of the LFG framework, too. Shiu & Huang (1989) was one of the first LFG accounts on sentential clitics (e.g. Mandarin question particles). Huang (1992, 1993b) applies the concept of functional uncertainty to account for Mandarin data. Her (2006b) introduces the concepts of interaction and optimality to LMT. Her (2012a,b) provides a full account of the classifier system. Finally, Bodomo (2001) and colleagues' work on Cantonese and Zhuang have added to the typological diversity of LFG research.

Acknowledgments

We thank our reviewers for their constructive comments, which have helped improve the chapter in many ways. All remaining errors are our own.

Abbreviations

Besides the abbreviations from the Leipzig Glossing Conventions, this chapter uses the following abbreviations.

EXP	experiential
M	measure word
PART	particle
ZAI	marker meaning 'now' or 'at the moment'

References

Ahrens, Kathleen. 1995. The meaning of the double object construction in Mandarin. In Jose Camacho & Lina Choueiri (eds.), *Proceedings of the Sixth North American Conference on Chinese Linguistics*, vol. 1, 1–10. Los Angeles: GSIL, University of Southern California.

Alsina, Alex. 1993. *Predicate composition: A theory of syntactic function alternations*. Stanford: Stanford University. (Doctoral dissertation).

Belyaev, Oleg. Forthcoming(a). Core concepts of LFG. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Belyaev, Oleg. Forthcoming(b). Introduction to LFG. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Bender, Emily. 2000. The syntax of Mandarin bǎ: Reconsidering the verbal analysis. *Journal of East Asian Linguistics* 9(2). 105–145.

Bodomo, Adams. 2001. LFG and the analysis of Chinese: Introduction to the workshop. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '01 conference*, 353–358. Stanford: CSLI Publications.

Bodomo, Adams. 2011. Reflexivity without apparent marking: The case of Mashan Zhuang. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '11 conference*, 68–88. Stanford: CSLI Publications.

Bodomo, Adams, Olivia S.-C. Lam & Natalie S. S. Yu. 2003. Double object and serial verb benefactive constructions in Cantonese. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '03 conference*, 59–74. Stanford: CSLI Publications.

Bodomo, Adams & Sophia Yat-Mei Lee. 2003. On the function COMP in Cantonese. *Journal of Chinese Linguistics Monograph Series* 19. 96–116.

Bodomo, Adams & Kang Kwong Luke. 2003. Lexical-functional grammar analysis of Chinese. *Journal of Chinese Linguistics* 19.

Bodomo, Adams, So-sum Yu & Dewei Che. 2017. Verb-object compounds and idioms in Chinese. In Ruslan Mitkov (ed.), *Europhras 2017: Computational and corpus-based phraseology*, 383–396. Dordrecht: Springer. DOI: 10.1007/978-3-319-69805-2_27.

Börjars, Kersti, Christopher Hicks & John Payne. 2018. Interdependencies in Chinese noun phrases. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '18 conference*, 109–128. Stanford: CSLI Publications.

Bresnan, Joan. 1982a. Control and complementation. *Linguistic Inquiry* 13(3). Reprinted in Bresnan & Kaplan 1982, 343–434.

Bresnan, Joan (ed.). 1982b. *The mental representation of grammatical relations*. Cambridge, MA: The MIT Press.

Bresnan, Joan. 1982c. The passive in lexical theory. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 3–86. Cambridge, MA: The MIT Press.

Bresnan, Joan. 1989. The syntactic projection problem and the comparative syntax of locative inversion. In Chu-Ren Huang & Keh-Jiann Chen (eds.), *Proceedings of ROCLING 1989*, 375–396. http://www.aclclp.org.tw/rocling/1989/M14.pdf.

Bresnan, Joan. 1994. Locative inversion and the architecture of Universal Grammar. *Language* 70(1). 72–131. DOI: 10.2307/416741.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & R. Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Bouma, Irene Krämer & Joost Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. KNAW.

Bresnan, Joan & Jonni M. Kanerva. 1989. Locative inversion in Chicheŵa: A case study of factorization in grammar. *Linguistic Inquiry* 20(1). 1–50.

Bresnan, Joan & Ronald M. Kaplan. 1982. Introduction: Grammars as mental representations of language. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, xvii–lii. Cambridge, MA: The MIT Press.

Burke, Michael, Olivia S.-C. Lam, Aoife Cahill, Rowena Chan, Ruth O'Donovan, Adams Bodomo, Josef van Genabith & Andy Way. 2004. Treebank-based acquisition of a Chinese lexical-functional grammar. In *Proceedings of the 18th Pacific Asia Conference on Language, Information and Computation*, 161–172. Tokyo.

Burusphat, Somsonge & Xiaohang Qin. 2012. Zhuang word structure. *Journal of Chinese Linguistics* 40(1). 56–83.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles N. Li (ed.), *Subject and topic*, 27–55. New York: Academic Press.

Chang, Li-Ping, Chu-Ren Huang & Keh-jiann Chen. 1988. 中文類及物動詞的現象 [Pseudo-transitive verbs in Mandarin Chinese]. In *Proceedings of the Second International Conference on the Teaching of Chinese as a Second Language*.

Chao, Yuen-Ren. 1968. *A grammar of spoken Chinese*. Berkeley/Los Angeles: University of California Press.

Che, Dewei. 2014. *The syntax of particles in Mandarin Chinese*. Hong Kong: University of Hong Kong. (Doctoral dissertation).

Che, Dewei & Adams Bodomo. 2018. A constraint-based analysis of the objects of VO verbal compounds in Mandarin Chinese. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '18 conference*, 171–190. Stanford: CSLI Publications.

Chen, I-Hsuan, Kathleen Ahrens & Chu-Ren Huang. 2022. The Chinese classifier system as a lexical-semantic system. In Chu-Ren Huang, Yen-Hwei Lin, I-Hsuan Chen & Yu-Yin Hsu (eds.), *Cambridge Handbook of Chinese Linguistics*. Cambridge, UK: Cambridge University Press.

Chen, Keh-jiann & Chu-Ren Huang. 1990. Information-based case grammar. In *COLING '90 volume 2: Proceedings of the 13th International Conference on Computational Linguistics*. DOI: 10.3115/997939.997949.

Chen, Keh-Jiann, Chu-Ren Huang, Liping Chang & Hui-Li Hsu. 1996. SINICA CORPUS: Design methodology for balanced corpora. In *Proceedings of the 11th Pacific Asia Conference on Language, Information and Computation (PACLIC11)*, 167–176.

Cheng, Lisa Lai Shen & Rint Sybesma. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30(4). 509–542. DOI: 10.1162/002438999554192.

Cheng, Y. H. & Chu-Ren Huang. 1988. Guoyu zhong de shuangbin dongci. [Double object verbs in Mandarin]. *The World of Chinese Language* 48. 34–48.

Chief, Lian-Cheng. 1996. An LFG account of Mandarin reflexive verbs. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '96 conference*. Stanford: CSLI Publications.

Chief, Lian-Cheng, Chu-Ren Huang, Keh-Jiann Chen, Mei-Chih Tsai & Li-Li Chang. 2000. What can near synonyms tell us? *International Journal of Computational Linguistics & Chinese Language Processing* 5(1). 47–59.

Chow, Pui-Lun. 2018. The malefactive topic role in Cantonese indirect passives. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '18 conference*, 191–207. Stanford: CSLI Publications.

Chow, Pui-Lun. 2019. *Passive and permissive constructions in Cantonese: An LFG account*. Hong Kong: University of Hong Kong. (Doctoral dissertation).

Cui, Can & Yulin Yuan. 2020. 物動詞施事的隱現問題研究 [A study on the emergent problems of transitive verb agents in existential sentences]. 當代語言學 *Contemporary Linguistics* 3. 381–394.

Dalrymple, Mary. 2001. *Lexical Functional Grammar* (Syntax and Semantics 34). New York: Academic Press. DOI: 10.1163/9781849500104.

Dalrymple, Mary & Helge Lødrup. 2000. The grammatical functions of complement clauses. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '00 conference*, 104–121. Stanford: CSLI Publications.

Dong, Hongyuan. 2016. An LFG analysis of pronominal binding in Mandarin Chinese. In *Proceedings of the Linguistic Society of America*, vol. 1. DOI: 10.3765/plsa.v1i0.3646.

Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67. 547–619. DOI: 10.2307/415037.

Dryer, Matthew S. 2005. Order of subject, object, and verb. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*. Oxford: Oxford University Press.

Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2020. *Ethnologue: Languages of the world*. 23rd edn. Dallas: SIL International.

Fang, Ji & Tracy Holloway King. 2007. An LFG Chinese grammar for machine use. In Tracy Holloway King & Emily M. Bender (eds.), *Proceedings of the GEAF07 workshop*, 144–160. Stanford: CSLI Publications. http://csli-publications.stanford.edu/GEAF/2007/geaf07-toc.html.

Feng, Shengli. 1995. Guanyue lilun yu hanyu de beidong ju [GB theory and passive sentences in Chinese]. *Zhongguo Yuyanxue Luncong [Studies in Chinese Linguistics]* 1. 1–28.

Feng, Zhiwei. 2004. Lfg 中從詞彙結構到功能結構的轉換 [Transfer from c-structure to f-structure in LFG]. 語言文字應用 *Applied Linguistics* 4. 105–112.

Fu, Cheng-de. 1993. 詞匯映射理論 [A review on Lexical Mapping Theory]. 當代語言學 *Contemporary Linguistics* 1. 18–25.

Fu, Ai-Ping. 1990a. 詞匯功能語法 (上) [Lexical Functional Grammar I]. 當代語言學 *Contemporary Linguistics* 1. 39–44.

Fu, Ai-Ping. 1990b. 詞匯功能語法 (下) [Lexical Functional Grammar II]. 當代語言學 *Contemporary Linguistics* 2. 32–39.

Gao, Ming-Le & Fang Li. 2009. 詞匯功能語法(LFG)解析 [Analysis of Lexical Functional Grammar]. 語文學刊 *Journal of Language and Literature Studies* 1. 1–6.

Gu, Yang. 1992. On the locative existential construction in Chinese. In Dawn Bates (ed.), *Proceedings of the 10th West Coast Conference on Formal Linguistics*, 183–195. Stanford: CSLI Publications.

Gu, Yang. 1997. 關於存現 結構的理論探討 [Theoretical discussion on the existential construction]. 現代外語 *Linguistics and Applied Linguistics* 3.

Guo, Yuqing. 2009. *Treebank-based acquisition of Chinese LFG resources for parsing and generation*. Dublin: School of Computing, Dublin City University. (Doctoral dissertation).

Guo, Yuqing, Haifeng Wang & Josef van Genabith. 2008. Accurate and robust LFG-based generation for Chinese. In *Proceedings of the Fifth International Natural Language Generation Conference*, 86–94. DOI: 10.3115/1708322.1708339.

Handel, Zev. 2015. The classification of Chinese: Sinitic (The Chinese language family). In William S-Y. Wang & Chaofen Sun (eds.), *The Oxford handbook of Chinese linguistics*, 34–44. Oxford: Oxford University Press.

Her, One-Soon. 1989. An LFG account for Chinese BEI sentences. *Journal of the Chinese Language Teachers Association* 24(3). 67–89.

Her, One-Soon. 1990. *Grammatical functions and verb subcategorization in Mandarin*. Mānoa: University of Hawaii. (Doctoral dissertation).

Her, One-Soon. 1991. Topic as a grammatical function in Chinese. *Lingua* 84(1). 1–23. DOI: 10.1016/0024-3841(91)90011-s.

Her, One-Soon. 1995. Generating function words in machine translation. *Journal of the National Chengchi University* 70. 25–46.

Her, One-Soon. 1996. Variation of the VO construction in Chinese: A synchronic account. *Linguistics* 34. 733–751. DOI: 10.1515/ling.1996.34.4.733.

Her, One-Soon. 1997. Interaction and explanation: The case of variation in Chinese VO construction. *Journal of Chinese Linguistics* 25. 146–157, 159–165.

Her, One-Soon. 2004. Argument-function linking in resultatives. *Concentric: Studies in Linguistics* 20(2). 1–34.

Her, One-Soon. 2006a. Justifying part-of-speech assignments for Mandarin *gei*. *Lingua* 116. 1274–1302.

Her, One-Soon. 2006b. Optimality-theoretic Lexical Mapping Theory: A case study of locative inversion. *International Journal of Technology and Human Interaction (Special Issue: Human Language Technology in Multilingual Perspectives)* 2(1). 67–94. DOI: 10.4018/978-1-87828-991-9.ch166.

Her, One-Soon. 2007. Argument-function mismatches in Mandarin resultatives: A lexical mapping account. *Lingua* 117(1). 221–246. DOI: 10.1016/j.lingua.2006.01.002.

Her, One-Soon. 2008. *Grammatical functions and verb subcategorization in Mandarin Chinese*. Revised. Taipei: Crane Publishing.

Her, One-Soon. 2009. Unifying the long passive and the short passive: On the *Bei* construction in Taiwan Mandarin. *Language and Linguistics* 10(3). 421–470.

Her, One-Soon. 2010. *Interaction and variation in the Chinese VO construction*. Revised. Taipei: Crane Publishing.

Her, One-Soon. 2012a. Distinguishing classifiers and measure words: A mathematical perspective and implications. *Lingua* 122(14). 1688–1691. DOI: 10.1016/j.lingua.2012.08.012.

Her, One-Soon. 2012b. Structure of classifiers and measure words: A Lexical Functional account. *Language and Linguistics* 3(6). 2–25.

Her, One-Soon. 2017. Deriving classifier word order typology, or Greenberg's Universal 20a and Universal 20. *Linguistics* 55(2). 265–303. DOI: 10.1515/ling-2016-0044.

Her, One-Soon, Dan Higinbotham & Joseph Pentheroudakis. 1991. An LFG-based machine translation system. *Computer Processing of Chinese and Oriental Languages* 5(3–4). 285–297.

Her, One-Soon & Hui-Chin Tsai. 2020. Left is right, right is not: On the constituency of the classifier phrase in Chinese. *Language and Linguistics* 21(1). 1–32.

Hsieh, Shu-Kai, Jia-Fei Hong & Chu-Ren Huang. 2022. The extreme poverty of affixation in Chinese: Rarely derivational and hardly affixational. In Chu-Ren Huang, Yen-Hwei Lin, I-Hsuan Chen & Yu-Yin Hsu (eds.), *Cambridge Handbook of Chinese Linguistics*. Cambridge, UK: Cambridge University Press.

Hsueh, Frank F. S. 1989. The structural meaning of *Ba* and *Bei* constructions in Mandarin Chinese. In Frank F. S. Hsueh & Timothy Light (eds.), *Functionalism and Chinese grammar*, 95–125. Chinese Language & Teachers Association.

Huang, C.-T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).

Huang, C.-T. James. 1999. Chinese passives in comparative perspective. *Tsing Hua Journal of Chinese Studies* 29(4). 423–509.

Huang, C.-T. James, Yen-hui Audrey Li & Yafei Li. 2009. *The syntax of Chinese*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/cbo9781139166935.

Huang, Chu-Ren. 1985. On pseudo-possessive NPs in Mandarin Chinese. In Wayne Harbert & Sarah Fagan (eds.), *Cornell working papers in linguistics*, 129–148. Department of Modern Languages & Linguistics, Cornell University.

Huang, Chu-Ren. 1986. Possessive objects — A case for explicit reference of grammatical relations. In *The second biennial conference on Relational Grammar and grammatical relations*. Ohio State University.

Huang, Chu-Ren. 1987. *Mandarin Chinese NP* de*: A comparative study of current grammatical theories*. Ithaca: Cornell University. (Doctoral dissertation).

Huang, Chu-Ren. 1988. 再析國語「領屬主語」結構—概化詞組結構語法(GPSG)及詞彙功能語法(LFG)之比較研究 [A reanalysis of Mandarin Chinese 'possessive subjects'—a comparative study of GPSG and LFG]. 漢學研究 *Chinese Studies* 6(2). 109–134.

Huang, Chu-Ren. 1989a. Subcategorized TOPICs in Mandarin Chinese. Paper presented at the 1989 CLTA Annual Meeting, Boston, MA.

Huang, Chu-Ren. 1989b. 詞彙功能語法簡介 [Introduction to Lexical Functional Grammar]. In *Proceedings of ROCLING II*, 73–94.

Huang, Chu-Ren. 1990. A unification-based LFG analysis of lexical discontinuity. *Linguistics* 28(2). 263–307. DOI: 10.1515/ling.1990.28.2.263.

Huang, Chu-Ren. 1991. Adjectival reduplication in Southern Min – A study of morpholexical rules with syntactic effects. In *Chinese languages and linguistics, Volume I: Chinese dialects*, 407–422. Taipei: Academia Sinica.

Huang, Chu-Ren. 1992. Certainty in functional uncertainty. *Journal of Chinese Linguistics* 20(2). 247–287.

Huang, Chu-Ren. 1993a. Mandarin Chinese and the Lexical Mapping theory: A study of the interaction of morphology and argument changing. *Bulletin of the Institute of History and Philology* 62(2). 337–388.

Huang, Chu-Ren. 1993b. Reverse long-distance dependency and functional uncertainty: The interpretation of Mandarin questions. In Chungmin Lee & Boemmo Kang (eds.), *Language. information, and computation*, 111–120. Seoul: Thaehaksa.

Huang, Chu-Ren. 2015. Notes on Chinese grammar and ontology: The endurant/perdurant dichotomy and Mandarin DM compounds. *Lingua Sinica* 1(1). DOI: 10.1186/s40655-015-0004-6.

Huang, Chu-Ren & Kathleen Ahrens. 1999. The function and category of *gei* in Mandarin ditransitive constructions. *Journal of Chinese Linguistics* 27(2). 1–26.

Huang, Chu-Ren & Kathleen Ahrens. 2000. The module-attribute representation of verbal semantics. In *Proceedings of the 14th Pacific Asia Conference on Language, Information and Computation*, 109–120.

Huang, Chu-Ren & Kathleen Ahrens. 2003. Individuals, kinds and events: Classifier coercion of nouns. *Language Sciences* 25(4). 353–373. DOI: /10.1016/S0388-0001(02)00021-9.

Huang, Chu-Ren, Li-Ping Chang, Kathleen Ahrens & Chao-Ran Chen. 1999. 詞彙語意和句式語意的互動關係. [The interaction of lexical semantics and constructional meaning]. In Yuen-Mei Yin, I-li Yang & Hui-chen Chan (eds.), *Chinese language and linguistics V: Interactions in language*, 413–438. Taipei: Academia Sinica.

Huang, Chu-Ren & Fu-Wen Lin. 1992. Composite event structures and complex predicates: A template-based approach to argument selection. In Laurel S. Stvan et al. (eds.), *FLSM III: Papers from the third annual meeting of the Formal Linguistics Society of Mid-America*, 90–101. Bloomington: IULC Publications.

Huang, Chu-Ren, Yen-Hwei Lin, I-Hsuan Chen & Yu-Yin Hsu (eds.). 2022. *Cambridge handbook of Chinese linguistics*. Cambridge, UK: Cambridge University Press.

Huang, Chu-Ren & Louis Mangione. 1985. A reanalysis of *de*: Adjuncts and subordinate clauses. In Jeffrey Goldberg, Susannah MacKaye & Michael Wescoat (eds.), *Proceedings of the 4th West Coast Conference on Formal Linguistics*, 80–91. Stanford: Stanford Linguistics Association.

Huang, Chu-Ren & Dingxu Shi (eds.). 2016. *A reference grammar of Chinese*. Cambridge, UK: Cambridge University Press.

Huang, Hui-Ting & One-Soon Her. 1998. Mandarin locative inversion and relation-changing rules. In Shuan-fan Huang (ed.), *Selected papers from the second international symposium on languages in Taiwan*, 287–304. Taipei: Crane Publishing.

Jiang, Li, Peter Jenks & Jing Jin. 2022. The syntax of classifiers in Mandarin Chinese. In Chu-Ren Huang, Yen-Hwei Lin, I-Hsuan Chen & Yu-Yin Hsu (eds.), *Cambridge Handbook of Chinese Linguistics*. Cambridge, UK: Cambridge University Press.

Jiang, Menghan & Chu-Ren Huang. 2018. Transitivity variations in Mandarin VO compounds: A comparable corpus-based approach]. In Yunfang Wu, Jia-Fei Hong & Qi Su (eds.), *Workshop on Chinese lexical semantics*, 564–575.

Jiang, Menghan, Nayalia Klyueva, Hongzhi Xu & Chu-Ren Huang. 2018. Annotating Chinese light verb constructions according to PARSEME guidelines. In Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis & Takenobu Tokunaga (eds.), *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)*, 2482–2486. European Language Resources Association (ELRA).

Kaplan, Ronald M. & Annie Zaenen. 1989. Long-distance dependencies, constituent structure, and functional uncertainty. In Mark Baltin & Anthony Kroch (eds.), *Alternative conceptions of phrase structure*, 17–42. Chicago: University of Chicago Press. Reprinted in **DKMZ:FormalIssues**.

Karttunen, Lauri. 1989. Radical lexicalism. In Mark Baltin & Anthony Kroch (eds.), *Alternative conceptions of phrase structure*, 43–65. Chicago: University of Chicago Press.

Kit, Chunyu. 1992. Parsing Chinese *Ba3* and *Bei4* constructions - An LFG approach. In *Proceedings of 3rd International Conference on Chinese Information Processing*. Beijing.

Kit, Chunyu. 1993a. *A Chinese Lexical-Functional Grammar for sentence parsing*. Hong Kong: City Polytechnic of Hong Kong. (MA thesis).

Kit, Chunyu. 1993b. Description of Chinese intransitive verbs and adjuncts within the LFG formalism. In Chu-Ren Huang (ed.), *Proceedings of the 1st Pacific Asia conference on Formal and Computational Linguistics (PACFoCol I)*.

Kit, Chunyu. 1998. *Ba* and *Bei* as multi-valence prepositions in Chinese - an LFG perspective. *Studia Linguistica Sinica* 1. 1–21.

Kit, Chunyu & Jonathan J. Webster. 1992. Parsing Chinese pivotal constructions in an LFG approach. In *Proceedings of 1992 International Conference on Computer Processing of Chinese and Oriental Languages*. Florida.

Kroeger, Paul R. 2004. *Analyzing syntax: A lexical-functional approach*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/cbo9780511801693.

Lam, Chit Fung. 2020. Functional and anaphoric control in Mandarin Chinese. In Miriam Butt & Ida Toivonen (eds.), *Proceedings of the LFG '20 conference*. Stanford: CSLI Publications.

Lam, Olivia S.-C. 2008. *Object functions and the syntax of double object constructions in Lexical Functional Grammar*. Oxford: University of Oxford. (D.Phil. Thesis).

Lam, Olivia S.-C., One-Soon Her, Jing Chen & Sophia Y.-M. Lee. Forthcoming. LFG and Sinitic languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Lee, On-Man. 2003. *The subject function in Cantonese*. Hong Kong: University of Hong Kong. (MA thesis).

Lee, Yat-Mei. 2002. *Complement functions in Cantonese: A Lexical-Functional Grammar approach*. Hong Kong: University of Hong Kong. (MA thesis).

Lefebvre, Claire. 2011. *Creoles, their substrates, and language typology*. Amsterdam: John Benjamins.

Li, Charles N. & Sandra A. Thompson. 1974. Historical change of word order: A case study of Chinese and its implications. In *Historical linguistics I: Syntax*, 199–218. Amsterdam: North Holland Publishing Company.

Li, Charles N. & Sandra A. Thompson. 1976. Subject and topic: A new typology. In Charles N. Li (ed.), *Subject and topic*, 457–489. New York: Academic Press.

Li, Charles N. & Sandra A. Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley/Los Angeles: University of California Press.

Li, Yafei. 1990. On V-V compounds in Chinese. *Natural Language & Linguistic Theory* 8. 177–207. DOI: 10.1007/bf00208523.

Li, Yafei. 1995. The thematic hierarchy and causativity. *Natural Language & Linguistic Theory* 13(2). 255–282. DOI: 10.1007/bf00992783.

Liao, Wei-tai Gwen. 2010. *An LFG account of empty pronouns in Mandarin Chinese*. Oxford: University of Oxford. (D.Phil. Thesis).

Light, Timothy. 1979. Word order and word order change in Mandarin Chinese. *Journal of Chinese Linguistics* 7. 149–180. http://www.jstor.org/stable/23752920.

Liu, Feng-hsi. 2006. Dative constructions in Chinese. *Language and Linguistics* 7(4). 863–904.

Liu, Feng-hsi. 2022. SVO as the canonical word order in modern Chinese. In Chu-Ren Huang, Yen-Hwei Lin, I-Hsuan Chen & Yu-Yin Hsu (eds.), *Cambridge Handbook of Chinese Linguistics*. Cambridge, UK: Cambridge University Press.

LSHK. 2002. 粵語拼音字表 *[Guide to LSHK Cantonese romanization of Chinese characters]*. 2nd edn. Hong Kong: Linguistic Society of Hong Kong.

Lü, Shuxiang. 1980. 現代漢語八百詞 *[Eight hundred words of modern Chinese language]*. Beijing: 商務印書館 [The Commercial Press].

Lui, Wilson. 2020. Locative inversion in Cantonese. In Miriam Butt & Ida Toivonen (eds.), *Proceedings of the LFG '20 conference*, 250–267. Stanford: CSLI Publications.

Luke, Kang Kwong, Adams Bodomo & Owen Nancarrow. 2001. The subject condition in Cantonese. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '01 conference*, 359–379. Stanford: CSLI Publications.

McCawley, James. 1992. Justifying part-of-speech assignments in Mandarin Chinese. *Journal of Chinese Linguistics* 20(2). 211–246.

Mei, Kuang. 1980. Is modern Chinese really a SOV language? *Cahiers de linguistique-Asie Orientale* 7. 23–45. DOI: 10.3406/clao.1980.1071.

Mo, R. 1990. *On Chinese subcategorized topics*. Fu-jen Catholic University. (MA thesis).

Packard, Jerome. 2000. *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/cbo9780511486821.

Pan, Haihua. 1996. Imperfective aspect *zhe*, agent deletion, and locative inversion in Mandarin Chinese. *Natural Language & Linguistic Theory* 14(2). 409–432. DOI: 10.1007/bf00133688.

Pan, Haihua. 1997. 詞彙映射理論在漢語句法研究的應用 [Lexical mapping theory and its application in Chinese]. *Modern Foreign Languages* 4. 3–18.

Pan, Haihua & Jianhua Hu. 2008. A semantic-pragmatic interface account of (dangling) topics in Mandarin Chinese. *Journal of Pragmatics* 40(11). 1966–1981. DOI: 10.1016/j.pragma.2008.03.005.

Pan, Yanhong. 2010. *On the verb phrase in Qinzhou Zhuang: An LFG analysis of serial verb constructions*. Hong Kong: University of Hong Kong. (Doctoral dissertation).

Shiu, Yu-Ling & Chu-Ren Huang. 1989. A unification-based approach to Mandarin questions. *Journal of Information Science and Engineering* 5(4). 395–420.

Shyu, Shu-Ing. 2016. Information structure. In Chu-Ren Huang & Dingxu Shi (eds.), *A reference grammar of Chinese*, 518–576. Cambridge, UK: Cambridge University Press.

Sun, Chaofen & Talmy Givón. 1985. On the so-called SOV word order in Mandarin Chinese: A quantified text study and its implications. *Language* 61. 329–351. DOI: 10.2307/414148.

Sun, Maosong. 2001. LFG for Chinese: Issues of representation and computation. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '01 conference*, 394–404. Stanford: CSLI Publications.

Tai, James. 1973. Chinese as a SOV language. In *Papers from the 9th regional meeting of the Chicago Linguistics Society*, 659–671. Chicago.

Tan, Fu. 1991. *Notion of subject in Chinese*. Stanford: Stanford University. (Doctoral dissertation).

Tao, Hongyin. 2022. Scalar pseudo-cleft constructions in Mandarin conversation: A multimodal approach. *Lingua* 266. DOI: 10.1016/j.lingua.2021.103202.

Tsai, Wei-Tien Dylan. 2022. Cambridge handbook of Chinese linguistics. In Chu-Ren Huang, Yen-Hwei Lin, I-Hsuan Chen & Yu-Yin Hsu (eds.), 616–634. Cambridge, UK: Cambridge University Press.

Tsao, Feng-fu. 1996. On verb classification in Chinese. *Journal of Chinese Linguistics* 24(1). 138–191.

Wang, William S.-Y. & Chaofen Sun. 2015. *The Oxford handbook of Chinese linguistics*. Oxford: Oxford University Press.

Webster, Jonathan J. & Chunyu Kit. 1995. Computational analysis of Chinese and English texts with the functional semantic processor and the C-LFG parser. *Literary and Linguistic Computing* 10(3). 203–211. DOI: 10.1093/llc/10.3.203.

Wei, Ai-yun. 2014. 國內外詞彙功能語法研究述評 [A review of domestic and international research on LFG]. 廣西師範大學學報: 哲學社會科學版 *Journal of Guangxi Normal University (Philosophy and Social Science Edition)* 3. 187–194.

Wurm, Stephen A. & Yongquan Liu (eds.). 1987. 中國語言地圖集 *[Language atlas of China]*. Hong Kong: Longman Group.

Xu, Jie & Sicong Dong. 2022. SOV as the canonical word order in modern Chinese. In Chu-Ren Huang, Yen-Hwei Lin, I-Hsuan Chen & Yu-Yin Hsu (eds.), *Cambridge Handbook of Chinese Linguistics*. Cambridge, UK: Cambridge University Press.

Xu, Liejiong & D. Terence Langendoen. 1985. Topic structures in Chinese. *Language* 61(1). 1–27. DOI: 10.2307/413419.

Yang, Fengrong. 2020. BEI-passive revisited: A constraint-based approach. In Miriam Butt & Ida Toivonen (eds.), *Proceedings of the LFG '20 conference*, 363–383. Stanford: CSLI Publications.

Yang, Ning & Geertje van Bergen. 2007. Scrambled objects and case marking in Mandarin Chinese. *Lingua* 117(9). 1617–1635. DOI: 10.1016/j.lingua.2006.06.009.

Yao, Yao & Feng-hsi Liu. 2010. A working report on modeling dative variation in Mandarin Chinese. In *COLING 2010: The 23rd International Conference on Computational Linguistics*, 1236–1244. Association for Computational Linguistics.

Yue, Anne O. 2003. Chinese dialects: Grammar. In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan languages*, 84–124. London: Routledge.

Zhang, Bojiang. 1999. 現代漢語的雙及物結構式 [The ditransitive construction in modern Mandarin]. 中國語文 Zhongguoyuwen 270. 175–184.

Zhu, De-xi. 1982. 語法講義 *[Lecture notes on grammar]*. Beijing: Commercial Press.

Chapter 6

LFG and Role and Reference Grammar

Delia Bentley
The University of Manchester

Nigel Vincent
The University of Manchester

LFG and Role and Reference Grammar have in common the goals of developing a formal model for the grammars of natural languages that both accommodates typological diversity and avoids syntax-centred derivationality. That said, the two frameworks differ in their choice of conceptual primitives and in the way the different components interact. In the present chapter we explore those differences in particular with respect to core sentence structure, information structure, cross-linguistic patterns and variety, and diachrony.

1 Historical Context

Both LFG and RRG emerged in the 1970's and 1980's in the context of the general reconsideration of possible models of grammar that took place at that time. These developments were driven in part by a concern to rethink the best way to capture the interaction between syntax, semantics and pragmatics, in part by a desire to reflect the typological diversity of natural languages and avoid a bias towards the sorts of structure found in 'standard average European', and in part by considerations of psychological plausibility and computational tractability. At the same time, the two frameworks differ in the relative priority to be assigned to these different lines of argument and evidence.

In the case of RRG the two principal motivating questions were the following: (a) What would a linguistic theory look like if it were based on the analysis of Lakhota, Tagalog and Dyirbal, instead of English? (b) How can the interplay

Delia Bentley & Nigel Vincent. Forthcoming. LFG and Role and Reference Grammar. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

of syntax, semantics and pragmatics in different grammatical systems best be captured and explained? Van Valin 2009: 704. Constraining the framework are, therefore, not only the classic Chomskyan criteria of descriptive and explanatory adequacy (on which see Rizzi 2016), but also those of typological and psychological adequacy, since in the words of Austin & Bresnan (1996: 263) 'theoretical economy and explanatory elegance are unreliable guides to truth'. Typological adequacy requires that the theory should grasp commonalities between different languages without attributing to a given language any features for which that language provides no evidence. Psychological adequacy, as formulated by Dik (1991: 248), states that a theory should be compatible with the results of psycholinguistic research on the acquisition, processing, production, interpretation and memorization of linguistic expressions. This is not to say that there are no postulated universal principles either in LFG or RRG, but rather that within neither framework is there the presumption of an innate, syntactically defined U(niversal) G(rammar).

This concern for psycholinguistic plausibility was shared with LFG, as discussed for example by Bresnan & Kaplan (1982), where it was linked to issues about the length and complexity of syntactic derivations within the transformational approach. Whereas at that time generative syntax was — and indeed still is — built on an exclusively categorial set of primitives, LFG and RRG in their different ways sought to explore in addition the use of relational concepts. Influential here had been, on the one hand, Relational Grammar with its definition of structures in terms of changing grammatical functions like subject and object, albeit while still retaining a derivational approach, and, on the other hand, Fillmorean Case Grammar with its set of semantically defined roles like agent and patient. For LFG this led to a much reduced, monostratal categorial component (c-structure) linked to but not derived from a set of grammatical relations (f-structure). RRG, by contrast, goes a step further and in addition sets aside notions like subject and object as also being in danger of biassing the system towards particular types and families of languages and opting instead for a core set of semantically defined relations. Despite these differences, RRG and LFG have in common the fact that, once the analytical burden is shared between categories and relations, grammatical structures are no longer required to respect the principles of endocentricity and binary branching which have become key parts of current Minimalist, cartographic and nanosyntactic approaches. A sentence can be represented simply as S rather than needing to be CP, TP or the like and if a language does not provide ready evidence of configurational structure, none needs to be imposed (Austin & Bresnan 1996; Van Valin & LaPolla 1997b: Chapter 2).

We move now to an overview of RRG (Section 2) before returning to a more detailed comparison of the two frameworks (Section 3) and consideration of the way they deal with issues involving language change, processing and acquisition (Sections 4, 5 and 6).

2 RRG: An overview

For RRG, grammar is a system in a traditional structuralist sense. However, RRG is not only interested in the syntagmatic and paradigmatic relations that characterize syntax, but also in the combinatorial relations between units of meaning within and out of context. This framework is, thus, like LFG, a parallel architecture theory (Jackendoff 2002: Chapter 5), which relies on three independent, albeit interacting, levels of analysis: discourse, lexical semantics and syntax.[1] Much of what other syntactic frameworks would explain in terms of syntactic derivation or movement is captured in RRG in terms of the mapping of these three dimensions. This reflects the assumption that grammatical structure can only be understood and explained with reference to the expressive and communicative functions of language.

Since it seeks explanation outside of the boundaries of syntax, RRG could thus be thought to lie on the functional side of the formalist-functional divide in theories of language (Butler 2005, 2006; Mairal Usón et al. 2012), and indeed Newmeyer (1998: 14-16) cites it as an example of what he calls 'external functionalism', adducing the description by Van Valin (1993: 1) of RRG as a 'structuralist-functionalist theory of grammar'. However, a preference for the explanation of linguistic phenomena in terms of meaning and external context by no means implies an absence of a formal notation. And indeed within RRG each of the levels of analysis is conceived of in terms of an articulated formalism and there are explicit constraints on the interaction of the three levels. In addition, in the last ten to fifteen years, an increasing number of scholars have attempted to apply RRG to language processing, both in the computational and the neurolinguistic domain. Such attempts have resulted in the development of new formalisms, which use the RRG framework as their basis (see Section 5 below).

The basic architecture of RRG is illustrated in Figure 1. While the two arrows in the middle show the bidirectionality of the semantics-syntax linking, the position of discourse-pragmatics with respect to this mapping indicates that discourse can

[1]A striking comparison in this connection is Sadock's independent realisation that a language like Greenlandic calls for a parallel or 'modular' architecture (Sadock 1991: ix–xi), which in turn led to his own model of Autolexical Syntax.

be relevant at every step in the linking (Bentley to appear and Section 3.5 below). Specifically, discourse-related meaning (for example, the distinction between the information that has already been given and the new information that is provided with the utterance) is not only expressed syntactically, but also in prosody, morphology and even in lexical choices. In fact, the encoding of discourse-related meaning in syntax varies across languages in important ways and this variation has been the object of much research in RRG (see among others Van Valin 1999, Shimojo 1995, 2008, 2009, 2010, 2011, Bentley 2008).

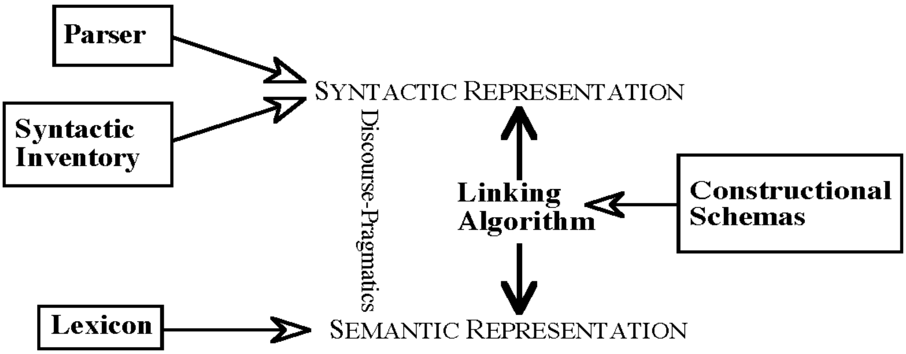


Figure 1: Organization of RRG (Van Valin 2005: 134)

We return to other properties of the linking at the end of Section 2.

2.1 The structure of the sentence and of reference phrases

There is a single syntactic representation for a sentence, which corresponds to the surface form of the sentence and appears in the Constituent Projection. As noted above, there is no requirement that the structure of the clause should be binary branching; the syntax of the sentence must be adequately represented in configurational and non-configurational, dependent-marking and head-marking languages alike. In clausal structure, a distinction is drawn between the semantically motivated positions, which are assumed to be universal, and other positions, which tend to be associated with particular pragmatic roles and are not universal. Together the two types of position form the Layered Structure of the Clause (see Figure 2).

There is no verb phrase in the Layered Structure of the Clause because not all languages offer evidence for it (for comparable considerations in LFG see Börjars

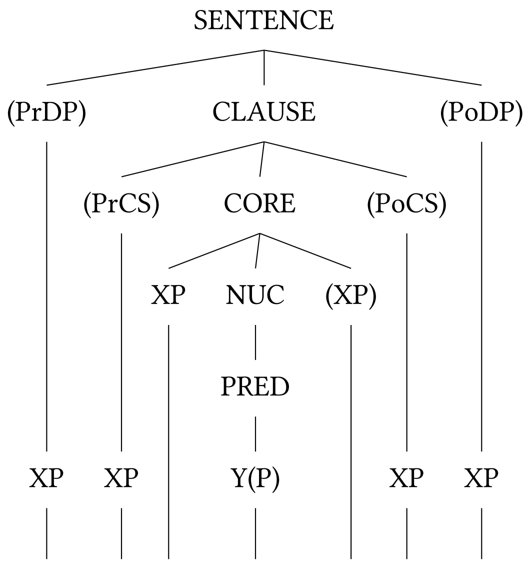


Figure 2: The Layered Structure of the Clause (from Van Valin to appear(b))

et al. 2019: 5-6). The Nucleus hosts the predicate, while the arguments drawn from the semantic representation of the predicate, called core arguments, figure within the Core and are labelled RPs (Reference Phrases).[2] No phonologically null elements are allowed in RRG syntax.[3] Neither the Nucleus nor the RP nodes are restricted to any particular lexical category, given that in some languages, such as Nootka and Tagalog, expressions that are verbs in categorial terms can have a referential function in the clause, in which case they behave as arguments, while nouns can have a predicative function (Van Valin 2008: 170).[4] In English too

[2]Within core arguments, RRG distinguishes between direct core arguments, which are unmarked or marked by case alone, and oblique ones, which are adpositionally marked.

[3]Genuine zero anaphora, i.e., the complete failure of expression of an argument, whether as a pronoun or in inflection, is dealt with in a system of direct mapping from discourse to the semantic representation of the clause, and vice versa, with the argument being represented in both of these domains, but not in syntax (see Section 3.5). Zero morphemes are, however, admitted in RRG in morphological paradigms, the key difference between these and phonologically null syntactic elements being that the latter type of element is redundant, on the assumption that the linking can occur directly from the semantic representation to discourse.

[4]This is not to say that nouns and noun phrases have no status in RRG. On the contrary, nouns and verbs are taken to be universal lexical categories, by contrast with adjectives, which are not found in all languages.

the predicate in the Nucleus can be an adjective, a noun phrase, or a prepositional phrase, although a verb is needed for the proper formation of the Nucleus of the clause. The Nucleus and the Core are taken to be universal positions because all languages predicate and refer. Any adjuncts that modify the nucleus, or the core, or indeed any of the more external syntactic layers figure in a Periphery as M(odifier) P(hrases). Every syntactic layer (Clause, Core, Nucleus) can have its own Periphery.

Both the core-internal positions and the peripheries of the various layers of the clause can host constituents with particular discourse roles. To give but one example, to the extent that they are overt, topical subjects normally occur in the core-initial pre-nuclear position in SVO languages. However, these positions are not defined in pragmatic terms, but rather in terms of the referential and predicative functions of language. The more external positions, instead, tend to be associated with pragmatically salient functions. The Pre- and Post-Core Slot normally host foci, although there can be language-specific restrictions on the kinds of foci that they admit. In a large number of languages the Pre-Core Slot hosts pre-verbal *wh*-words and the same position has been claimed to be involved in contrastive focus fronting in some Romance languages (Bentley 2008). The Pre-Core Slot hosts topics, as well as foci, in languages with a V2 constraint on word order (Diedrichsen 2008). The Post-Core Slot is the position of secondary foci which non-canonically occur in post-verbal position in Japanese, a verb-final language (Shimojo 1995). The Pre-Detached Position (formerly called Left-Detached Position) is the position of detached topics and can iterate, thus allowing the utterance to have several topics, while the Pre- and Post-Core Slot cannot be repeated.[5] The Post-Detached Position (formerly called Right-Detached Position) hosts afterthoughts or topics.[6] The pragmatically salient positions are not universal: the languages that provide no evidence for these positions are not assumed to have them. The building blocks of the Layered Structure of the Clause are the building blocks of complex predicates and clauses, as will be explained in Section 3.4.

The structure of the RP and of adpositional phrases is built following the same principles as the structure of the clause (see Cortéz-Rodriguez to appear, Ibañez

[5]This raises the question of the position of initial sequences of *wh*-words in languages which allow them, for example Bulgarian, an issue which to our knowledge has only been addressed from an RRG perspective by Eschenberg (1999).

[6]The reason for the relabelling of the Left- and Right-Detached positions is that these names reflect a bias towards western languages, which are written from left to right. The problem does not arise with Pre-Detached and Post-Detached, which reflect the before and after dimensions of speech.

Cerda to appear.). Thus, RPs have their own Constituent Projection, with Nucleus and Core, and their respective peripheries. RPs also have their Operator Projection, which defines the scope of the functional categories of definiteness, deixis, quantification, and number.

In RRG the functional markers of closed-class grammatical categories such as aspect, modality, tense and illocutionary force are not mapped to the Constituent Projection, but rather to the Operator Projection, and hence the framework does not incorporate an inventory of functional heads. The Operator Projection is the mirror image of the Constituent Projection because RRG assumes that the order of the morphemes that express grammatical categories is a function of their syntactic and semantic scope (Foley & Van Valin 1984; Bybee 1985). Thus, the Nucleus comes first, in the Operator Projection, as the domain of aspect, nuclear negation and directionals. Core negation and root modality have scope over the Core. Finally, status (epistemic modality), tense, evidentials, and illocutionary force have scope over the Clause.

The Constituent Projection is not built incrementally in the linking. Rather, the syntactic structure of the clause, and of the RPs and PPs contained in it, are drawn as templates from the syntactic inventory of the given language at the relevant stage in the linking. The selection of syntactic templates in the linking is governed by the *Syntactic Template Selection Principle* (Van Valin & LaPolla 1997b: 324, Van Valin 2005: 130) and by discourse considerations (Section 3.6), to which we return below. The syntactic inventory of a language comprises all the templates that are necessary to form grammatical sentences in that language. It reflects universal linearization principles concerning the position of the extra clausal positions shown in Figure 2, as well as the word order preferences of the language: primarily, its branching directionality, in the sense of Dryer (1992). Broad typological properties, such as head and dependent marking, and configurationality, also play a role in word order. Instead, the position of the operators in the clause largely depends on their semantic scope (see Figure 3). The syntactic inventory complements the lexical inventory as well as an inventory of constructional schemas, to which we shall also return.

2.2 Logical Structures, semantic roles and macroroles

The lexicon is an important component of grammar in RRG, since the semantic representation of the clause is based on the semantic representation of the verb and any other predicating elements figuring in it, for example, any predicative adpositional phrases. The semantic representation, or Logical Structure, of a verb

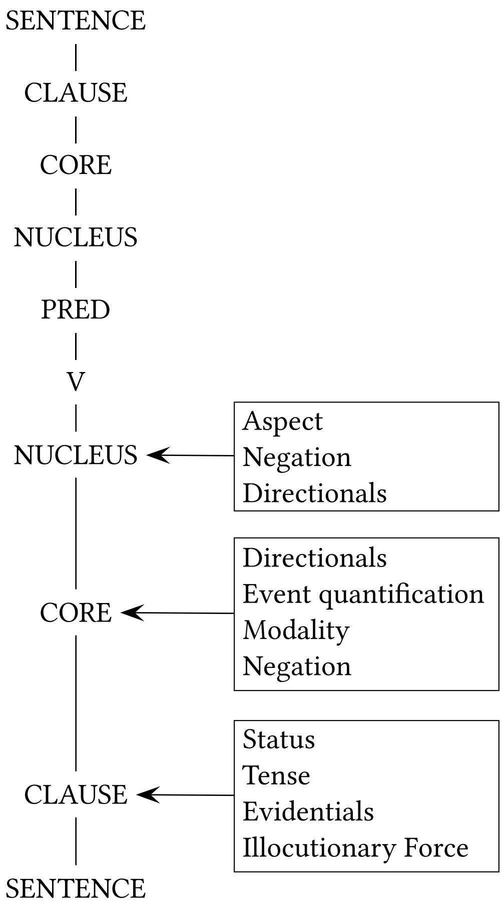


Figure 3: The operator projection in the layered structure of the clause

is based on a theory of lexical decomposition which relies on Vendler's (1967: 97-121) Aktionsart types *state*, *activity*, *achievement* and *accomplishment*, to which Van Valin & LaPolla add the distinction between plain and active accomplishment (see below) and Van Valin (2005: 32) adds the non-Vendlerian class of *semelfactives* (Smith 1997: 55-58). State and activity are the basic types upon which all the others are built. Both states and activities are [−telic] and [−punctual]. However, states describe static situations, whereas activities describe dynamic ones, that is, situations that involve change, albeit not of the type leading to a result state. We provide below the semantic representations of the states 'red' and 'know' and of the activities 'march' and 'sing'.

(1) States
  a. **be′**(x, [**red′**]) ‘red’
  b. **know′**(x, y) ‘know’

(2) Activities
  a. **do′**(x, [**march′**(x)]) ‘march’
  b. **do′**(x, [**sing′**(x, (y))]) ‘sing’

Predicates are presented in bold, followed by a prime, and English is the metalanguage used to represent them; **be′** figures in the Logical Structure of attributive, identificational and specificational states, alongside the constant identifying the given state. Instead, **do′** marks the Logical Structure of all activities.

Achievements and accomplishments are [+telic], which means that they describe change leading to the attainment of a result state. The former predicate type, being [+punctual], does not include a PROC(ess) component (cf. (3)), which instead characterises the latter (cf. (4)).[7] PROC and the other Logical Structure components in capital letters are operators, or markers of templatic facets of meaning, which combine with the constants representing the idiosyncratic meaning of individual lexical items. The RRG theory of lexical decomposition stands out from others in differentiating accomplishments from active accomplishments (cf. (5)). These are built on the basis of the logical structures of an activity plus an accomplishment. The process that is part of the accomplishment is simultaneous with the activity, and both are followed by the attainment of a result state (Van Valin 2018). Simultaneity is represented with the notation ∧, whereas the symbol & stands for “and then”.

(3) Achievements
  a. INGR **appear′**(x) ‘appear’
  b. INGR **be-at′**(x) ‘arrive’

(4) Accomplishments
  a. PROC INGR **dead′**(x) ‘die’
  b. PROC INGR **know′**(x, y) ‘learn’

(5) Active accomplishments

[7] See Bentley (2019) and Van Valin (to appear[b]) for proposals on the differentiation of quantized and non-quantized change in the Logical Structure of accomplishments.

a. **do′**(x, [**run′**(x)]) ∧ PROC **cover.path.distance′**(x, (y)) & INGR **be-at′(path.endpoint**, x) 'run to a location'
b. **do′**(x, [**write′**(x, y)]) ∧ PROC **create′**(y) & INGR **exist′**(y) 'write (tr.)'

Semelfactives (Smith 1997: 55–58) describe repeatable punctual events, which may be [+static] or [+dynamic] (cf. (6a) vs. (6b)), and do not lead to a result state, as testified by the absence of result state participles of these verbs in attributive function in the noun phrase.

(6) Semelfactives
a. SEML **see′**(x, y) 'glimpse'
b. SEML **do′**(x, [**cough′**(x)]) 'cough'

There are standard diagnostics to determine the Aktionsart of the predicate of a clause, based on Dowty's (1979) seminal work. No *a priori* assumption is made as to whether verbs describing comparable eventualities should belong to the same Aktionsart type across languages, although it is acknowledged that there are striking cross-linguistic similarities of this kind, whose rationale can be captured on the basis of a system of lexical decomposition like the RRG one.

The predicate types discussed above have causative counterparts, which in principle combine any logical structure α with any logical structure β by means of the operator CAUSE. The causal event may, however, remain unspecified, as is shown in (7b), which is built upon (7a).

(7) Accomplishment vs. causative accomplishment
a. PROC INGR **dead′**(x) 'die'
b. [**do′**(x, ∅)] CAUSE PROC INGR **dead′**(y) 'kill'

Traditional thematic role labels, like *theme* or *patient*, are mere mnemonics for the position which an argument occupies in Logical Structure as determined by applying the standard tests for the Aktionsart of the predicate. It is purely on the basis of its position that a core argument derives its thematic role (Jackendoff 1976; Van Valin & LaPolla 1997b: 82-138). There are five relevant positions.

(8) Semantic positions which are relevant to the linking

Arg of DO	1st arg of **do′**(x, ...)	1st arg of **pred′**(x,y)	2nd arg of **pred′**(x,y)	arg of state **pred′**(x)

The two leftmost positions in (8) are both positions for the argument of an activity: only the latter is usually found in the Logical Structure of an activity, unless agentivity is lexicalized (Van Valin & Wilkins 1996), in which case the first position (arg of DO) co-occurs with the second one (1st arg of **do′**(x, …)). The other positions are found in the Logical Structure of bivalent (1st and 2nd argument of **predicate**′(x,y)) and monovalent (argument of state **predicate′**(x)) states. As can be seen in (3) to (7) these positions combine with each other and with operators of cause, semelfactivity, process and change.

Importantly, the positions in (8) are not grammatically salient *per se*, but only to the extent that they determine which generalized semantic role, or macrorole, an argument is assigned in the linking. The relation between argument positions and macroroles is captured by the Actor-Undergoer Hierarchy in (9), while the macrole assignment principles are spelled out in (10).

(9) The Actor - Undergoer Hierarchy and its mapping onto argument positions (Van Valin 2005: 61)

ACTOR				UNDERGOER
⟶	⟶	⟶		
		⟵	⟵	⟵
Arg of DO	1st arg of **do′**(x,…)	1st arg of **pred′**(x, y)	2nd arg of **pred′**(x, y)	Arg of state **pred′**(x)

[‘→’ = increasing markedness of realization of argument as macrorole]

(10) Default Macrorole Assignment Principles (Van Valin 2005: 63)

a. Number: the number of macroroles a verb takes is less than or equal to the number of arguments in its logical structure.
   i. If a verb has two or more arguments in its logical structure, it will take two macroroles.
   ii. If a verb has one argument in its logical structure, it will take one macrorole.

b. Nature: for verbs which take one macrorole,
   i. If a verb has an activity predicate in its logical structure, the macrorole is actor.
   ii. If a verb has no activity in its logical structure, the macrorole is undergoer.

Actor and undergoer are the two primary arguments of transitive predications. Two-place verbs belonging to different Aktionsart types, say an active accomplishment (e.g., *write a book*) and a state (e.g., *know the answer*), are not differentiated in terms of macrorole assignment: both take an actor (the highest or leftmost argument in Logical Structure) and an undergoer (the lowest or rightmost argument in Logical Structure). There is, however, a fundamental asymmetry between the two macroroles, in that the highest core argument will always be the actor, whereas the lowest one is only the default choice for undergoer. Indeed, variable selection of the undergoer from the two lower arguments of three-place predicates is allowed in some languages. This is exemplified by English *present* in (11). In addition, two-place predicates may be intransitive, in which case this is specified in the lexicon, as exemplified with English *belong (to)* in (12).

(11) a. [**do′**(x, ∅)] CAUSE [INGR **have′**(y, z)]
b. x presents z to y
c. x presents y with z

(12) **have′**(x, y) [MR1] 'belong (to)'

In (11b) z is the undergoer, whereas in (11c) the undergoer is y. The actor is x in both cases. As for (12), [MR1] lexically specifies that this verb only takes one macrorole despite being bivalent. Finally, whether the only core argument of a one-place predicate is an actor or an undergoer is established by the principles in (10b).

An important claim of RRG is that no subcategorization requirements need to be specified for a verb, other than the argument positions in its Logical Structure and its transitivity, which is defined as the number of macroroles it takes. The prepositions that mark the oblique arguments required by some verbs (e.g., *load x with y, load y on x*) are argued to be predictable from general principles, for which we refer to Van Valin & LaPolla (1997b: 376-384).

Macrorole assignment plays a key role in the linking, allowing RRG to capture how syntactically different, but semantically comparable, structures are related. Thus, starting from the assumption that languages with nominative-accusative alignment select the actor, whereas languages with ergative-absolutive alignment select the undergoer, as the default privileged grammatical relation (Section 2.3), passive and antipassive are constructions with the marked macrorole selection as grammatical relation: undergoer in the passive and actor in the antipassive. We return below to the notion of subject, which is not considered to be a universal of grammar in RRG.

Macrorole assignment, or failure thereof, also captures the different syntax of verbs with the same number of arguments. Consider (13a) and (13b).

(13) Italian

a. Mario, la matematica, l' ha sempre amata.
Mario the maths.FSG ACC.CL.FSG has always love.PTCP.FSG

b. (A Mario), la matematica gli è sempre piaciuta.
to Mario the maths.FSG DAT.CL is always please.PTCP.FSG
'Mario, maths, he always loved/liked it.'

The contrast between nominative and dative experiencer verbs (e.g., Italian *amare* 'love' vs. *piacere* 'please, like') depends on whether both arguments are assigned a macrorole, with the result being a transitive structure, as testified by the accusative clitic and the perfect auxiliary 'have' in (13a), or the experiencer being denied macrorole status, in which case the structure has a single macrorole and is intransitive, as testified by the selection of a different auxiliary, 'be', and the dative clitic in (13b).

2.3 Grammatical relations

RRG rejects the traditional notions of subject and object as primitives or universals of syntactic theory. Following Durie's (1985; 1987) analysis of Acehnese, an Austronesian language, Van Valin & LaPolla (1997b: 255-260) claim that there are languages which group arguments in terms of their macrorole status without assigning them a syntactic function. In Acehnese, all actors are marked in the same way, as illustrated by the proclitic pronoun in (14a)-(14b), whereas undergoers are marked differently, as illustrated by the ungrammaticality of (14c) and the optional enclitic pronoun in its grammatical counterpart in (14d).[8]

(14) Acehnese (Van Valin & LaPolla 1997b: 255-256)

a. (Gopnan) geu-mat lôn.
(3SG) 3-hold 1SG
'(S)he holds me.'

b. Geu-jak (gopnyan).
3-go (3SG)
'(S)he goes.'

[8] The reader should note, on the one hand, that Acehnese is a head-marking language and, on the other, that the Logical Structure of the verb 'go', in this and other languages, includes an activity. Therefore, the macrorole assigned to the direct core argument is actor, following (10b-i).

c. *(Lôn) lôn-rhët.
   1sg 1sg-fall

d. Lôn rhët(-lôn).
   1sg fall-1sg
   'I fall.'

The contrast between (14a)-(14b) and (14c)-(14d) suggests that arguments are only grouped in terms of their macrorole, as is the case with active-vs.-inactive alignment, and there is no marking that defines a syntactic function. Acehnese also has no voice constructions, such as passive or antipassive, which follows from the absence of grammatical relations.

From this it also follows that if a grammatical relation is to be postulated for a given language or construction, evidence will be required of restricted neutralizations of semantic roles for grammatical purposes (see LaPolla to appear for an in-depth discussion of this point). Such neutralizations can be, and indeed often are, found at the level of specific constructions, although the well-known Indo-European languages tend to be consistent across constructions. With reference to the Acehnese examples in (14a)-(14d), the fact that the obligatory pre-verbal clitic only cross-references the actor indicates that this type of cross-referencing involves no such neutralization, but merely a restriction to actor. Contrastingly, the controller of person and number agreement on the English verb can be characterized as a restricted neutralization, specifically [A, S, d(erived)-S], because only the actor of a transitive (cf. (15a)), the actor or undergoer of an intransitive (cf. (15b)-(15c)) or the derived intransitive S of a passive (cf. (15d)) can control this kind of agreement. The undergoer of a transitive structure cannot (contrast (15a)-(15d) with (15e)).

(15) a. Mary$_i$ (A) has$_i$ eaten all the biscuits$_j$ (U).
  b. Mary$_i$ (Sa) has$_i$ eaten.
  c. Mary$_i$ (Su) has$_i$ fallen.
  d. All the biscuits$_j$ (d-S) were$_j$ eaten by Mary$_i$ (A).
  e. *Mary$_i$ (A) have$_j$ eaten all the biscuits$_j$ (U).

The fact that the grouping [A, S, d-S] is insensitive to the distinction between Sa and Su indicates that the control of person and number agreement on the English verb neutralizes the semantic role of the controller. The fact that the undergoer of a transitive (U) is banned from this syntactic function, and indeed a special voice construction, the passive, is needed for this argument to control agreement as a d-S, indicates that the neutralization under discussion is restricted.

RRG calls this kind of restricted neutralization a privileged syntactic argument (PSA).

Importantly, there are languages that provide no evidence for such restrictions. Thus, Mandarin Chinese (LaPolla 1990, 1993, 1995, to appear) has no conventionalized associations between syntactic position, agreement on the verb, case marking on the noun, etc. and particular semantic roles. The claim in RRG is, therefore, that Mandarin Chinese is a language which does not have any grammatical relations.

PSAs can have the syntactic functions of controller or pivot. The latter is the missing argument in a construction, whereas the controller is the argument that supplies its interpretation. Observe that the pivot of the English construction with *want* is defined as [A, S, d-S].

(16) a. Mary$_i$ [CONTROLLER] wants $\_\__i$ [PIVOT, A] to eat the biscuits.
  b. Mary$_i$ [CONTROLLER] wants $\_\__i$ [PIVOT, Sa] to eat.
  c. Mary$_i$ [CONTROLLER] wants $\_\__i$ [PIVOT, Su] to die.
  d. Mary [CONTROLLER] wants $\_\__i$ [PIVOT, d-S] to be loved.
  e. *Mary$_i$ [CONTROLLER] wants you to love $\_\__i$ [PIVOT, U].

Similar considerations are valid for the missing argument in conjunction reduction. This suggests that English is consistent in how it constrains the PSA across constructions. Nonetheless, there are English constructions in which different restrictions apply. For instance, the controller of the non-finite complementation with *persuade* is the undergoer, and cannot be the actor (cf. (17)).

(17) a. Mary$_j$ persuaded Paul$_i$ [CONTROLLER] $\_\__{i/*j}$ [PIVOT] to stay.
  b. Paul$_i$ [CONTROLLER] was persuaded by Mary$_j$ to $\_\__{i/*j}$ [PIVOT] to stay.

Given that there is a restriction in (17), but no neutralization, this is a case of semantic control, comparable to the control of the pre-verbal clitic in Acehnese.

While being comparable to English, in that they have restricted neutralizations of the kind described above, other languages define the PSA differently. Thus, Kalkatungu, an Australian aboriginal language (Blake 1979), provides evidence of the restricted neutralization [U, S, d-S], which defines ergative-absolutive alignment. The participial construction exemplified below illustrates this kind of PSA.

(18) Kalkatungu (Van Valin 2005: 97-98)

a. Ṯuaṭu pa-ij maṛapai-∅ icaij [iŋka-ʎ-iŋka-cin-∅].
snake.ERG that-ERG woman-ABS bite go-LNK-go-PTCP-ABS
'The snake bit the woman$_i$ [as __$_i$ was walking along].

b. [Jaṛikajan-ati-ɲin-tu] caa ŋa-ṯu ḻaji ∅
hungry-VBLZ-PTCP-ERG here 1SG-ERG kill 3SG.ABS
'[__$_i$ Being hungry] I$_i$ killed it.'

c. Kuntu caa ḷuŋa-ṉa ∅ [ŋa-ṯu ḻa-ɲin-ka-∅]
NEG here cry-PST 3SG.ABS 1SG-ERG hit-PTCP-SUFF-ABS
'He$_i$ didn't cry [when I hit __$_i$].'

d. *Nga-ṯu ṇaɲa macumpa-∅ [aṛi-ɲin-∅ kaṯir-∅]
1SG-ERG saw kangaroo-ABS eat-PTCP-ABS grass-ABS
'I saw the kangaroo$_i$ [__$_i$ eating grass].'

The pivot or missing argument of the Kalkatungu participial construction can be an intransitive S (SA in (18a) and SU in (18b)) or a transitive U (cf. (18c)), but it cannot be the actor of a transitive structure (A) (cf. (18d)). Therefore, there is a neutralization of semantic macroroles in this construction and this is restricted to S and U, leaving out A. In fact, if the verb in the participial construction is antipassivized, then the construction is grammatical.

(19) Kalkatungu (Van Valin 2005: 98)
Nga-ṯu ṇaɲa macumpa-∅ [aṛi-li-ɲin-∅ kaṯir-ku]
1SG-ERG saw kangaroo-ABS eat-ANTIP-PTCP-ABS grass-DAT
'I saw the kangaroo$_i$ [__$_i$ eating grass].'

The data in (19) indicate that d-S is also admitted in the Kalkatungu participial construction. The PSA of this structure is thus to be defined as [U, S, d-S].

It should further be noted that some languages do not have special voice constructions, in which case they may have the restricted neutralizations [A, S] or [U, S], although the latter is claimed to be very rare. In addition, in other languages, the PSA need not be a macrorole argument. We refer to Van Valin & LaPolla (1997b: 352-363) for relevant discussion.

The RRG conception of grammatical relations poses very strong constraints on the analysis of correspondences such as the ones that other frameworks conceive of as relations between active objects and passive subjects or between transitive objects and unaccusative subjects. Not only is it not possible to rely on movement or derivation, but the very construct of object is not available either. As was briefly mentioned above, the passive, as well as the antipassive, are captured in terms of the PSA selection hierarchy that is at work in the linking in a given language, or a given construction. Starting from the ranking of arguments in (20), which reflects the argument positions in Logical Structure (cf. (8) and (9)), the default PSA is selected in accordance of the two main principles in (21).[9]

(20) Arg of DO > 1st arg of **do′** > 1st arg of **pred′**(x,y) > 2nd arg of **pred′**(x,y) > arg of **pred′**(x)

(21) PSA Selection Principles
  a. Accusative construction: the default PSA is the highest-ranking direct core argument in terms of (20).
  b. Ergative construction: the default PSA is the lowest-ranking direct core argument in terms of (20).

The principle in (21a) captures the fact that, in English and many other languages, the actor is the PSA of a transitive construction, whereas (21b) captures the selection of undergoer as default PSA in Dyirbal transitive constructions. Conversely, the marked PSA selection found in the English passive is undergoer, while the marked PSA selection found in the antipassive is actor. The principles in (21) mention direct core arguments (see footnote 2), as opposed to macroroles, because of the existence of languages in which non-macrorole arguments can be PSAs (Icelandic, Georgian, Japanese, etc.). In the present context, however, we will not dwell on this difference.

At this point we should mention constructional templates or *schemas* (Van Valin & LaPolla 1997b: 430-436; Van Valin 2005: 132-135). These are constellations of syntactic, morphological, semantic and pragmatic instructions, which, while making reference to the general principles of grammar, complement them with the language-particular information that is necessary to form and parse the constructions of a given language. In the formation of the English passive, it is the passive constructional schema that specifies that the PSA is not chosen in accordance with the default PSA selection principle (cf. (21a)), usually owing to discourse-pragmatic factors. In addition, the constructional schema establishes that the actor cannot occur within the syntactic Core, although it can be expressed in a *by*-phrase, and that the verb carries special, passive, morphology. In *wh*-questions, it is a constructional schema that instructs the speaker on the

[9] The Logical Structure of the predicate in the clause is ascertained by applying a number of standard tests, which include Dowty's (1979) ones, as mentioned above. Therefore, there are independent criteria to establish the status of the candidates for PSA-hood vis-à-vis the hierarchy in (20).

default position of the *wh*-word in the given language and whether the *wh*-word is subject to any restricted neutralizations (Van Valin 2005: 132-133). In languages that have different PSAs in different constructions (for example Jakaltek), constructional schemas specify what the PSA is in the given construction.

To return to grammatical relations, in the absence of a notion of object, the correlation between the functions that in other frameworks are the transitive object and the unaccusative subject is captured in RRG in terms of the thematic properties of the PSA, with some unaccusative patterns being restricted to undergoers and others to the lowest ranking argument, regardless of whether this is assigned a macrorole or the status of PSA. Thus, unaccusative subjects in RRG are not underlying objects, but rather PSAs which are linked from the two rightmost positions in (20), similarly to passive PSAs.[10] It is to the linking that we now turn, as the final topic of Section 2.

2.4 The linking

As can be seen in Figure 4, the linking is bidirectional, to account for both language production and language comprehension, and includes both universal and language-specific steps. Whereas logical structures and macrorole assignment, which is based on the hierarchy in (9) and the principles in (10), are universal, languages differ substantially in how arguments link to syntax.

The linking is governed by the Completeness Constraint, which ensures that there is a match between the referring expressions in the clause and the arguments in the semantic representation of the clause.

(22) Completeness Constraint
All the arguments explicitly specified in the semantic representation of a sentence must be realized syntactically in the sentence, and all the referring expressions in the syntactic representation of a sentence must be linked to an argument position in a logical structure in the semantic representation of the sentence.

The semantic representation of the sentence is built on the basis of the Logical Structures of the predicators in the clause (including the predicating adpositions of adjunct modifiers). These Logical Structures are drawn from the lexicon, although the semantics of the predicate is also subject to compositional rules,

[10] It would not be possible to review here to the wide range of crosslinguistic variation in unaccusativity. We refer to Centineo (1986, 1995), Van Valin (1990), Bentley (2006), among others, for some of the RRG treatments of this topic.

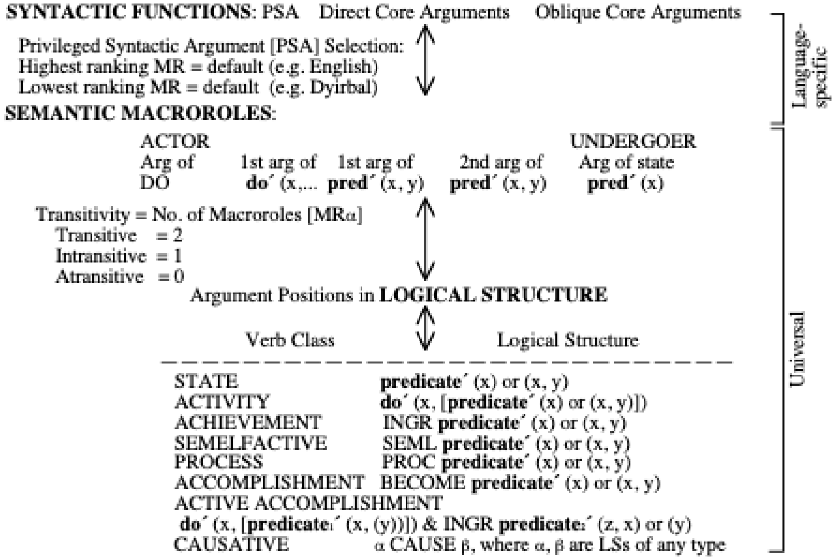


Figure 4: The linking of semantic and syntactic representation (Van Valin & LaPolla 1997b: 177)

which we omit here for the sake of brevity.[11] In the mapping from semantics to syntax, the information in the semantic representation of the clause is key for the retrieval of the appropriate syntactic templates from the syntactic inventory. The selection of the core template is governed by the principle in (23a).

(23) a. Syntactic Template Selection Principle
The number of syntactic slots for arguments and argument-adjuncts within the core is equal to the number of distinct specified argument positions in the semantic representation of the core.[12]

[11]We refer here to alternations between activities and active accomplishments which depend on whether the activity in the Logical Structure of the predicate combines with the Logical Structure of an adpositional phrase describing an endpoint.

[12]An argument-adjunct is an adposition which introduces an argument of the verb, at the same time contributing its semantics to the clause. The locative adposition required by *put* is an argument-adjunct, since it is part of the valence of the verb, and hence is an argument, but it can vary independently of the verb (e.g., *put the book on/under/next to,* etc. *the desk*) in the same way an adjunct can (e.g. *dance on/next to/beside,* etc. *the desk*). The semantic representation of *x puts y in z* ([**do′**(x, [**act.on′**(x, y)])] CAUSE [INGR **be-in′** (z, y)]) reflects the argument sharing between the verb and the adposition in a way that the semantic representation of *x dances on*

b. Language-specific qualifications of the Principle in (23a):
   i. All cores in the language have a minimum syntactic valence of 1.
   ii. Argument-modulation voice constructions reduce the number of core slots by 1.
   iii. The occurrence of a syntactic argument in the Pre-/Post-Core Slot reduces the number of core slots by 1 (may override 23b-i).

The Principle in (23a) follows from the Completeness Constraint and is universal, whereas the qualifications in (23b) are language-specific (though they all apply to English). An additional, universal, qualification of (23a) is needed to capture non-subordinate complex constructions, and we refer to Van Valin & LaPolla (1997b: 546) and Paris (to appear) for this.

In the syntax to semantics linking the syntactic representation of the clause is created by a Parser on the basis of the overt syntactic structure of a sentence. The Parser appears alongside the syntactic inventory in the general architecture of RRG shown in Figure 1 (Van Valin 2005: 131). The constructional schemas also appear in the RRG architecture, since they play a key role in providing language- and construction-specific information in both directions of the linking. The step-by-step procedures that characterize the linking, in both directions, are detailed in the Linking Algorithm(s), which are rather complex, to capture language-specific variation (Van Valin 2005: 136-158).

Having introduced how RRG is conceived and how the parts of the model fit together, in the next sections we shall engage in a more detailed comparison of the different ways things are done within RRG and LFG.

3 LFG and RRG compared

As noted above, both LFG and RRG fall within the class of linguistic models defined as parallel correspondence or level-mapping. There are nonetheless significant differences between them with respects to various dimensions of linguistic analysis and description. We consider some of these differences in a little more detail in the present section.

y does not: in **be-on**′(y, [**do**′(x, [**dance**′(x)])]) the Logical Structure of the adjunct *on* modifies the Logical Structure of *dance* taking this as one of its arguments, but there is no argument sharing between the two predicates.

3.1 Grammatical relations and control

A, perhaps the, key difference between the two frameworks concerns the status of grammatical relations like subject and object. These are at the heart of LFG, where they constitute the ingredients of f-structure, a level which stands as a crucial point of intersection between lexical argument structure, sentential syntax and meaning. By contrast, as we have seen, RRG regards grammatical relations as construction and language particular instantiations of possible argument relations and as such to be defined at the level of individual grammars rather than as an intrinsic part of the cross-linguistically applicable theoretical framework. Within LFG this reliance on functional structure has meant that the inventory of functions has had to be extended to include (x)COMP and (x)ADJ in order to accommodate the full range of embedded or subordinate clauses. Although the desirability of such as extended inventory has not gone unchallenged — see for example the discussion of COMP in Patejuk & Przepiórkowski (2016) — the fact remains that some f-structural account of all the parts of a sentence is required in LFG but not in RRG, where the semantically defined primitives suffice.

One place where this difference can be seen is in the treatment of control. The RRG treatment of these constructions has its roots in Foley & Van Valin's (1984: 307-308) theory of obligatory control, which is defined in semantic terms:

1. Causative and jussive verbs have undergoer control.
2. All other (M-)transitive verbs have actor control.[13]

Examples of causative verbs are *make*, *force* and *cause*, whereas *tell*, *persuade* and *order* are examples of jussive verbs, the latter group being distinct from the former in that it describes an eventuality that relies on verbal means. Examples with *persuade* were provided in (17). Here we provide an example with *tell*. The fact that the controller remains the same regardless of passivization (cf. (24b)) indicates that this construction has a semantic controller (undergoer).

(24) a. Mary told Paul$_i$ [CONTROLLER]__$_i$ to leave.
 b. Paul$_i$ [CONTROLLER] was told by Mary$_j$ __$_{i/*j}$ to leave.

[13]M(acrorole-)transitivity is the number of macrorole arguments that a verb takes. It is syntactically more salient than S(yntactic-)transitivity, which is the number of direct core arguments a verb takes. The difference between the two is clear in the case of activity verbs with active accomplishment counterparts (*eat/eat the cake*). Whereas the active accomplishments (*eat the cake*) are M-transitive (and therefore also S-transitive), the activities can have an inherent argument that has no macrorole status (*eat pasta*), in which case they are S-, but not M-transitive.

The control constructions with transitive verbs that are neither causative nor jussive also have a semantic controller, although here the controller is the actor:

(25) Paul$_i$ [CONTROLLER] promised Mary$_j$ __$_{i/*j}$ to leave.

The intuition behind the theory of control introduced above is that the lexical semantics of the verbs providing the controller determines the type of semantic control. Indeed, the theory is also valid in syntactically ergative languages (for example, Dyirbal), languages with active-inactive alignment (Acehnese) and head-marking languages (Lakhota) (Van Valin 2005: 241). In addition, if a verb can have causative and non-causative or jussive and non-jussive semantics (see, for example, *ask*) the semantics of the controller varies accordingly. If the verb providing the controller is intransitive, as for instance is the case with *try*, there is no issue of selection.

(26) Paul$_i$ [CONTROLLER] tried __$_i$ to leave.

The controlled missing argument, or pivot, on the other hand, is a PSA in all of the constructions above, in that it is characterized by the restricted neutralization [A, S, d-S].

(27) Mary told Paul$_i$ / Paul$_i$ promised Mary / Paul$_i$ tried __$_i$ to leave / __$_i$ to see a doctor / __$_i$ to be seen by a doctor / *a doctor to see__$_i$.

An important feature of control constructions is highlighted by the ungrammaticality of passivization of the first verb when this is neither causative nor jussive (cf. (28) vs. (24b)).

(28) *Paul was promised by Mary to leave.

The finding in (28) is explained by the type of semantic control that the structure requires (actor), combined with the type of syntactic linkage that the structure involves. This is a non-subordinate core juncture (Section 3.4), which independently requires that an argument of the second core be shared with — and realized within — the first core. The latter requirement is the additional, universal, qualification of the Syntactic Template Selection Principle (cf. (23)), which was mentioned in passing above (Van Valin & LaPolla 1997b: 546, Van Valin 2005: 244-245, Paris to appear).

(29) Universal qualification of (23a)
The occurrence of a core as the linked core in a non-subordinate core juncture reduces the number of core slots by 1.

The actor of the passive is independently claimed not to occur within its core in RRG, and, therefore, the specific argument sharing required cannot take place in (28). This results in a violation of the Completeness Constraint in the linking (cf. (22) in Section 2) and, hence, in ungrammaticality.

Raising to subject/raising to object/Exceptional Case Marking constructions are called Matrix Coding constructions in RRG. We give an example of matrix coding as PSA in (30a) and of matrix coding as non-PSA in (30b).[14]

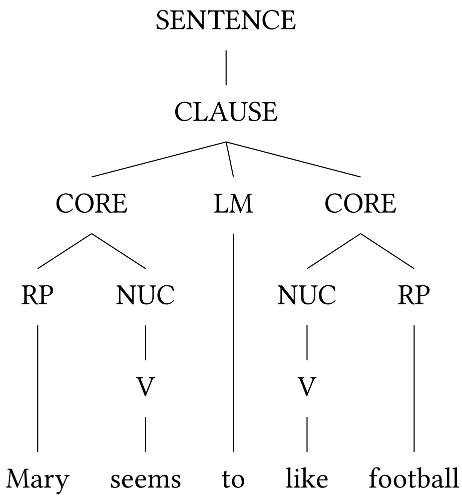
(30) a. Mary seems to like football.
  b. John believes Mary to like football.

Although these structures are characterized by the sharing of an argument between two cores, similarly to control constructions, the shared argument is not a pivot. In matrix coding to PSA, the matrix verb is bivalent, but atransitive, which means that it takes no macroroles. An example is **seem′**(x, y) [MR0], where x is an optionally realized experiencer and y is a propositional argument. In English, if *seem* is followed by a finite complement (*It seems that Mary likes football*), a non-argumental expletive pronoun (*it*) fills the initial position in the core of *seem*, satisfying the language-specific requirement of a nominative-marked RP in that position.[15]

Whether finite or non-finite, the propositional argument as such is not assigned a macrorole or a grammatical relation in RRG. Instead, the individual arguments within the propositional argument have macrorole status and play a key role in the linking in the matrix coding construction with a non-finite propositional argument (cf. (30a)). This construction coordinates two cores in the syntax: the core of *seem* and that of *like* in (30a) (see Section 3.5 and Figure 5). The predicate in the second core contributes an argument to the first core in the linking. This takes the place of the direct core argument in the first core, satisfying the universal qualification in (29), as well as the language-specific requirement of a nominative RP in the core-initial position. If an argument of the second core were not linked to the first core, the Completeness Constraint would be violated, given that, to satisfy (29), an argument specified in the Logical Structure of the verb in the second core could not have any syntactic expression.

[14] Although other epistemic predicates figure in matrix coding as PSA (for example, *be likely, be certain*) this structure is not in principle limited to epistemic predicates: modality impersonal (*be necessary, must*) and factitives (*be sad, be fascinating*) are also known to figure in matrix coding crosslinguistically (Kimenyi 1980; Bentley 2003).

[15] The optionally expressed experiencer argument cannot satisfy this requirement because it is not a direct core argument and hence it cannot be marked with nominative.



seem′ (∅, [**like′** (Mary, football)])

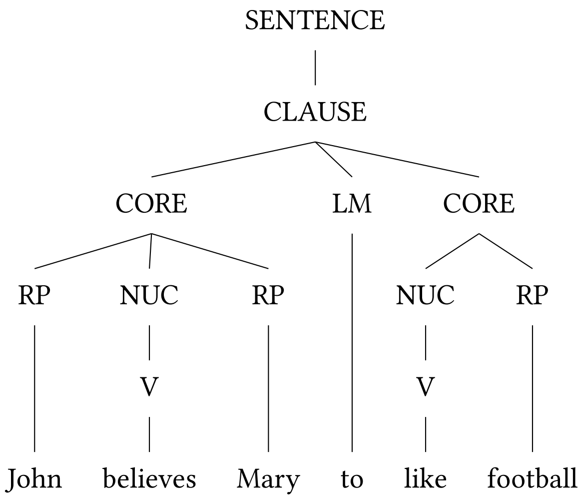
Figure 5: Semantic and syntactic representations of (30a)[16]

As for matrix coding to non PSA, the relevant verbs (*believe, expect, find, consider*, etc.) are M-transitive: an example is **believe′**(x, y). The second argument can be an NP or a proposition, i.e., a full clause (*John believes that Mary likes football*) or a core (cf. (30b)). In the latter case, an argument provided in the Logical Structure of the verb in the second core is linked to the first core to satisfy (29), again avoiding a violation of the Completeness Constraint.[17]

Therefore, in RRG, argument sharing in matrix coding is captured by an independent property of non-subordinate core junctures, i.e., (29). The difference between the two matrix coding constructions is a function of the lexical properties of the verbs occurring in the first core. Matrix coding as PSA characterizes two-place verbs which have no direct core argument to satisfy the requirement of a nominative RP in core-initial position. With these verbs, (29) is satisfied by

[16]LM in Figure 5 and following figures stands for Linkage Marker.

[17]What should be noted here is that this sharing does not result in the absence of the propositional argument. In other words, in the matrix coding construction, the second argument position of *believe* is taken by one of the arguments in the second core, while the latter does not fill an argument slot itself. This differentiates the matrix coding construction from other verbs which can take either an NP or a clausal complement: *Phil told Dana a story* -> *Phil told Dana to leave* (*to leave* takes the place of one arguments); *Eileen remembered her purse* (2 arguments) -> *Eileen remembered to leave* (again, *to leave* takes the place of one argument). We refer to Van Valin (2005: 255-256) for relevant discussion.



believe′ (John, [**like′** (Mary, football)])

Figure 6: Semantic and syntactic representations of (30b)

an argument from the predicate in the second core taking the function of PSA in the first core. The other type of matrix coding characterizes M-transitive verbs which provide an argument of their own as PSA. With these verbs, (29) is satisfied by an argument of the second core taking the second argument slot in the matrix core.

In more general terms, the contrast between control and matrix coding constructions depends on the lexical properties of the verbs involved in them, with the function of the shared argument, as well as actor or undergoer control, being determined lexically.[18] Syntactically, all of these constructions are non-subordinate core junctures and they all abide by the constraints on this type of linkage.

By contrast, control in LFG makes fundamental use of the relations SUBJ and XCOMP (see Dalrymple et al. 2019: Chapter 15 and Vincent forthcoming [this volume]). To take two classic instances, the entries in (31) are those proposed by Bresnan (1982) for the functional control verbs *seem* and *try*, the difference between the two lying in the fact that for *seem* the SUBJ function is not at the same

[18]The same is true of the structure that is commonly known as tough movement. This involves matrix coding as PSA with propositional attitude adjectives, which only have a propositional argument (*it is easy to please Mary, Mary is easy to please*), and control with psych action adjectives, which have a nominative RP of their own (*Mary is eager to please*) (see Van Valin & LaPolla 1997a: Chapter 9, exercise 6).

time a semantic argument whereas for *try* it is. That said, the crucial equivalence of the embedded and matrix SUBJ is formally the same in both cases.

(31) a. *seem* V (↑ PRED) = 'SEEM⟨XCOMP⟩SUBJ'
(↑ SUBJ) = (↑ XCOMP SUBJ)

b. *try* V (↑ PRED) = 'TRY⟨SUBJ, XCOMP⟩'
(↑ SUBJ) = (↑ XCOMP SUBJ)

At the same time we should emphasise that in neither framework are the particular analyses necessarily unchallenged or unchallengeable. Although it would clearly go against the principles of LFG for an analysis not to be mediated by grammatical functions, just as it would not be consistent within RRG for direct reference to be made to subject or object, the preferred account may vary from scholar to scholar. Thus, Dalrymple et al. (2019: 561-566) offer the alternative entry for *try* as in (32):

(32) *try* V (↑ PRED) = 'TRY⟨SUBJ, COMP⟩'
(↑ COMP SUBJ PRED) = 'PRO'

On this version, *try* is considered to be an instance of obligatory anaphoric control rather than functional control, and hence the second argument is COMP rather than XCOMP. This in turn means that the PRED value of the embedded subject is PRO; in other words a kind of null function whose value is interpreted by reference to another argument in the clause. We return to the issue of null arguments in Section 3.3 below.

3.2 Predicate types and theta roles

The presence vs. absence of grammatical relations is also evident in the different way the two models characterise the individual lexical predicates. In LFG the lexical entry for a given item includes mention of the grammatical relations with which it is associated, which in turn are related to the appropriate theta roles by means of lexical mapping. In RRG, by contrast, the lexical entries are stated directly in logico-semantic terms (see Section 2), and it is these which are then linked to the argument structure as indicated in (8) and (9) above. Moreover, RRG adopts a distinct theory of theta roles which incorporates the two macrocroles Actor and Undergoer, which are defined in terms of positions in Logical Structure. These are generalizations across argument types and have no direct analogue in LFG, or indeed in any other framework which we are familiar with. Macroroles constitute the primary interface between logical structure and syntax, and their assignment is governed by universal principles, and no further argument structure is postulated. Meanwhile, the logical structure based on Dowty's and Vendler's approach to Aktionsart types allows for a more fine-grained classification of predicates than is to be seen in LFG, while at the same time allowing the argument structure to emerge from the logical structure rather than having to be defined in a separate structural dimension.

One consequence of the decision within RRG not to admit as theoretical primitives a separate set of grammatical functions means that the question of how to relate these to semantic roles does not arise. In short there is no RRG equivalent of lexical mapping theory.

3.3 Argument realization

So far we have discussed for the most part issues relating to the content side, whether syntactic or semantic, of arguments. There are, however, differences between the two approaches when it comes to the way those arguments and associated clausal structure are given realization. One case in point concerns the treatment of null arguments, or so-called pro-drop, as in the contrast between French *il/elle arrive* beside Italian *arriva* 'he/she/it arrives'. Here LFG and RRG agree on rejecting the categorial solution but LFG has instead recourse to the null function seen above in the analysis of control. A verb form such as Italian *arriva* will have PRO as the value of its SUBJ function with the person/number values being determined by the appropriate features which are independently required by the language's system of verbal inflection. Not surprisingly in the literature this kind of account has been labelled 'pronoun incorporation' (Börjars et al. 2019: 68-75, Toivonen forthcoming [this volume]). At the same time it is also possible for the same verb form to have an overt argument as in Italian *arriva Giorgio* 'George is arriving' and hence the PRO value for the SUBJ constitutes an optional part of the verb's lexical entry triggered only when there is no overt argument. In other languages such as Chicheŵa this optionality also extends to the OBJ, but the formal mechanism is the same in both instances. For further discussion and exemplification, see Dalrymple et al. (2019: 179-85, 500-502) and Bresnan et al. (2016: Chapter 8).

In RRG, when the argument is not expressed independently of the verb, the verb inflection bearing its person/number features is linked to the Constituent Projection, similarly to the verbal affixes of head-marking languages (Van Valin & LaPolla 1997b: 331-332). In cases of extensive discourse-driven zero anaphora, found in Thai, Mandarin and Japanese, pro-drop is dealt with as a direct linking from discourse to Logical Structure, and, following Van Valin (2005: 171-174), Kamp & Reyle's (1993) Discourse Representation Theory has been adopted to formalize this linking.

A different issue concerns the treatment of long-distance dependencies as in *wh*-questions. There is considerable cross-linguistic variation here (for a typological survey see Mycock 2006) but the crucial point is that the questioned item need not occur in the position of the corresponding answer. Within derivational models this can be straightforwardly handled by a rule of *Wh*-movement which shifts the relevant item to the initial position in the clause in a language like English, while a language like Chinese has no such rule and therefore question and answer occupy the same slot. For RRG the position of the *wh*-item depends on a language-specific aspect of the linking, which is specified in a constructional schema (Section 2.3), and is directly activated with the selection of an appropriate syntactic template (Section 2.4). LFG relies instead on a further function FOCUS with the functional value of the questioned item being set as equivalent to the FOCUS and therefore being realised in that slot wherever in the language that may occur. In this way it is possible to accommodate not only languages like English with a single initial slot or Chinese where the interrogative item remains *in situ* but also languages like Bulgarian which allow several different *wh*-items to occur in sequence at the beginning of the clause.

More generally then, as we have noted in various places, LFG tends where possible to avoid the proliferation of functional heads which is a characteristic of cartographic and nanosyntactic approaches. Thus, for example, although recourse is standardly had to CP to label clauses with a fronted question word or an embedded complementizer, there is no automatic assumption that all simple main clauses are CP, nor is there any attempt to split C into separate functional heads to host topics and other fronted elements. And while some LFG accounts incorporate K as the category to be associated with items such as the Hindi-Urdu ergative particle *ne* (Dalrymple et al. 2019: 102-103), this is not the general practice (see Vincent 2021 and Przepiórkowski & Patejuk 2021 for further discussion and exemplification). By contrast, such analytical strategies have no analogue within RRG, where the categorial inventory is reduced to a minimum and functional heads do not figure at all.

Finally, in this connection, an instructive case concerns the treatment of the phenomenon of co-subordination (on which see Section 3.5 below). This is a concept unique to RRG and which has no analogue either in traditional grammar or in LFG, both of which distinguish simpy co-ordination, marked by items such as *and* and *or*, and subordination, signalled by various kinds of finite and non-finite complementation patterns. Foley (2010) argues against the necessity of postulating such a third mode of clause combining and offers instead an account within

LFG based on a categorial distinction between the functionally headed IP and the plain S or small clause. A response by Van Valin (2021) argues against Foley's account and more generally against the postulation of categorial solutions to what are functional/semantic problems.

3.4 Syntax and configurationality

Another dimension of linguistic realization concerns constituency and configurationality. In the various versions of Minimalism and cartography all structures are by definition configurational, and thus data such as the following Warlpiri example (cited from Austin & Bresnan 1996, example (1)) are problematic.

(33) Warlpiri

kurdu-jarra-rlu	=**ka-pala**	maliki	wajili-pi-nyi	wita-jarra-rlu
child-DUAL-ERG	PRES-3DU.SBJ	dog-ABS	chase-NPST	small-DUAL-ERG

'Two small children are chasing the dog.'

According to Hale (1983), after whom this example is cited, native speakers accept any order of the words here provided that the auxiliary element (highlighted in bold) is cliticised to the first item. Moreover, the adjective 'small' and the noun 'child' may, but do not have to, go together and if they do they can count as a constituent and occupy first position before the cliticised auxiliary. A fully configurational model can only handle this kind of data by postulating one structure as underlying and deriving the other options by movements to predetermined slots, some of which will inevitably be unfilled. In addition, the arguments of the verb in Warlpiri may remain unexpressed if derivable from context. In that case the relevant position in the tree is still present but is filled by a null pro. However, in a model such as LFG, once f-structure and c-structure are separated and not required to map onto each other in a one-to-one fashion, as Austin & Bresnan (1996) show, it is a straightforward matter to distinguish the argument structure from the way those arguments are realised in terms of linear order. Strict configurationality is then a requirement of particular languages such as English or Arabic, but it is not a property of universal grammar.

Within RRG the thinking is very similar. Not only is endocentricity not a principle of RRG, but there is also no expectation that the components of individual constituents, or units of meaning within the clause, should be contiguous. The flat structure of the RRG Layered Structure of the Clause (Section 2.1), therefore, caters straightforwardly for non-configurational languages, as in its own way does LFG by not requiring all constituents to be endocentric and thus arriving at flat structures by a different but equally satisfactory route.

3.5 Predicate and clause linkage

The RRG theory of predicate and clause linkage relies on the key notions of nexus and juncture. Nexus is the relationship established between two layers of the Layered Structure of the Clause (Section 2.1): instead of the traditional coordination vs. subordination dichotomy, RRG makes a trifold distinction between coordination, co-subordination, and subordination. Each of these types of nexus can in principle occur at three levels of juncture, nucleus, core, or clause, as can be seen in Table 1, although it is not the case that all languages exhibit all the possible nexus-juncture types.

Table 1: Nexus-juncture combinations

Juncture	**Nexus**
Nucleus	Coordination
	Co-subordination
	Subordination
Core	Coordination
	Co-subordination
	Subordination
Clause	Coordination
	Co-subordination
	Subordination

Nuclear junctures involve a single core containing two or more nuclei, core junctures normally feature two cores within a clause, and, finally, clausal junctures are typically characterised by two clausal nodes within a sentence. We discuss below some more complex constructions whereby a core joins with a clause. Since the operators expressing grammatical categories such as aspect, modality and tense are assumed to have scope over specific layers of the clause, operator scope is an important criterion to diagnose the level of juncture of a linkage.

We will not discuss each nexus-juncture type in detail (for exhaustive treatments see Van Valin & LaPolla 1997b: 441-492, Ohori to appear). Instead, we shall first deal with co-subordination, which is not a construct of LFG, and we shall contrast it with coordination, exemplifying at the same time the key diagnostics of linkage used in RRG. Then, we shall move on to subordination, which is subdivided into the complement and the adverbial type, in accordance with assumptions made in other frameworks.

The notion of co-subordination originated with scholarship on Papuan languages, where a type of clause linkage was found which can neither be analysed as coordination nor as subordination: on the one hand, the linked clauses cannot stand alone and are dependent on a matrix clause for the expression of clausal operators; on the other hand, they fail to exhibit the marking of subordination that obligatorily occurs elsewhere. We provide here some exemplification from Chuave (Thurmann 1975).

(34) Chuave (Papuan, Van Valin & LaPolla 1997b: 448)
Yai kuba i-re kei si-re fu-m-e.
man stick get-SEQ.SP dog hit-SEQ.SP go-3SG-IND
'The man got a stick, hit the dog, and went away.'

Although (34) translates as a coordination in English, it is not a coordination in Chuave because the first two clauses cannot stand alone, which would be expected if they were coordinated main clauses, and because they lack their own illocutionary force morpheme. Every independent utterance requires an illocutionary force marker in Chuave (see *-e* in (34), which is glossed as indicative), and the fact that this marker is shared by the clauses in (34) suggests that they are not coordinated, but rather stand in a dependence relation.

RRG thus distinguishes co-subordination from coordination, assuming that the former nexus type involves operator sharing. Specifically, the non-matrix unit(s) must depend on the matrix unit for the expression of at least one operator at the relevant level of juncture. An important corollary of this assumption is that when nuclei, cores, and clauses are joined together in a co-subordinate nexus, the first node that joins them is not of the higher type, but rather constitutes the same layer as the linked layers, as shown in Figure 7, which contrasts with Figure 8, representing coordination.

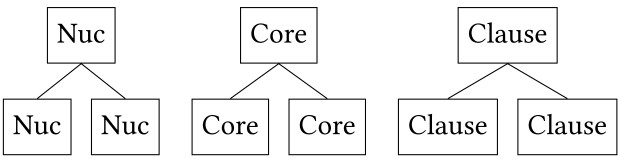


Figure 7: Nuclear, core and clausal co-subordination

We refer to Van Valin & LaPolla (1997b: 455) for exemplification of all the co-subordination and coordination linkages that are available in English. Here we should mention that since there are no sentence-level operators, sentences allow coordination and subordination (Van Valin 2005: 192), but not co-subordination.

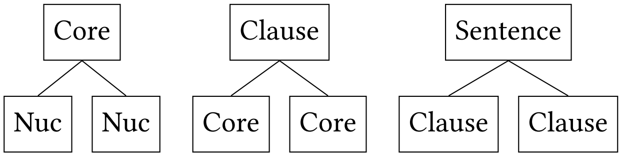


Figure 8: Nuclear, core and clausal coordination

Sentential coordination and subordination can thus be added to the nexus-juncture types shown in Table 1, although, again, it is not predicted that all languages will allow these types of linkage.

The contrast between cosubordination and coordination emerges in non-finite complementation. Compare the English constructions in (35).

(35) a. Mary tried to open the door.
b. Mary told Paul to open the door.

In both cases the relevant level of juncture is the Core, as suggested by the fact that in neither structure do the two predicates share the nuclear, aspectual, operators: the perfect and progressive operators only have scope over *try* and, respectively, *tell*, in (36a) and (36b).

(36) a. Mary has been trying to open the door.
b. Mary has been telling Paul to open the door.

Sharing of all the arguments, as evidenced by passivization, would also indicate a nuclear juncture, but in both constructions *the door* is an argument of *open* alone (**The door is tried to open by Mary*, **The door is told Paul to open by Mary*).

The two predicates do share one argument (Mary and, respectively, Paul), which is suggestive of a Core juncture. Yet, there is a key difference between the two constructions. The non-matrix predicate depends on *try* for the expression of deontic modality in (35a), and, therefore, (37a) can be read as (37b).

(37) a. Mary must try to open the door.
b. Mary must open the door.

Although *must* cannot be embedded under *tell* for independent reasons (it rejects the *to* infinitive), importantly, the same operator sharing as in (37) does not apply to the structure with *tell* in (35b).

(38) Mary must tell Paul to open the door ≠ Mary must open the door.

In light of the above evidence, the linkage with English *try* is analysed in RRG as a case of core cosubordination, as opposed to the one with *tell*, which is a case of core coordination.

Similarly to the construction in (35a), the one in (39a) illustrates core co-subordination, as testified by operator sharing at the level of the core. Not only does the embedded predicate *waiting* depend on the matrix predicate *sit* for the expression of deontic modality (see (35b)), but a deontic modal operator and core negation with scope on the matrix predicate can also have scope on the embedded one (see (39c)-(39d)). In addition, the two predicates share one argument, *Mary*.

(39) a. Mary sat waiting for your call.
b. Mary sat (*must) wait(ing) for your call.
c. Mary must sit waiting for your call > Mary must wait for your call.
d. Mary didn't sit waiting for your call > Mary didn't wait for your call.

In nuclear junctures all arguments of the linked predicates are pulled together as the arguments of a single nucleus. In Italian, this is evidenced by the occurrence of accusative or locative clitics to the left of the matrix predicate, even though such clitics express arguments of the second predicate. This structure is referred to as clitic climbing in frameworks which allow movement.

(40) Italian

a. Maria lo è tornata a prendere.
Mary OBJ.CL be.3SG return.PTCP to get
'Mary went back to get it.'

b. Maria ci è cominciata ad andare.
Mary LOC.CL be.3SG start.PTCP to go
'Mary started to go there.'

Since Rizzi (1976), the structures in (40a)-(40b)) have been known to be monoclausal. In RRG, they must be considered to be nuclear junctures, since the two predicates share all their arguments. The selection of the perfect auxiliary *essere* 'be' in (40) would at first seem to suggest that these are nuclear co-subordinations, whereby the non-matrix predicate depends on the matrix one for the expression of the perfect operator. This is clearly the case with (40a), since transitive *prendere* 'get' would otherwise select the perfect auxiliary *avere* 'have'.

(41) Italian

Maria lo ha preso.
Mary OBJ.CL have.3SG got
'Mary got it.'

The case in (40b) is more puzzling, since it is *andare* 'go' that selects *essere* 'be' in the perfect, whereas *cominciare* 'begin, start' would select either auxiliary *essere* 'be' or *avere* 'have', when occurring alone, and, in fact, it would not occur with 'be' with an animate PSA.

(42) Italian

a. Lui è cambiato? – Ha cominciato.
He be.3SG change.PTCP have.3SG started
'Has he changed?' – 'He has started.'

b. Il film è cominciato.
the film be.3SG start.PTCP
'The film has started.'

In Bentley (2006: 82-83), we proposed that the structure in (40b) is a case of ad(verbial)-nuclear subordination, where *cominciare* 'start' is not a predicate because it does not contribute any arguments of its own, but merely an aspectual operator. This is represented with a Nucleus which lacks a predicate node but links to the Operator Projection to contribute aspectual information (see Figure 9). In the Constituent Projection, this Nucleus occurs in the periphery of the predicative Nucleus of the clause.

It is not uncommon for aspectuals, modals and indeed other classes of predicates to enter into more than one nexus-juncture type with other predicates in a given language.

RRG thus understands ad(verbial)-subordination as a structure whereby a given layer of the Layered Structure of the Clause has a peripheral modifier. While in (40b), the peripheral modification occurs at the level of the Nucleus, clausal ad-subordination — or ad-clausal subordination, as it is normally called — is illustrated in (43), which would also be analysed as a structure with an adverbial subordinate clause in other frameworks, including LFG. In this case, the subordinate clause *because you arrived* occurs in the Periphery of the main clause *Mary left*.

(43) Mary left because/even though you arrived.

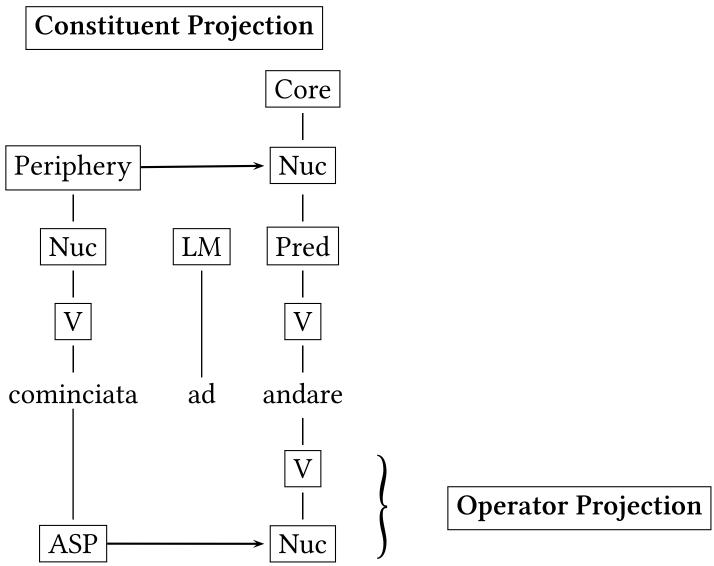


Figure 9: Ad-nuclear subordination with *cominciare* 'begin' in Italian

Observe that, in contrast with the cases of ad-subordination illustrated above in (40), the subordinate clause in (43) has a full-fledged predicate, which contributes its argument to the clause. It should not be assumed that by definition ad-subordination requires a modifier that lacks a predicate of its own. This can, but need not, be the case and it certainly is not the case with ad-clausal subordination.

Different semantic classes of verbs lend themselves to different nexus-juncture types (see Van Valin 2005:205-213 for a discussion of the rationale of the relevant patterns). To give but one, important, example, crosslinguistically, perception verbs lend themselves to forming less cohesive linkage types than causative verbs. No predictions are made in RRG on the exact nexus-juncture type that each predicate class will require in a given language. However, building upon Silverstein (1976) and Givón (1980), RRG has developed the *Interclausal Relations Hierarchy* (Van Valin & LaPolla 1997b: 481-483, Van Valin 2005: 209), which juxtaposes a scale of semantic relations with a range of nexus-juncture types, both being arranged in decreasing order of cohesion. The mapping between the two sides of the Interclausal Relations Hierarchy is many to one. However, RRG makes the strong falsifiable prediction that the tightest syntactic linkage realizing a particular semantic relation in a given language should be higher than, or as high as, the tightest syntactic linkage realizing lower semantic relations on the hierarchy in the same language. Although this prediction has been tested in work on specific languages (see, e.g., Casti 2012 on Sardinian), it ought to be further investigated in future work.

Where RRG has developed an innovative system of clausal organization and inter-clausal relations, LFG has remained more closely linked both to traditional grammar and work over the years in the generative tradition. Complex sentences involve the embedding of the c-structure of the subordinate clause within that of the matrix clause, but both are defined in terms of syntactic categories and in particular the concept of CP has been taken over wholesale from work in the Minimalist framework. It is true that the proliferation of functional heads within the clause has been avoided through recourse to the new grammatical relations (x)COMP and (x)ADJ, but, as noted above, categorial structure is still central in a way that it is not within RRG.

3.6 Pragmatics and information structure

The treatment of information structure in RRG and LFG is comparable, insofar as both frameworks consider it to be a module of grammar in its own right, which is independent from, but interacts with, the other modules. Both RRG and LFG allow information structure to be encoded in syntax (the layered structure of the clause, and, respectively, c-structure), morphology (Shimojo 1995), or prosody (see O'Connor 2006 for LFG and O'Connor 2008 for RRG). In addition, in RRG, the organization of grammar explicitly acknowledges the pervasive role of information structure at all stages of the bidirectional linking (see Figure 1). Broadly defining information structure as the organization of information in grammar, in this section we will address two principal issues, placing particular emphasis on RRG: (a) which information structure notions are adopted, and how they are defined, and (b) the place of information structure in the architecture of grammar.

Starting with the key notions, Zaenen forthcoming [this volume] draws a distinction between information structure proper, or the sentence-internal organization of information, and discourse structure, which is concerned with the packaging of information in larger textual units. This contrast does not find a parallel in RRG. While in both frameworks Lambrecht's (1994) notions of presupposition and assertion play a key role in the definition of topic and focus (Van Valin 2005: 68-73 and Zaenen forthcoming [this volume]), various different feature decomposition analyses have been developed in LFG to capture the nuances of salience, topic-worthiness, and contrastiveness (see Zaenen forthcoming [this volume] and references therein). RRG, instead, does not make use of feature decomposition, which is not to say that it does not attempt to capture the gradualness of the relevant notions, as will be explained in due course.

In RRG there is general consensus on which notions are relevant and how they should be labelled. The framework relies heavily on Lambrecht's (1994: 49) distinction between relational and non-relational constructs in information structure. The non-relational constructs are concerned with the status of the denotata of the discourse referents in the minds of the discourse participants: whether a given referent is already established or new for the hearer or both interlocutors, and, if it is new, whether it can be uniquely individuated or, alternatively, related to other referents. Although a referent is by definition brand-new, when it is first introduced into discourse, it may be possible for the interlocutors to identify it, in which case it is normally encoded as definite, in languages with overt marking of definiteness (e.g., *This morning I saw your sister / the Head of Department / the student you were taking about*). Otherwise it is unidentifiable and encoded as indefinite. Following Prince (1981) and Chafe (1987), RRG assumes that unidentifiable discourse referents can be anchored, i.e., related to established referents (e.g., *This morning I saw a student from the Physics Department)*, or, otherwise, unanchored (e.g., *This morning I saw a student*). Once a referent has been introduced, it becomes identifiable: if it is in the current focus of attention, it will be active; otherwise, it can be textually, inferentially or situationally accessible, or, alternatively, temporarily outside the focus of attention. The last type of discourse status is called inactive. Researchers in RRG have over the years investigated the grammatical correlates of the aforementioned notions in a large variety of languages (see, for example, Shimojo 1995, 2009, 2010, 2011; Pavey 2001; Belloro 2004, 2015; Matić et al. 2014; Latrouite & Riester 2018; Balogh 2021a, among others). The set of non-relational constructs which are universally adopted in the RRG treatment of information structure is illustrated in Figure 10, although we should note that other pragmatic states have been investigated by individual RRG researchers, for example saliency, or persistence in discourse (Shimojo 2009).

As for the relational notions, following Gundel (1988) and Lambrecht (1986, 1994, 2000), RRG defines topic as what the speaker wants to request information about, or increase the addressee's knowledge of, or get the addressee to act with respect to (Van Valin 2005: 68). The definition of topic is, therefore, inherently relational, in that it makes reference to the information unit about which new information is being requested or conveyed in the utterance. Importantly, the topic is also traditionally assumed to be part of the pragmatic presupposition, or the set of relevant propositions, and ultimately the information, which is shared by speaker and hearer prior to the utterance. Drawing on Reinhart (1981), Frascarelli

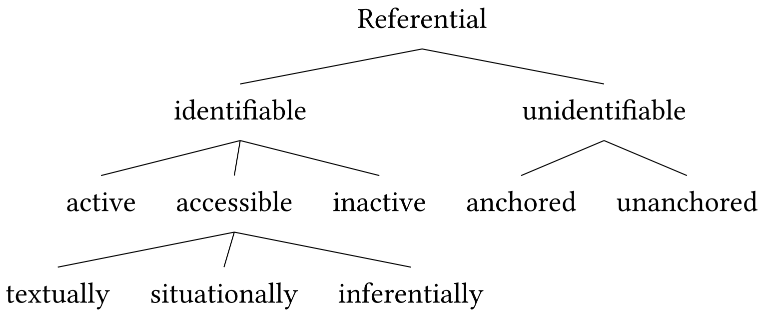


Figure 10: The cognitive states of referents in discourse (Van Valin & LaPolla 1997b: 201)

& Hinterhölzl (2007), Cruschina (2012), among others, in recent years, a distinction has been introduced in the framework between referential and aboutness topics, the first type being referentially old and part of the presupposition, the latter being introduced anew with the utterance, but nonetheless relational, in that it can be defined as what the utterance increases the addressee's knowledge about (see Bentley et al. 2015, Bentley (to appear)).

The gradualness of the notion of topic is captured in RRG at the interface with the non-relational notions mentioned above. In particular, it is assumed that topics align with active discourse referents, as can be seen in Figure 11.

Active	Most acceptable
Accessible	
Inactive	↕
Brand-new anchored	
Brand-new unanchored	Least Acceptable

Figure 11: The Topic Acceptability Scale (Van Valin & LaPolla 1997b: 204)

The morphosyntactic correlates of the alignment shown in Figure 11 are captured in Figure 12, which expresses the likelihood of marking of the topic by means of strategies that code referents in terms of their degree of accessibility.

Focus is defined in RRG as the part of a declarative utterance that is asserted (i.e., the component of that utterance whereby the assertion differs from the presupposition) or, in an interrogative utterance, the part that is questioned (Van Valin 2005: 69). The distinction between broad and narrow focus is made in the

Zero	Clitic/ bound pronoun	Pronoun [−stress]	Pronoun [+stress]	Definite NP	Indefinite NP

Markedness of occurrence as topic

Figure 12: Coding of referents as topic (adjusted from Van Valin & LaPolla 1997b: 205)

context of Lambrecht's (1994: 221-238) theory of focus structure, which has been extremely influential in RRG scholarship.

Focus structure can be defined as the conventional association of information meanings with sentence forms or the way that presupposed and asserted information are packaged in the sentences of a given language. While all grammars have strategies to differentiate sentences which provide new information on an established topic from sentences which occur out of the blue, and would seem to be topicless, there is a great deal of crosslinguistic variation in such strategies, and such variation has received attention in connection with the broader issue of the relative language-specific flexibility of the syntactic positions of predicates and arguments, and of focal information units (Van Valin 1999, Bentley 2008, etc.). Although Lambrecht's (1994) tripartition into predicate-, argument- and sentence focus is generally adopted (argument focus being renamed as narrow focus), the assumption that sentence focus lacks a topic altogether has been challenged, in light of Erteschik-Shir's (1997) theoretical work, which finds empirical support in the study of a number of seemingly topicless constructions, such as existentials and presentationals (Bentley et al. 2015, Bentley 2018).

An important distinction made in RRG is that between the Potential Focus Domain, which is the syntactic domain in the sentence of a given language in which focus can occur, and the Actual Focus Domain, which is the syntactic component of a given sentence that is in focus. The Potential Focus Domain differs across languages, as is clearly shown by the comparison of languages that heavily rely on prosody for the encoding of focus (e.g., English, see Vallduví 1992, Van Valin 1999) with languages that rely on syntactic position (e.g., Sicilian) or on the constructional choices (e.g., French). The Actual Focus Domain differentiates the three principal types of focus structure mentioned above.

To conclude the discussion of the relational notions that have received attention in RRG, we should mention contrastiveness. This is orthogonal to the notions of topic and focus, in that the alternatives that are contrasted can be selected from the presupposition or introduced anew within the assertion. Importantly, topical and focal contrasted units exhibit the same marking in some languages, whether by syntactic or morphological means (see Shimojo 2009, 2010, 2011 for Japanese and De Cia 2019 for North-Eastern Italo-Romance). In Japanese, for example, contrastive units can be marked as topics with *-wa*. To capture the inherent informational complexity of contrastiveness, Shimojo (2011) borrows Erteschik-Shir; Erteschik-Shir's (1997; 2007) notion of subordinate f(ocus) structure. The essence of his claim is that *-wa* marked contrastive units in Japanese are foci, because they are selected or highlighted from a finite set, but they are embedded in and selected from a topical, contextually available, set. A Japanese clause with contrastive *–wa* marking of the argument is thus represented as follows (Shimojo 2011: 275): $[\{x_{foc}, y\}_{top}]$-wa_{top} $[\text{predicate}]_{foc}$.

As for the place of information structure in the architecture of grammar, RRG, similarly to LFG, considers information structure to be an independent module of grammar. In terms of how this view is represented in each framework, King (1997) (cited in Zaenen forthcoming [this volume]) introduced an information structure projection in LFG, i-structure, and various proposals were subsequently advanced to model the flow of information from the other modules to i-structure. Similarly to LFG, RRG has a separate Speech Act Projection, which, however, does not participate in the flow of information, but rather represents the Potential and Actual Focus domain, and hence the focus structure, of an utterance in a given language.

In RRG the accessibility status of the discourse referents is conventionally represented in Logical Structure, the idea being that this status is significant in the construction of the meaning of the sentence. To give but one example, we showed in Figure 11 that an active referent lends itself more readily to the role of topic than an inactive or unidentifiable one. The topicworthiness of an active discourse referent may thus play a role in the selection of a specific lexical item as the predicate. Consider the lexical pair *fear* vs. *frighten*, or its rough Italian counterpart, *temere* vs. *spaventare*: an active stimulus will tend to be construed as the topic, which in turn will favour the choice of the *frighten* member of the pair in language production.

The activation status of the discourse referents also plays a role in the construction of meaning in language comprehension. Consider the case of an utterance which lacks an overt expression for one of the arguments of the predicate. Zero marking suggests that the position of that argument in Logical Structure can only be filled with an argument value that denotes an active discourse referent, or a referent that is textually, inferentially or situationally accessible. This referent must be retrieved from discourse.

The flow of information from the discourse context to linguistic expression is modelled in RRG by means of the tools offered by Discourse Representation Theory (Kamp & Reyle 1993), particularly in the analysis of zero anaphora phenomena, such as pro-drop (Section 3.3), in the absence of relevant morphological exponence, but also in the case of the silent predicates of Japanese and other languages. Importantly, the flow is supposed to occur directly between discourse and Logical Structure, without the intervention of syntax (or the Constituent Projection), given that empty syntactic arguments and positions are disallowed in RRG.

As should be clear from Figure 1, information structure plays a key role in the bidirectional linking of RRG. This view has already been illustrated in the discussion of the lexical choices for predicators and the filling of silent positions in Logical Structures. PSA choice, alongside voice alternations, are also heavily affected by the informational status of the arguments, as is the morphological marking of topics and foci. To conclude, we will briefly mention the stage in the linking which requires the selection of a syntactic template for the sentence. This stage involves language-specific considerations regarding a number of pragmatically-motivated positions: Pre- and Post-Core Slot and the Pre- and Post-Detached Position (see Section 2.1 and Balogh 2021b for further, language-specific, positions of Hungarian).

3.7 Semantic structure

Both LFG and RRG pay explicit attention to sentence semantics and, unlike Minimalism, neither theory requires the meaning of a sentence to be constructed one-to-one off syntactic heads and phrases. However, they differ in the way semantic and syntactic structure are integrated and in the type of semantic framework deployed. Within LFG, there is a separate dimension of s(emantic)-structure, which connects directly to f-structure rather than via c-structure. Although there is no strict directionality involved, it is nonetheless the case that f-structures, in turn built on the basis of the functional representations associated with lexical items, are input to the meaning construction, which is similar to the way, as described above, lexical semantics within RRG determines both the structure and overall meaning of the clause. At the same time two differences between the frameworks stand out. First as we have seen, RRG does not use grammatical functions as an intermediary point of analysis between argument structure and sentential meaning. Second, RRG relies solely on classical predicate logic and builds the semantic representation of a sentence in the lexical phase of the semantics-syntax (or syntax-semantics) linking, retrieving the meanings of the predicates from the lexicon and combining them without recourse to specific instructions other than the rules of predicate logic. There is no linear or resource logic equivalent to the role of Glue within LFG and thus RRG corresponds more closely to what Findlay (2021: 346) describes as the 'pre-Glue' stage in the development of LFG.

4 LFG, RRG and diachrony

Within LFG there has been relatively little historical work to date (for recent overviews see Börjars & Vincent 2017 and Booth & Butt forthcoming [this volume]) and in RRG even less (though see Ohori 1992, Eschenberg 2005, the contributions to Kailuweit et al. (2008) and Matasović (to appear)). However, both approaches have much to offer in the diachronic as well as the synchronic domain, as will be explored and exemplified in this section.

Given the traditional distinction between linguistic form/*signifiant* and content/*signifié*, changes can be broadly classified into three types: changes in form, changes in content and changes in the relation between the two. As far as the first is concerned, simple change of form or sound change, neither LFG nor RRG have anything special to say. Let us start then with the last and in particular the way these changes play out in the development of the Romance causatives, and where we can detect some instructive differences in the LFG-based account in Börjars & Vincent (2017: 651-655) compared to the RRG version in Kailuweit (2008: 79-83). The basic facts are fairly straightforward. Most Romance languages have a causative construction invoving the DO verb + infinitive (see Labelle 2017, Alsina forthcoming [this volume]) as in the French example (44).

(44) French

Je	ferai	manger	les	gâteaux	à	Jean
I	make.FUT.1SG	eat.INF	the	cake.PL	to	John

'I'll make John eat the cakes'

This structure, which has parallels across the whole of Romance from the earliest attestations (Vincent 2016) is monoclausal, as evidenced among other things by the fact that if the arguments are clitics they precede the higher verb (*je les lui ferai manger* 'I will make him eat them') and that the structure cannot be iterated (**je ferai faire manger les gâteaux à Jean à ses enfants* – contrast the biclausal English causative *I will make John make his children eat the cakes*). In other terms, what we have here is a complex predicate construction. There are similar examples in early Romance and late Latin texts.

(45) Old French (*Chanson de Roland* 852, 12th cent.)
en Sarraguce fait suner ses taburs
in Saragoss make.PRS.3SG sound.INF his drum.PL
'in Saragossa he makes his drums sound'

(46) Latin (*Vulgate,* Numbers 11.24, late 4th cent. CE)
quos stare fecit circa tabernaculum
who.ACC.MPL stand.INF make.PST.3SG around tabernacle.ACC
'who he made stand around the tabernacle'

However, if we go back further to an earlier stage we find a biclausal accusative and infinitive construction as in:

(47) Latin (Lucilius 1224, 2nd cent. CE)
purpureamque uvam facit albam pampinum habere
purple.ACC-and grape.ACC make.PRS.3SG white.ACC vine.ACC have.INF
'and it (the sun) causes the pale vine-shoot to have purple grapes'

That this is biclausal is evidenced by the fact that there are two accusatives here, one for the actor of the embedded clause and one for the undergoer, whereas in examples like (44) the embedded actor is marked by the preposition *à*, that is to say the usual marker of the non-macrorole core argument of ditrasitive verbs.

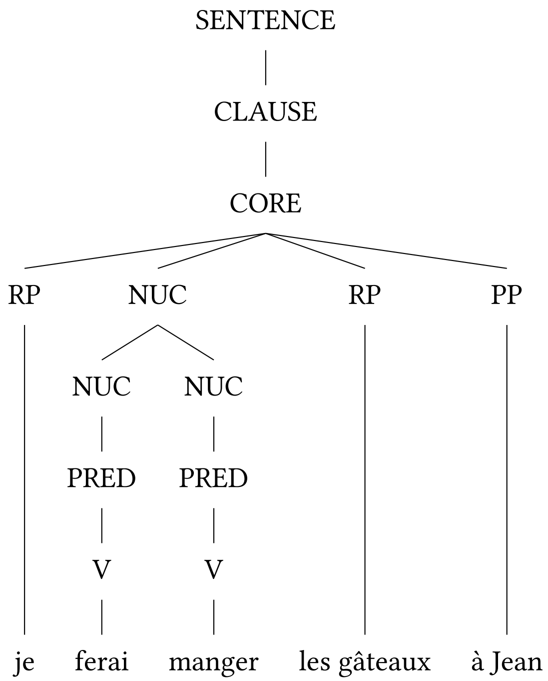
Two questions now arise:

1. how do the two frameworks model such constructions?
2. what diachronic trajectories do these synchronic analyses imply?

For example (44), Van Valin (2009: Figure 28.13) proposes the following structure:[19]

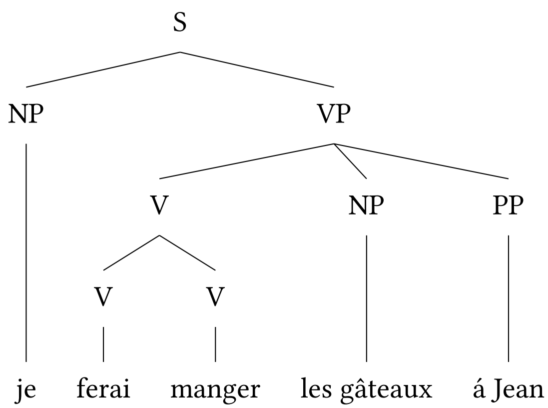
(48) Constituent projection of (44):

[19] Kailuweit (2008: 81) has essentially the same structure but with the verbal arguments dominated by [ARG [NP]] rather than, as here, by RP. Nothing of essence for the present issue hangs on this difference.



By contrast, the LFG representation in c-structure would be:

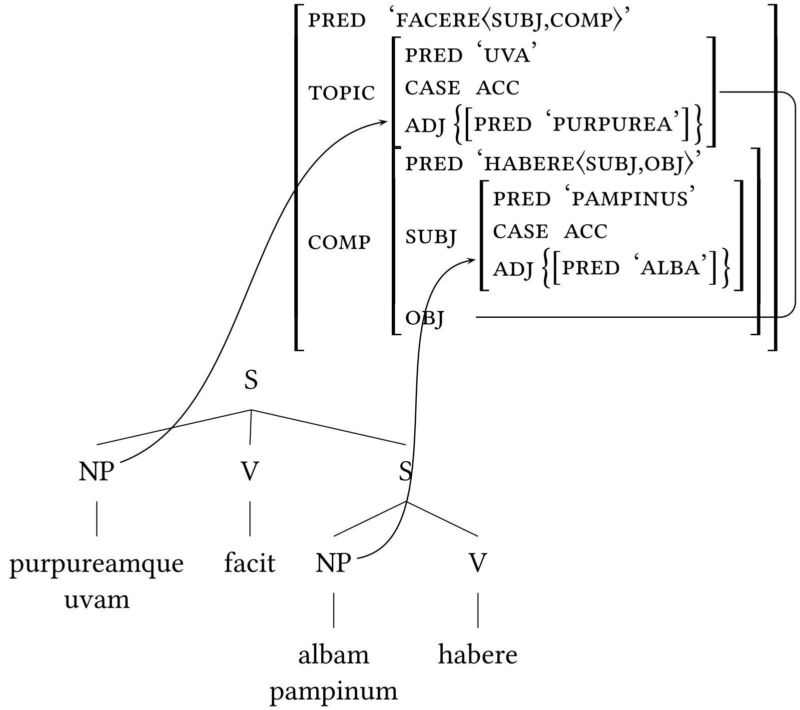
(49) C-structure of (44):



And to this would be linked an f-structure, where there is a single PRED value for the verbal complex: PRED ‘FAIRE.MANGER⟨SUBJ,OBJ,OBJ$_{\theta}$⟩’ with the three arguments respectively *je*.SUBJ, *les gâteaux*.OBJ and *Jean*.OBJ$_{\theta}$. The most appropriate c-structure for Romance causatives has been a matter of some discussion within the LFG literature ever since the early work of Alsina (1997) and is discussed in Alsina forthcoming [this volume] and Andrews forthcoming [this volume]. The structure in (49) is the one put forward in Börjars & Vincent (2017: 652) and is modelled on the proposal for Urdu complex predicates advanced by Butt (1997). The crucial property is that the 'make' verb and its dependent infinitive constitute a complex lexical item within a monoclausal construction and in this way accounting for the non-iterability of Romance causatives when compared to their English counterparts.

By contrast, the LFG tree for an example such as (47) would be as follows:

(50) C-structure and f-structure of (47):



A comparison of the two LFG representations shows that the change here has been modelled at the level of f-structure; where in the Latin example there were two separate predicates *facere* 'do' and *habere* 'have', in French we have rather a single complex predicate *faire manger*. In other words, there is a shift from a biclausal to a monoclausal pattern modelled through the changing functional structures of the relevant predicates. In the RRG account by contrast there is a change from core juncture to nuclear juncture (see Section 3.4), that is to say a similar pattern of structural conflation but achieved without reference to grammatical functions.

There are, then, parallels between the accounts within the two systems of both the earlier biclausal and the later monoclausal structures, but when it comes to describing and explaining the change from the one to the other over time there is a striking difference. As Börjars & Vincent (2017: 659) note, LFG has no inherent means of accounting for the directionality of change compared to for example the Minimalist framework. The latter includes a constraint that derivational movement, in the synchronic sense, is always upwards. Since the layers of functional structure always dominate the lexical layers, it follows that shifts can only be from lexical to functional exactly as the data from studies of grammaticalization predict. RRG by contrast, rather than relying on an abstract distinction between functional and lexical heads, incorporates the semantic-syntactic directionality directly into its overall structure via the Interclausal Relations Hierarchy (see Section 3.4 and Van Valin 2005: 209, Van Valin 2009: Fig 28.20, Matasović 2008). According to this view, there is an inherent link between semantic type and clausal structure. It is predicted therefore that a pattern containing the ingredients of causativity, if it is not already monoclausal, should move in that direction, exactly as the data we have reviewed above suggest. What neither model easily accounts for is the reversion to bicausality that is attested in some modern Romance varieties. The example in (51) is from the Piedmontese dialect of Borgomanero:

(51) Borgomanero (Tortora 2014: 155, ex. 154d)

al	farissa	vônga-ti	lü,	la	strija
SBJ.CL	make.COND.3SG	see.INF-you.SG	he	DEF	witch

'He would make you see the witch.'

In standard Italian or in French the clitic subject *ti* 'you' of 'see' would precede the causative in a monoclausal construction, while the fact that is attached here to the embedded infinitive leads Tortora (2014) to propose a biclausal account. Davies (1995) adduces similar evidence from modern Spanish and contrasts it with the monoclausal patterns found in the earlier stages of the language. Changes such as this suggest that it is not necessary to expect all diachronic developments to follow from asymmetries built into particular analytic frameworks, but some changes may be due to independently motivated external factors.

That said, diachrony does frequently show directionality, as is clear from the third type of change, namely those patterns that fall within the domain of grammaticalization. The emergence of grammatical markers such as tense/aspect auxiliaries, (in)definite articles and the like from former lexical items suggests that there are inherent links between different types of meaning, though the question remains open as to whether these should be attributed to forces external to language rather than to inherent properties of particular models. In this connection, Eschenberg (2005: Chapter 6), basing herself on earlier work by Rankin (2004), documents a striking series of changes in a set of particles in Umo^{n}ho^{n} (Omaha), which serve as both articles within the NP and evidentials within the clause. Here is not the place to go into detail but Figure 13 (Eschenberg 2005: 186) demonstrates the two functions of the item *k^{h}e* as a marker of deixis and subsequently as indicating the evidential basis for the speaker's assertion. She con-

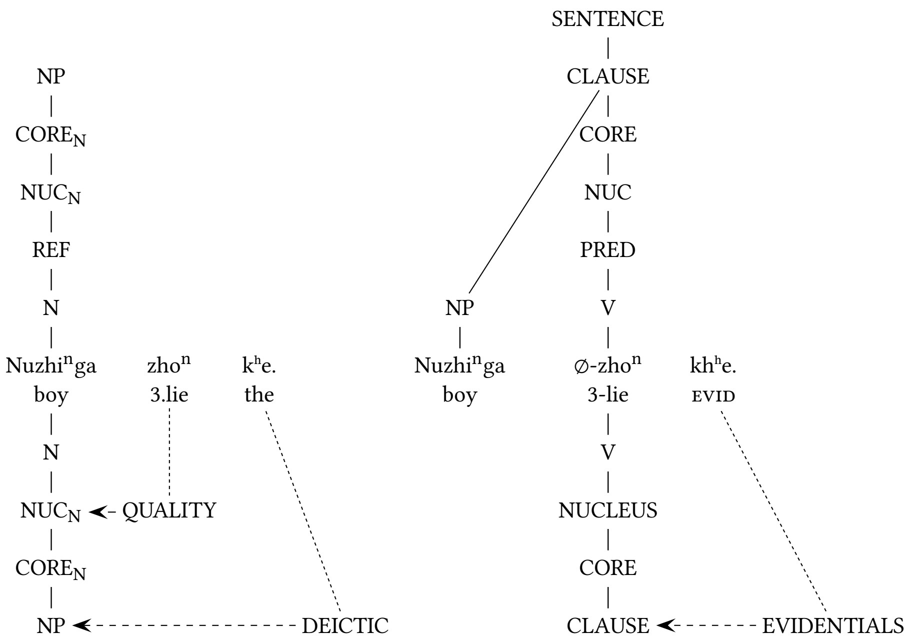


Figure 13: The marker *k^{h}e* in Umo^{n}ho^{n} (Omaha) (Eschenberg 2005: 186)

cludes that the structural parallels which an account along RRG lines suggests open up the potential for grammaticalization, though in fact no directionality is predicted and indeed over time some items within this class show a shift from auxiliary to article and back to auxiliary. The general conclusion, therefore, is that whatever the analytical framework, historical and synchronic data can and do complement each other.

5 Computational linguistics

Computational work has been a key component of LFG right from the outset (see the chapters in Part 5 of the present volume). By contrast, in the early stages the implementations of RRG were traditionally fewer, although there have been many relevant proposals in the last few years, and RRG now aims to offer an explanatory framework for the study of computational linguistics. While both approaches have been interested in parsing and sentence comprehension, the goal of developing more or less complete computational grammars of a range of languages has been a specific focus of LFG work, particularly but not exclusively via the ParGram project (Forst & King forthcoming: ?? [this volume]). The languages that figured within this project are typologically varied and in addition to the initial choice of English, French and German, the project has now been extended to include not only other Indo-European languages such as Norwegian, Polish, Urdu and Welsh but also a representative selection of languages from other families and parts of the world such as Georgian, Tigrinya, Japanese and Wolof. Comparable to the LFG ParGram project, albeit smaller in scale, are the RRG parbank project, a parallel treebank under development, which currently covers a small text corpus of German, English, Farsi, French, and Russian (Arps et al. 2021) and the RRG Biblical Hebrew treebank project (Cany Højgaard & Nielsen 2021). Moreover, Guest (2008) developed a parser which has been used to analyse a large corpus of English sentences and a somewhat smaller corpus of Dyirbal sentences (see also Nolan to appear). In addition, the cognitive scientist John Ball has, in the last decade, applied RRG in various Artificial Intelligence domains (see https://medium.com/pat-inc for details).

6 Psycho- and neuro-linguistics

Language acquisition and processing are domains in which RRG and LFG line up with each other in the sense that neither requires, nor finds evidence for, an innate UG (Pinker 1989, Van Valin 2002, Weist to appear and references therein). When it comes to acquisition both argue for the importance in the first instance of recurrent semantic patterns with syntactic structures only emerging at a later stage. Pinker (1982, 1989) in particular used LFG as a framework for the investigation and modelling of language acquisition, while Van Valin (1994, 1998, 2001, 2002) offers case studies from the perspective of RRG.

As for neurolinguistic research, RRG has been used as the grammar component of a sentence comprehension model developed in Bornkessel et al. (2004) and

Bornkessel & Schlesewsky (2006). Van Valin (to appear[a]) uses the RRG machinery to explain the ability of split-brain patients to provide grammaticality judgements with their isolated right hemisphere, developing a proposal which could potentially also capture the decoupling of grammaticality judgements and interpretation in agrammatic aphasics. For an overview of the relevant LFG-inspired work see Dalrymple et al. (2019: 726-728). Jones (2019) develops a new line of thinking for an 'incremental' version of LFG which addresses issues in relation to language processing and artificial intelligence.

7 Concluding remarks

In our introduction we alluded to the fact that both RRG and LFG share a commitment to formal architectures involving parallel structures and no derivations. In terms of the threefold classification of models proposed by Francis & Michaelis (2003) — a) derivational, b) licensing, as with the various kinds of construction grammar, and c) level-mapping, in which each level has its own structures and theoretical primitives — LFG and RRG both fall into their third class. At the same time, in his comments on an earlier draft of this chapter, Van Valin observes that 'RRG could be considered a kind of (generic) construction grammar, given its construction-specific theory of grammatical relations and use of constructional schemas to represent language-specific information'. That said, it must be noted that constructions are only deemed to be necessary in RRG when the general principles of the linking algorithm allow scope for variation, and thus can be applied in a construction-specific way. In similar vein, within LFG although proposals exist for integrating specific contructional types and idioms (see for example Asudeh et al. 2013), the model as a whole remains solidly based on words and phrases. The allusion to the sound-meaning link also suggests another dimension along which theories can be compared, namely the scale from syntax through semantics to pragmatics. At one extreme, there is cartography/nanosyntax with its insistence on the centrality of syntactic configurations and features while at the other there lies a purely pragmatics-driven model such as Dynamic Syntax (Kempson et al. 2016, 2017), which was set beside LFG in the workshop reported in Vincent (2009). Both LFG and RRG fall between these two extremes, but with LFG, given the importance of c-structure and the grammatical functions of f-structure, sitting nearer the syntactic end of the spectrum while RRG is more firmly based in semantic territory. However, there are signs of moves towards a larger role for semantics within LFG, as evidenced by Dalrymple et al. (2019: Chapter 8) and Asudeh forthcoming [this volume] and in a different way by Findlay forthcoming [this volume]. Only time will tell what the outcome of such a rapprochement might be.

Acknowledgements

Our thanks to Robert Van Valin, Adams Bodomo and a third (anonymous) reviewer for their comments and suggestions. Special thanks too to Mary Dalrymple for her advice and patience. Note that we provide full bibliographical references in this chapter on the RRG side whereas for LFG we often refer readers to the relevant chapter within the present volume.

Abbreviations

Besides the abbreviations from the Leipzig Glossing Conventions, this chapter uses the following abbreviations.

CL	clitic	SP	same PSA (privileged syntactic argument)
LNK	linker	SUFF	suffix
SEQ	sequential		

References

Alsina, Alex. 1997. Causatives in Bantu and Romance. In Alex Alsina, Joan Bresnan & Peter Sells (eds.), *Complex predicates*, 203–246. Stanford: CSLI Publications.

Alsina, Alex. Forthcoming. LFG and Romance languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Andrews, Avery D. Forthcoming. Clause structure and configurationality. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Arps, David, Tatiana Bladier, Kilian Evang, Laura Kallmeyer, Robin Möllemann, Rainer Osswald & Simon Petitjean. 2021. *Elementary trees in RRGparbank*. Paper presented at the 16th International Conference on Role and Reference Grammar, York University, Toronto, Canada.

Asudeh, Ash. Forthcoming. Glue Semantics. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Asudeh, Ash, Mary Dalrymple & Ida Toivonen. 2013. Constructions with lexical integrity. *Journal of Language Modelling* 1(1). 1–54. DOI: 10.15398/jlm.v1i1.56.

Austin, Peter K. & Joan Bresnan. 1996. Non-configurationality in Australian aboriginal languages. *Natural Language & Linguistic Theory* 14(2). 215–268. DOI: 10.1007/bf00133684.

Balogh, Kata. 2021a. Additive particle uses in Hungarian: A Role and Reference Grammar account. *Studies in Language* 45(2). 428–469.

Balogh, Kata. 2021b. *Negative particles and negative pronominals in Hungarian*. Paper presented at the 16th International Conference on Role and Reference Grammar, York University, Toronto, Canada.

Belloro, Valeria. 2004. *A Role and Reference Grammar account of third person clitic clusters in Spanish*. State University of New York at Buffalo. (MA thesis).

Belloro, Valeria. 2015. *To the right of the verb: An investigation of clitic doubling and right dislocation in three Spanish dialects*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Bentley, Delia. 2003. Sur la force d'une approche non-dérivationnelle de l'analyse linguistique: quelques donnés de l'italo-roman. In Jacques François (ed.), *Aspects de la role and reference grammar*, 51–73. Cahier du CRISCO.

Bentley, Delia. 2006. *Split intransitivity in Italian*. Berlin/New York: De Gruyter Mouton. DOI: 10.1515/9783110896053.

Bentley, Delia. 2008. The interplay of focus structure and syntax: Evidence from two sister languages. In Robert Van Valin Jr. (ed.), *Investigations of the syntax-semantics-pragmatics interface*, 263–84. Amsterdam: John Benjamins. DOI: 10.1075/slcs.105.20ben.

Bentley, Delia. 2018. Grammaticalization of subject agreement on evidence from Italo-Romance. *Linguistics* 56(6). 1245–1301. DOI: 10.1515/ling-2018-0022.

Bentley, Delia. 2019. The logical structure of verbs of quantized and non-quantized change. Paper presented at the 2019 International Role and Reference Grammar Conference, University at Buffalo.

Bentley, Delia. to appear. The RRG approach to information structure. In Delia Bentley, Ricardo Mairal-Usón, Wataru Nakamura & Robert Van Valin Jr. (eds.), *The Cambridge handbook of Role and Reference Grammar*. Cambridge, UK: Cambridge University Press.

Bentley, Delia, Francesco Maria Ciconte & Silvio Cruschina. 2015. *Existentials and locatives in Romance dialects of Italy*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780198745266.001.0001.

Bentley, Delia & Nigel Vincent. Forthcoming. LFG and Role and Reference Grammar. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Blake, Barry J. 1979. *A Kalkatungu grammar*. Vol. 57 (B). Canberra: Pacific Linguistics.

Booth, Hannah & Miriam Butt. Forthcoming. LFG and historical linguistics. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Börjars, Kersti, Rachel Nordlinger & Louisa Sadler. 2019. *Lexical-Functional Grammar: An introduction*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/9781316756584.

Börjars, Kersti & Nigel Vincent. 2017. Lexical-Functional Grammar. In Adam Ledgeway & Ian Roberts (eds.), *The Cambridge handbook of historical syntax*, 642–663. Cambridge, UK: Cambridge University Press.

Bornkessel, Ina & Matthias Schlesewsky. 2006. The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review* 113. 787–821. DOI: 10.1037/0033-295x.113.4.787.

Bornkessel, Ina, Matthias Schlesewsky & Robert Van Valin Jr. 2004. *Syntactic templates and linking mechanisms: A new approach to grammatical function asymmetries*. https://rrg.caset.buffalo.edu/rrg/vanvalin_papers/rrg_adm_CUNY04.pdf.

Bresnan, Joan. 1982. Control and complementation. *Linguistic Inquiry* 13(3). Reprinted in Bresnan & Kaplan 1982, 343–434.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-Functional Syntax*. 2nd edn. (Blackwell Textbooks in Linguistics 16). Malden, MA: Wiley-Blackwell.

Bresnan, Joan & Ronald M. Kaplan. 1982. Introduction: Grammars as mental representations of language. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, xvii–lii. Cambridge, MA: The MIT Press.

Butler, Christopher S. 2005. Functional approaches to language. In Christopher S. Butler, María de los Ángeles Gómez-González & Susana M. Doval-Suárez (eds.), *The dynamics of language use: Functional and contrastive perspectives*, 3–17. Amsterdam/Philadelphia: John Benjamins. DOI: 10.1075/pbns.140.04but.

Butler, Christopher S. 2006. Functionalist theories of language. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn., vol. 1–14, 696–704. Oxford: Elsevier. DOI: 10.1016/b0-08-044854-2/04771-4.

Butt, Miriam. 1997. Complex predicates in Urdu. In Alex Alsina, Joan Bresnan & Peter Sells (eds.), *Complex predicates*, 107–150. Stanford: CSLI Publications.

Bybee, Joan. 1985. *Morphology: A study of the relationship between form and meaning*. Amsterdam: John Benjamins.

Cany Højgaard, Christian & Nicolai Winther Nielsen. 2021. *Towards an RRG treebank of Biblical Hebrew*. Paper presented at the 16th International Conference on Role and Reference Grammar, York University, Toronto, Canada.

Casti, Francesco. 2012. *Testing the interclausal relations hierarchy: Aspectual and modal periphrases in modern Sardinian*. Manchester: University of Manchester. (Doctoral dissertation).

Centineo, Giulia. 1986. A lexical theory of auxiliary selection in Italian. In *Davis working papers in linguistics*, vol. 1, 1–35.

Centineo, Giulia. 1995. The distribution of *si* in Italian transitive-inchoative pairs. In Mandy Simmons & Teresa Galloway (eds.), *Proceedings of SALT 5*, 54–71. Ithaca: Cornell University Press. DOI: 10.3765/salt.v5i0.2708.

Chafe, Wallace L. 1987. Cognitive constraints on information flow. In Russell S. Tomlin (ed.), *Coherence and grounding in discourse*, 21–51. Amsterdam: John Benjamins.

Cortéz-Rodriguez, Francisco José. to appear. Morphology in RRG: The layered structure of the word, inflection and derivation. In Delia Bentley, Ricardo Mairal-Usón, Wataru Nakamura & Robert Van Valin Jr. (eds.), *The Cambridge handbook of Role and Reference Grammar*. Cambridge, UK: Cambridge University Press.

Cruschina, Silvio. 2012. *Discourse-related features and functional projections*. New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780199759613.001.0001.

Dalrymple, Mary, John J. Lowe & Louise Mycock. 2019. *The Oxford reference guide to Lexical Functional Grammar*. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198733300.001.0001.

Davies, Mark E. 1995. The evolution of the Spanish causative construction. *Hispanic Review* 63. 57–77. DOI: 10.2307/474378.

De Cia, Simone. 2019. *The syntactic-pragmatic interface in north-eastern Italian dialects: Consequences for the geometry of the left periphery*. Manchester: University of Manchester. (Doctoral dissertation).

Diedrichsen, Elke. 2008. Where is the precore slot? Mapping the layered structure of the clause and German sentence typology. In Robert Van Valin Jr. (ed.), *Investigations of the syntax-semantics-pragmatics interface*, 203–24. Amsterdam: John Benjamins. DOI: 10.1075/slcs.105.16die.

Dik, Simon. 1991. Functional grammar. In Flip G. Droste & John E. Joseph (eds.), *Linguistic theory and grammatical description*, 247–274. Amsterdam/Philadelphia: John Benjamins.

Dowty, David R. 1979. *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ* (Synthese Language Library 7). Dordrecht: Reidel.

Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68(1). 81–138. DOI: 10.2307/416370.

Durie, Mark. 1985. *A grammar of Acehnese on the basis of a dialect of North Aceh*. Dordrecht: Foris.

Durie, Mark. 1987. Grammatical relations in Acehnese. *Studies in Language* 11(2). 365–399. DOI: 10.1075/sl.11.2.05dur.

Erteschik-Shir, Nomi. 1997. *The dynamics of focus structure*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/cbo9780511519949.

Erteschik-Shir, Nomi. 2007. *Information structure: The syntax-discourse interface*. Oxford: Oxford University Press.

Eschenberg, Ardis. 1999. *Polish focus structure*. SUNY Buffalo. (MA thesis). https://rrg.caset.buffalo.edu/.

Eschenberg, Ardis. 2005. *The article system of Umo^{n}ho^n (Omaha)*. State University of New York at Buffalo. (Doctoral dissertation).

Findlay, Jamie Y. 2021. Meaning in LFG. In I Wayan Arka, Ash Asudeh & Tracy Holloway King (eds.), *Modular design of grammar: Linguistics on the edge*, 340–374. Oxford: Oxford University Press. DOI: 10.1093/oso/9780192844842.003.0020.

Findlay, Jamie Y. Forthcoming. LFG and Tree Adjoining Grammar. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Foley, William. 2010. Clause linkage and nexus in Papuan languages. In Isabelle Bril (ed.), *Clause linking and clause hierarchy: Syntax and pragmatics*, 27–50. Amsterdam/Philadelphia: John Benjamins. DOI: 10.1075/slcs.121.02fol.

Foley, William A. & Robert Van Valin Jr. 1984. *Functional syntax and universal grammar*. Cambridge, UK: Cambridge University Press.

Forst, Martin & Tracy Holloway King. Forthcoming. Computational implementations and applications. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Francis, Elaine J. & Laura A. Michaelis. 2003. Mismatch: A crucible for linguistic theory. In Elaine J. Francis & Laura A. Michaelis (eds.), *Mismatch: Form-function incongruity and the architecture of grammar*, 1–27. Stanford: CSLI Publications.

Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topics in German and Italian. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning and form*, 87–116. Amsterdam: John Benjamins. DOI: 10.1075/la.100.07fra.

Givón, Talmy. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4. 333–377. DOI: 10.1075/sl.4.3.03giv.

Guest, Elizabeth. 2008. Parsing for Role and Reference Grammar. In Robert Van Valin Jr. (ed.), *Investigations of the syntax-semantics-pragmatics interface*, 435–453. Amsterdam: John Benjamins. DOI: 10.1075/slcs.105.30gue.

Gundel, Jeanette. 1988. Universals of topic-comment structure. In Michael Hammond, Edith A. Moravcsik & Jessica Wirth (eds.), *Studies in syntactic typology*, 209–239. Amsterdam: John Benjamins. DOI: 10.1075/tsl.17.16gun.

Hale, Ken. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language & Linguistic Theory* 1. 5–47. DOI: 10.1007/bf00210374.

Ibañez Cerda, Sergio. to appear. Adposition assignment and adpositional phrase types in RRG. In Delia Bentley, Ricardo Mairal-Usón, Wataru Nakamura & Robert Van Valin Jr. (eds.), *The Cambridge handbook of Role and Reference Grammar*. Cambridge, UK: Cambridge University Press.

Jackendoff, Ray. 1976. Toward an explanatory semantic representation. *Linguistic Inquiry* 7(1). 89–150.

Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Jones, Stephen. 2019. *Modelling an incremental theory of Lexical Functional Grammar*. Oxford: University of Oxford. (D.Phil. Thesis).

Kailuweit, Rolf. 2008. Some remarks on RRG and grammaticalization: French verbal periphrases. In Rolf Kailuweit, Björn Wiemer, Eva Staudinger & Ranko Matasović (eds.), *New applications of Role & Reference Grammar: Diachrony, grammaticalization, Romance languages*, 69–86. Newcastle upon Tyne: Cambridge Scholars Publishing.

Kailuweit, Rolf, Björn Wiemer, Eva Staudinger & Ranko Matasović (eds.). 2008. *New applications of Role & Reference Grammar: Diachrony, grammaticalization, Romance languages*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic: An introduction to modeltheoretic semantics of natural language, formal logic and Discourse Representation Theory*. Dordrecht: Kluwer Academic Publishers.

Kempson, Ruth, Ronnie Cann, Eleni Gregoromichelaki & Stergios Chatzikyriakidis. 2016. Language as mechanisms for interaction. *Theoretical Linguistics* 42(3-4). 203–276. DOI: 10.1515/tl-2016-0011.

Kempson, Ruth, Ronnie Cann, Eleni Gregoromichelaki & Stergios Chatzikyriakidis. 2017. Action-based grammar. *Theoretical Linguistics* 43. 141–167. DOI: 10.1515/tl-2017-0012.

Kimenyi, Alexandre. 1980. *A relational grammar of Kinyarwanda*. Vol. 91. Berkeley: University of California Press.

King, Tracy Holloway. 1997. Focus domains and information structure. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '97 conference*, 2–13. Stanford: CSLI Publications.

Labelle, Maria. 2017. Causative and perception verbs. In Andreas Dufter & Elisabeth Stark (eds.), *Manual of Romance morphosyntax and syntax*, 299–331. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110377088-008.

Lambrecht, Knud. 1986. *Topic, focus and the grammar of spoken French*. Berkeley: University of California, Berkeley. (Doctoral dissertation).

Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representation of discourse*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/cbo9780511620607.

Lambrecht, Knud. 2000. When subjects behave like objects: An analysis of the merging of S and O in sentence-focus constructions across languages. *Studies in Language* 24. 611–682. DOI: 10.1075/sl.24.3.06lam.

LaPolla, Randy J. 1990. *Grammatical relations in Chinese: Synchronic and diachronic considerations*. Berkeley: University of California, Berkeley. (Doctoral dissertation).

LaPolla, Randy J. 1993. Arguments against 'subject' and 'direct object' as viable concepts in Chinese. *Bulletin of the Institute of History and Philology* 63(4). 759–813.

LaPolla, Randy J. 1995. Pragmatic relations and word order in Chinese. In Pamela Downing & Michael Noonan (eds.), *Word order in discourse*, 297–329. Amsterdam/Philadelphia: John Benjamins. DOI: 10.1075/tsl.30.11lap.

LaPolla, Randy J. to appear. Grammatical relations. In Delia Bentley, Ricardo Mairal-Usón, Wataru Nakamura & Robert Van Valin Jr. (eds.), *The Cambridge handbook of Role and Reference Grammar*. Cambridge, UK: Cambridge University Press.

Latrouite, Anja & Arndt Riester. 2018. The role of information structure for morphosyntactic choices in Tagalog. In Sonja Riesberg, Asako Shiohara & Atsuko Utsumi (eds.), *Perspectives on information structure in Austronesian languages*, 247–284. Berlin: Language Science Press.

Mairal Usón, R., L. Guerrero & C. González (eds.). 2012. *El funcionalismo en la teoría lingüística. la gramática del papel y la referencia. introducción, avances y aplicaciones*. Akal: Madrid.

Matasović, Ranko. 2008. Patterns of grammaticalization and the layered structure of the clause. In Rolf Kailuweit, Björn Wiemer, Eva Staudinger & Ranko Matasović (eds.), *New applications of Role and Reference Grammar*, 45–57. Newcastle upon Tyne: Cambridge Scholars Publishing.

Matasović, Ranko. to appear. RRG and diachronic linguistics. In Delia Bentley, Ricardo Mairal-Usón, Wataru Nakamura & Robert Van Valin Jr. (eds.), *The Cambridge handbook of Role and Reference Grammar*. Cambridge, UK: Cambridge University Press.

Matić, Dejan, Rik van Gijn & Robert D. Van Valin Jr. 2014. Information structure and reference tracking in complex sentences: An overview. In Rik van Gijn, Jeremy Hammond, Dejan Matić, Saskia van Putten & Ana Vilacy Galucio (eds.), *Information structure and reference tracking in complex sentences*, 1–42. Amsterdam: John Benjamins. DOI: 10.1075/tsl.105.01mat.

Mycock, Louise. 2006. *The typology of wh-questions*. Manchester: University of Manchester. (Doctoral dissertation).

Newmeyer, Frederick J. 1998. *Language form and language function*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/4115.001.0001.

Nolan, Brian. to appear. Computational implementation and applications of RRG. In Delia Bentley, Ricardo Mairal-Usón, Wataru Nakamura & Robert Van Valin Jr. (eds.), *The Cambridge handbook of Role and Reference Grammar*. Cambridge, UK: Cambridge University Press.

O'Connor, Robert. 2006. *Information structure in Lexical-Functional Grammar: The discourse-prosody correspondence*. Manchester: University of Manchester. (Doctoral dissertation).

O'Connor, Robert. 2008. A prosodic projection for Role and Reference Grammar. In Robert Van Valin Jr. (ed.), *Investigations of the syntax-semantics-pragmatics interface*, 227–244. Amsterdam: John Benjamins.

Ohori, Toshio. 1992. *Diachrony in clause linkage and related issues*. Berkeley: University of California, Berkeley. (Doctoral dissertation).

Ohori, Toshio. to appear. The structure and semantics of complex sentences. In Delia Bentley, Ricardo Mairal-Usón, Wataru Nakamura & Robert Van Valin Jr. (eds.), *The Cambridge handbook of Role and Reference Grammar*. Cambridge, UK: Cambridge University Press.

Paris, Luis. to appear. Cleft sentences and relative clauses. In Delia Bentley, Ricardo Mairal-Usón, Wataru Nakamura & Robert Van Valin Jr. (eds.), *The Cambridge handbook of Role and Reference Grammar*. Cambridge, UK: Cambridge University Press.

Patejuk, Agnieszka & Adam Przepiórkowski. 2016. Reducing grammatical functions in LFG. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-Driven Phrase Structure Grammar and Lexical Functional Grammar*, 541–559. Stanford: CSLI Publications.

Pavey, Emma. 2001. *Information structure in Acadian French*. University of North Dakota. (MA thesis).

Pinker, Steven. 1982. A theory of the acquisition of lexical-interpretive grammars. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 655–726. Cambridge, MA: The MIT Press.

Pinker, Steven. 1989. *Learnability and cognition: The acquisition of argument structure*. Reprinted in 2013 with new preface. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/9700.001.0001.

Prince, Ellen. 1981. Toward a taxonomy of given/new information. In Peter Cole (ed.), *Radical pragmatics*, 223–255. Academic Press.

Przepiórkowski, Adam & Agnieszka Patejuk. 2021. Coordinate structures without syntactic categories. In I Wayan Arka, Ash Asudeh & Tracy Holloway King (eds.), *Modular design of grammar: Linguistics on the edge*, 205–220. Oxford: Oxford University Press.

Rankin, Robert L. 2004. The history and development of Siouan positionals with special attention to polygrammaticalization in Dhegiha. *Sprachtypologie und Universalienforschung* 57(2-3). 201–227. DOI: 10.1524/stuf.2004.57.23.202.

Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27.

Rizzi, Luigi. 1976. Ristrutturazione. *Rivista di Grammatica Generativa* 1. 1–54.

Rizzi, Luigi. 2016. The concept of explanatory adequacy. In Ian Roberts (ed.), *The Oxford handbook of universal grammar*, 97–113. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199573776.013.5.

Sadock, Jerrold M. 1991. *Autolexical Syntax: A theory of parallel grammatical representations*. Chicago: University of Chicago Press.

Shimojo, Mitsuaki. 1995. *Focus structure and morphosyntax in Japanese: Wa and ga, and word order flexibility*. State University of New York at Buffalo. (Doctoral dissertation).

Shimojo, Mitsuaki. 2008. How missing is the missing verb? The verb-less numeral quantifier construction in Japanese. In Robert Van Valin Jr. (ed.), *Investigations of the syntax-semantics-pragmatics interface*, 285–304. Amsterdam: John Benjamins. DOI: 10.1075/slcs.105.21shi.

Shimojo, Mitsuaki. 2009. Focus structure and beyond: Discourse-pragmatics in Role and Reference Grammar. In Lilián Guerrero, Sergio Ibáñez Cerda & Valeria A. Belloro (eds.), *Studies in Role and Reference Grammar*, 111–141. Universidad Nacional Autónoma de México.

Shimojo, Mitsuaki. 2010. The left periphery and focus structure in Japanese. In Wataru Nakamura (ed.), *Proceedings of the 10th international conference on Role and Reference Grammar*, 315–335.

Shimojo, Mitsuaki. 2011. The left periphery and focus structure in Japanese. In Wataru Nakamura (ed.), *New perspectives in Role and Reference Grammar*, 266–93. Newcastle upon Tyne: Cambridge Scholars Publishing.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R. M. W. Dixon (ed.), *Grammatical categories in Australian languages*. Australian Institute of Aboriginal Studies.

Smith, Carlota S. 1997. *The parameter of aspect*. 2nd edn. Dordrecht: Kluwer Academic Publishers. DOI: 10.1007/978-94-011-5606-6.

Thurmann, Robert. 1975. Chuave medial verbs. *Anthropological linguistics* 17(7). 342–352.

Toivonen, Ida. Forthcoming. Pronoun incorporation. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Tortora, Christina. 2014. *A comparative grammar of Borgomanerese*. Oxford/New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780199945627.001.0001.

Vallduví, Enric. 1992. *The informational component*. New York: Garland.

Van Valin, Robert, Jr. 1990. Semantic parameters of split intransitivity. *Language* 66. 221–260. DOI: 10.2307/414886.

Van Valin, Robert, Jr. 1993. *Advances in Role and Reference Grammar*. Amsterdam/Philadelphia: John Benjamins.

Van Valin, Robert, Jr. 1994. Extraction restrictions, competing theories and the argument from the poverty of the stimulus. In Susan D. Lima, Roberta Corrigan & Gregory K. Iverson (eds.), *The reality of linguistic rules*, 243–259. Amsterdam/Philadelphia: John Benjamins. DOI: 10.1075/slcs.26.17van.

Van Valin, Robert, Jr. 1998. The acquisition of WH-questions and the mechanisms of language acquisition. In Michael Tomasello (ed.), *The new psychology of language: Cognitive and functional approaches to language structure*, 221–249. LEA. DOI: 10.4324/9781315085678-9.

Van Valin, Robert, Jr. 1999. A typology of the interaction of focus structure and syntax. In Ekatarina Raxilina & Yakov G. Testelec (eds.), *Typology and linguistic theory: From description to explanation*, 511–524. Moscow: Languages of Russian Culture.

Van Valin, Robert, Jr. 2001. The acquisition of complex sentences: A case study in the role of theory in the study of language development. In *Proceedings from the panels of the chicago linguistic society's thirty-sixth meeting*, 511–531.

Van Valin, Robert, Jr. 2002. The development of subject-auxiliary inversion in English wh-questions: An alternative analysis. *Journal of Child Language* 29. 161–175. DOI: 10.1017/s0305000901004974.

Van Valin, Robert, Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/cbo9780511610578.

Van Valin, Robert, Jr. 2008. RPs and the nature of lexical and syntactic categories in Role and Reference Grammar. In Robert Van Valin Jr. (ed.), *Investigations of the syntax-semantics-pragmatics interface*, 161–178. Amsterdam: John Benjamins. DOI: 10.1075/slcs.105.14van.

Van Valin, Robert, Jr. 2009. Role and Reference Grammar as a framework for linguistic analysis. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 703–738. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199677078.013.0028.

Van Valin, Robert, Jr. 2018. Some issues regarding (active) accomplishments. In Rolf Kailuweit, Lisann Künkel & Eva Staudinger (eds.), *Applying and expanding Role and Reference Grammar*, 5–26. Freiburg: Freiburg Institute for Advanced Studies.

Van Valin, Robert, Jr. 2021. Cosubordination. In Robert Van Valin Jr. (ed.), *Challenges at the syntax-semantics-pragmatics interface: A Role and Reference Grammar perspective*, 241–254. Newcastle upon Tyne: Cambridge Scholars Publishing.

Van Valin, Robert, Jr. to appear(a). Grammatical aspects of language processing in the brain. In Delia Bentley, Ricardo Mairal-Usón, Wataru Nakamura & Robert Van Valin Jr. (eds.), *The Cambridge handbook of Role and Reference Grammar*. Cambridge, UK: Cambridge University Press.

Van Valin, Robert, Jr. to appear(b). Principles of Role and Reference Grammar. In Delia Bentley, Ricardo Mairal-Usón, Wataru Nakamura & Robert Van Valin Jr. (eds.), *The Cambridge handbook of Role and Reference Grammar*. Cambridge, UK: Cambridge University Press.

Van Valin, Robert, Jr. & Randy J. LaPolla. 1997a. Instructor's guide to Syntax: Structure, meaning, and function. Unpublished manuscript.

Van Valin, Robert, Jr. & Randy J. LaPolla. 1997b. *Syntax: Structure, meaning, and function*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/cbo9781139166799.

Van Valin, Robert, Jr. & David Wilkins. 1996. The case for 'effector': Case roles, agents, and agency revisited. In Masayoshi Shibatani & Sandra A. Thompson (eds.), *Grammatical constructions: Their form and meaning*, 289–322. Oxford: Oxford University Press.

Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca: Cornell University Press. DOI: 10.7591/9781501743726. Previously published in 1957 as 'Verbs and times', *Philosophical Review* 56: 143–160.

Vincent, Nigel. 2009. LFG and Dynamic Syntax: Two non-derivational theories. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '09 conference*, 587–603. Stanford: CSLI Publications.

Vincent, Nigel. 2016. Causatives in Latin and Romance. In J. N. Adams & Nigel Vincent (eds.), *Early and late Latin: Continuity or change?*, 294–312. Cambridge, UK: Cambridge University Press.

Vincent, Nigel. 2021. The 'of' word. In I Wayan Arka, Ash Asudeh & Tracy Holloway King (eds.), *Modular design of grammar: Linguistics on the edge*, 238–248. Oxford: Oxford University Press.

Vincent, Nigel. Forthcoming. Raising and control. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Weist, Richard. to appear. Functional acquisition processes in child language: Links to Role and Reference Grammar. In Delia Bentley, Ricardo Mairal-Usón, Wataru Nakamura & Robert Van Valin Jr. (eds.), *The Cambridge handbook of Role and Reference Grammar*. Cambridge, UK: Cambridge University Press.

Zaenen, Annie. Forthcoming. Information structure. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Name index

O’Connor, Robert, 274

Language index

Handbook of Lexical Functional Grammar

Set blurb on back with \BackBody{my blurb}