Chapter 9

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 'non-configurationality'. 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 'non-configurationality', 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.

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 $\overline{\text{V}}$.

[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), 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. This theory constrains how c-structure nodes can be annotated to produce f-structures, and shows the influence of Ken Hale's ideas about 'configurational' vs. '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 non-configurational was Malayalam, predominantly verb-final. Most of the SVO languages were analysed with distinct S and VP rules like these:

(2) a. S   ⟶   NP   VP
                  (↑ SUBJ)=↓   ↑=↓
     b. VP   ⟶   V   NP   …
                  ↑=↓   (↑ OBJ)=↓

But for Icelandic, Andrews used a flat S rule with subject first (consistent with Bresnan's 1982a: 354/296[4] 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 non-configurational 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 non-configurational language to which considerable attention was paid in this period; discussion of Warlpiri was included in Bresnan (1982a) and several other chapters in the Bresnan (1982b) collection.[5]

2.1 Malayalam

Non-configurational 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   ⟶   NP*   V

---

[4]The page before the '/' is for the journal version, the one after for the page in the 1982 volume.
[5]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 *-ʈə* 'is' to the verb. Some relevant examples are:
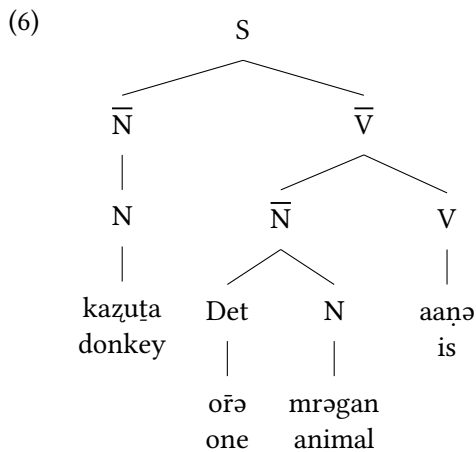
(5) Malayalam (Mohanan 1982: 528-529)

   a. Kuʈʈi iŋŋale   ammakkə   anaayey-aanə koʈuʈʈa-ʈə.
       child yesterday mother.DAT elephant-is   gave-it
       'It was an elephant that the child gave to the mother yesterday.'

   b. kuḷaʈʈil wecc-aaṇə jooṇinte kuʈʈi aanaye   ṇuḷḷia-ʈə.
       pond   at-is   John.GEN child elephant pinched-it
       'It was at the pond that John's child pinched the elephant.'

   c. *kuḷaʈʈil aaṇə weccə jooṇinte kuʈʈi aanaye   ṇuḷḷia-ʈə
       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 2023 [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: 513-534). An example is:

(6)

$$\begin{array}{c}
\text{S} \\
\overline{\text{N}} \qquad \overline{\text{V}} \\
\text{N} \qquad \overline{\text{N}} \qquad \text{V} \\
\text{kazụta} \qquad \text{Det} \quad \text{N} \qquad \text{aaṇə} \\
\text{donkey} \qquad \text{oṛə} \quad \text{mrəgan} \qquad \text{is} \\
\text{one} \quad \text{animal}
\end{array}$$

'The donkey is an animal.' (Mohanan 1982: 513–534)

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 similar 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 2023 [this volume]. One example is Complex Predicate constructions in Japanese (Ishikawa 1985, Matsumoto 1996), discussed in Andrews 2023 [this volume], and another is Warlpiri preverbs, considered shortly below.

Another important characteristic of Malayalam and other non-configurational 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 (↑ 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 non-configurational languages, and is subject to numerous variations in different languages. In English, anaphoric control is predominantly used with subjects of nonfinite verbs, but also applies in some other, more limited, circumstances. 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 non-configurationality 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 2023 [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, an $\overline{\text{N}}$, 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):[6]

(11)

(Simpson 1991: 283)
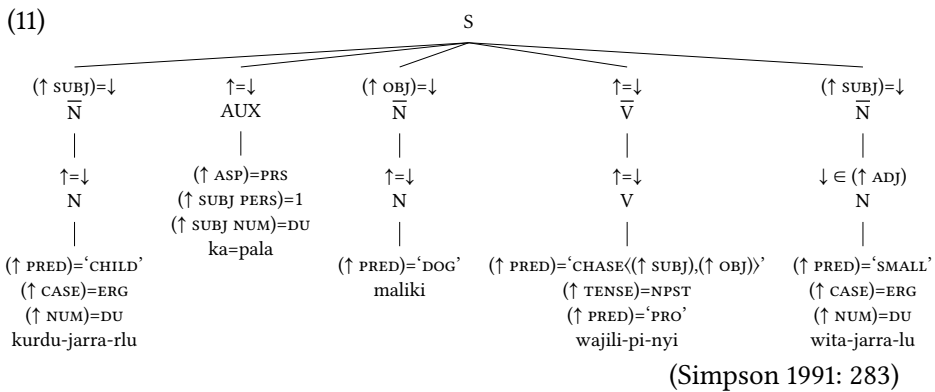
The annotations on the first and last $\overline{\text{N}}$s allow them to unify into a single f-substructure, the value of SUBJ:

---

[6]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)
$$
\begin{bmatrix}
\text{SUBJ} & \begin{bmatrix}
\text{PRED} & \text{`CHILD'} \\
\text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{`SMALL'} \\ \text{CASE} & \text{ERG} \end{bmatrix} \right\} \\
\text{CASE} & \text{ERG}
\end{bmatrix} \\
\text{PRED} & \text{`CHASE}\langle(\uparrow \text{SUBJ}),(\uparrow \text{OBJ})\rangle\text{'} \\
\text{ASP} & \text{PRS} \\
\text{TENSE} & \text{NPST} \\
\text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`DOG'} \end{bmatrix}
\end{bmatrix}
$$

Similarly to Malayalam, Warlpiri also makes extensive use of anaphoric control, although many arguments not expressed as NPs 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[7] 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'.[8] Finally come agreement 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.[9]

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

---

[7]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).

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

[9]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.

(13) S $\longrightarrow$ (AUX) α $(\alpha)^*$
       $\alpha = \overline{\text{V}}, \overline{\text{N}}$, Particle
       $(\uparrow \text{ASPECT}) \leftrightarrow (\uparrow \text{TENSE})$
       Assign $(\uparrow \text{G})=\downarrow$ 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 ($\uparrow=\downarrow$) or arguments bearing any grammatical function, providing a high degree of non-configurationality, including generating multiple $\overline{\text{N}}$ nodes with the $(\uparrow \text{SUBJ})=\downarrow$ 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 a position after the first phonological unit by a rule of sentence-phonology:

(14) Encliticization Rule:
      ]AUX $[\alpha]$ $[\alpha]^*$ $\rightarrow$ α+AUX $[\alpha]^*$
      (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),[10] nor in the discussion of discontinuous NPs in 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. These will not be discussed here.

Concluding our discussion of the first period, we find a basic distinction between configurational and non-configurational 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 ternary 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 it is still being explored (Kuhn 2023 [this volume]). The dating of the period is difficult, since the 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),[11] 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

---

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

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

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 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 & 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:[12]

---

[12]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.

(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.'

    b.
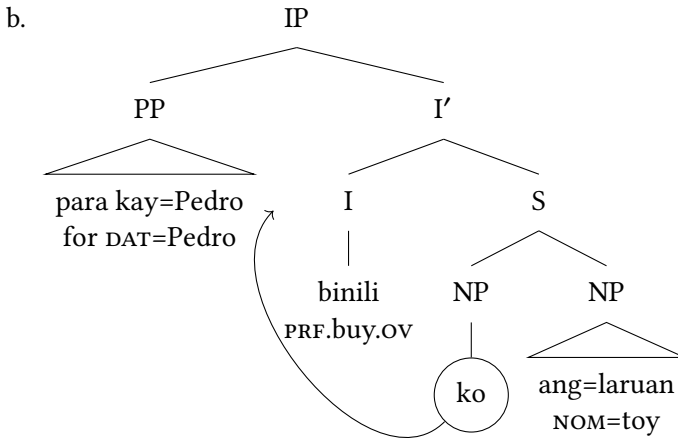
In this tree, the *ko* is ascribed to a position under S where a full NP argument could appear (initial in accord with the ordering tendencies noted above), with the arrow indicating some kind of clitic displacement to after the first constituent of IP. Another thing to note is the use of prime notation rather than bars, so I′ instead of Ī. 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 analyse 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)

```
                              IP
          ┌───────────────────┼──────────────┐
    ↓ ∈ (↑ ADJ)         (↑ A-SB)=↓          ↑=↓
    (↑ FOC)=↓                Cl              I′
        PP                    │         ┌─────┴─────┐
   ┌─────┴─────┐              I       ↑=↓         ↑=↓
para kay=Pedro               ko        I           S
    for DAT                            │           │
                                    binili    (↑ SUBJ)=↓
                                  PRF.buy.OV       NP
                                              ┌────┴────┐
                                             ang=laruan
                                             NOM=toy
```
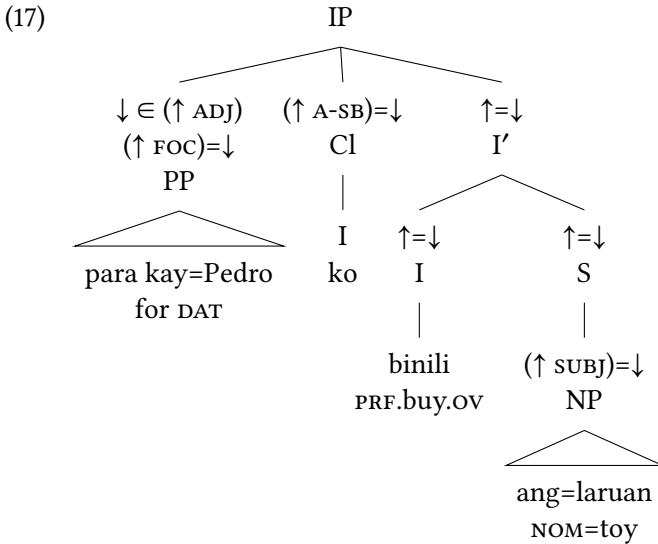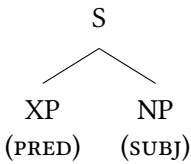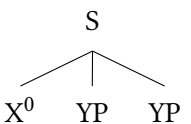
Given appropriate lexical entries, this will produce the following f-structure:[13]

(18)

$$
\begin{bmatrix}
\text{FOC} & \begin{bmatrix} \text{PRED 'FOR'} \\ \text{OBJ} \ [\text{PRED 'PEDRO'}] \end{bmatrix} \\[2ex]
\text{SUBJ} & \begin{bmatrix} \text{PRED 'TOY'} \\ \text{CASE NOM} \\ \text{SPEC DEF} \end{bmatrix} \\[3ex]
\text{PRED} & \text{'BUY}\langle (\uparrow \text{A-SB}) (\uparrow \text{SUBJ}) \rangle\text{'} \\
\text{ASP} & \text{PRF} \\[1ex]
\text{A-SB} & \begin{bmatrix} \text{PRED 'PRO'} \\ \text{PERS 1} \\ \text{NUM SG} \\ \text{CASE GEN} \end{bmatrix} \\[3ex]
\text{ADJUNCTS} & \{[\ ]\}
\end{bmatrix}
$$

---
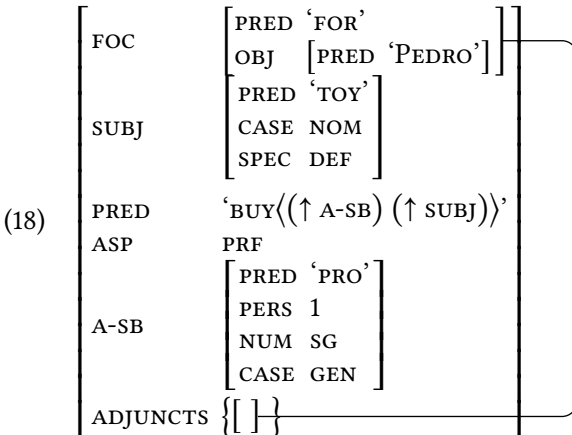
[13]This structure uses the older treatment of discourse functions such as FOCUS as grammatical functions in f-structure, along the lines of Bresnan (2001: 97-98) or Bresnan et al. (2016: 97). Kroeger (1993) does not provide any specific f-structures. For contemporary views, see Zaenen 2023 [this volume].

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):

(19) a.

```
            S
          /   \
        XP      NP
      (PRED)   (SUBJ)
```

   b.

```
          S
        / | \
     X⁰  YP  YP
```

Unfortunately, Kaufman (2010: 259-260), working within the Minimalist framework, finds that the clitic facts cited by Kroeger[14] 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:

---

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

(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:

(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 Randriamasi-
manana use it in their XLE grammar of Malagasy,[15] 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 argumenta-
tion 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 associa-
tion 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).

---

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

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: non-configurationality in Australian Languages

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

(23)   a. In Warlpiri, the verb does not appear in I.[16]

      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*.[17] 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

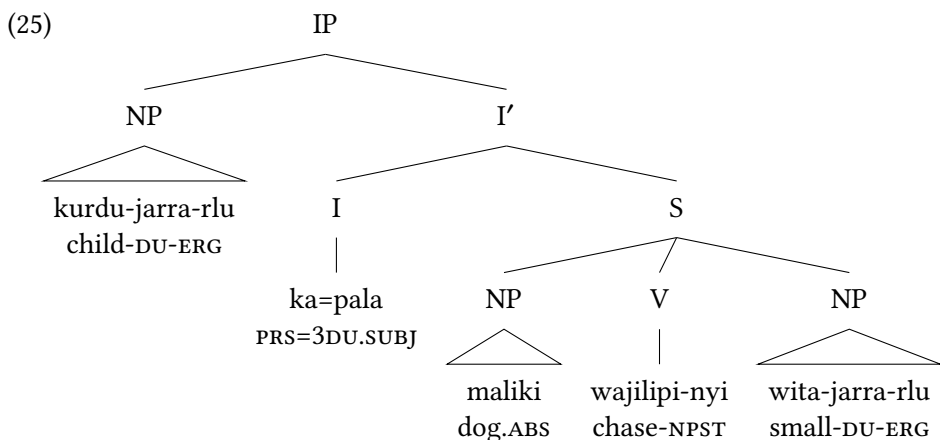   b. be phonologically well-formed as a word (initial *-lp* clusters are not allowed)

---

[16]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.

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

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 (1996) propose for Warlpiri, with an initial NP in the SPEC of IP position, and auxiliary material in the following I:

(25)

```
                           IP
                ┌──────────┴──────────┐
               NP                     I′
          ┌────┴────┐          ┌──────┴──────┐
     kurdu-jarra-rlu   I                      S
     child-DU-ERG      │          ┌───────────┼───────────┐
                       │         NP            V          NP
                    ka=pala   ┌───┴──┐         │       ┌───┴───┐
                    PRS=3DU.SUBJ maliki    wajilipi-nyi   wita-jarra-rlu
                              dog.ABS    chase-NPST    small-DU-ERG
```

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 & 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.):[18]

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

   a. Rambal-luwa-rnu=**rna=rla=jinta**   marlu-ku
      mistake-shoot-PST=1SG.SUBJ-1SG.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 or 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):

---

[18]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.

(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.[19]

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 & Bresnan (1996) 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.[20] 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:

---

[19]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.

[20]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.

(30)   Warlpiri (Legate 2008: 34)

Kuturu-ju   ka=npa=nyanu   nyarrpara-wiyi marda-rni?
nullanulla-ᴛᴏᴘ ᴘʀꜱ.ɪᴘꜰᴠ=2ꜱɢ.ɴᴏᴍ=ᴀɴᴀᴘʜᴏʙᴊ where-first   have-ɴᴘꜱᴛ
'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

   a. Kaji=ka=rna   nyarrpara-kurra ya-ni.
       ɴꜰᴀᴄᴛᴄ=ᴘʀꜱ.ɪᴘꜰᴠ=1ꜱɢ where-ᴀʟʟ   go-ɴᴘꜱᴛ
       'I might go somewhere.'/'Where might I go?' (Legate 2008: 17)

   b. Nyarrpara-kurra kaji=ka=rna   ya-ni.
       where-ᴀʟʟ   ɴꜰᴀᴄᴛᴄ=ᴘʀꜱ.ɪᴘꜰᴠ=1ꜱɢ go-ɴᴘꜱᴛ
       *'I might go somewhere.'/'Where might I go?' (Legate 2008: 17)

   c. Nyiya=rlangu   kaji=ka=rlu   nyina
       what-for.example ɴꜰᴀᴄᴛᴄ=ᴘʀꜱ.ɪᴘꜰᴠ=3ᴘʟ.ᴏʙᴊ be.ɴᴘꜱᴛ
       wampana-piya-ju.
       spectacled.hare.wallaby-like-ᴛᴏᴘ
       'What ones for example might be like this spectacled hare wallaby?'
       (Legate 2008: 18)

   d. Kaji=lpa=ngu   wanti-yarla nyiya-rlangu
       ɴꜰᴀᴄᴛᴄ=ᴘꜱᴛ.ɪᴘꜰᴠ=2ꜱɢ.ᴏʙᴊ fall=ɪʀʀ   what-for.example
       milpa-kurra.
       eye-ᴀʟʟ
       '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 & 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.[21] 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).

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 on on either side of the auxiliary, while in Legate's, interrogatives 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 non-configurational, 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 and 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.'

---

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

b. Opera, net drugogo vida musykal'nogo iskusstva, kotoryj privlekal
   opera   not other   type musical   art   which   attract
   by   k sebe takoe vnimanie.
   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)   EP   ⟶   XP   CP
                 (↑ E-TOP)=↓   ↑=↓

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,[22] with some kind of further restriction, perhaps essentially semantic, preventing them from occurring in embedded positions (King 1995: 106):

(34)   CP   ⟶   XP   CP
                 (↑ E-TOP)=↓   ↑=↓

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, ↑=↓ annotations omitted):[23]

(35)   Russian

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

---

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

[23]In the tree, the annotation over the first constituent should be ↓ ∈ (↑ ADJUNCTS), but in the structure I give the original.

b.

$$
\begin{array}{c}
\text{CP} \\
\end{array}
$$

```
                          CP
         ┌─────────────────┼──────────────┐
   (↑ ADJ)=↓        (↑ SUBJ)=↓          C′
   ↓ ∈ (↑QFOC)      ↓ ∈ (↑QFOC)          │
      AP                NP               IP
      │                 │                │
    kogda              kto               I′
    when               who        ┌──────┴──────┐
                                  I            VP
                                  │             │
                                udaril          V′
                                hit.PST         │
                                           (↑ OBJ)=↓
                                               NP
                                               │
                                             Borisa
                                            Boris.ACC
```
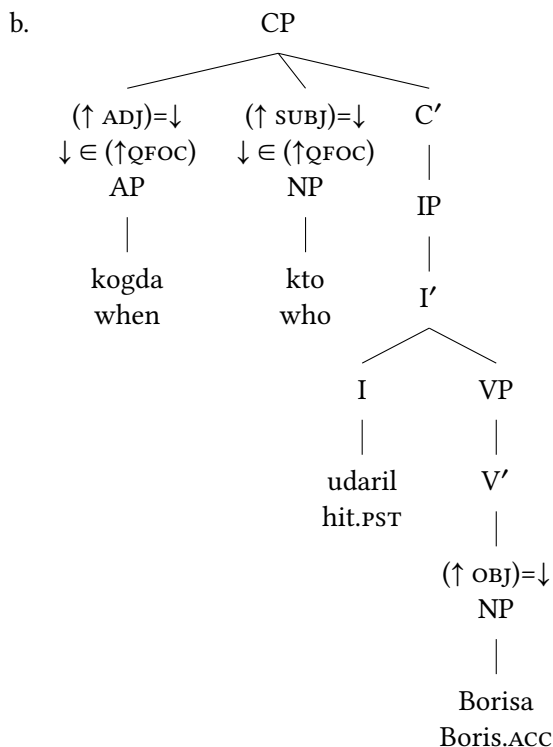
Yes-no questions can be marked either by intonation, or with the marker *li*, which appears in second position, after either an XP or the verb. An XP in front of *li* is interpreted as a focus, with the body of the question presupposed (King 1995: 236–237, King 1994), and items in SPEC of IP are Topics, 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.

```
                           CP
              ┌─────────────┴──────────────┐
        (↑ obj)=↓                          C′
        ↓ ∈ (↑qfoc)            ┌────────────┴────────────┐
           NP                  C                         IP
            │                  │           ┌─────────────┴──────┐
            N                  li     (↑ subj)=↓              I′
            │                  Q      ↓ ∈ (↑ topic)            │
          knigu                          NP                    I
         book.ACC                         │                    │
                                         ona              pročitala
                                       she.NOM            read.PST
```
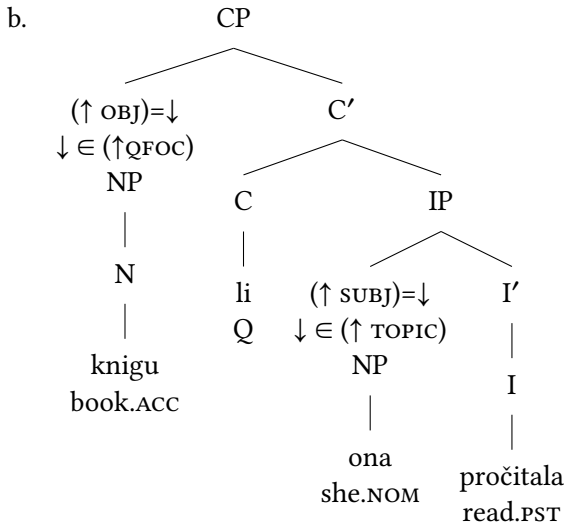
If the yes-no question has no focus, then the verb appears before *li*, and King proposes that the verb is adjoined to C:

(37)

```
                              CP
                               │
                              C′
                  ┌────────────┴──────────────┐
                  C                           IP
          ┌───────┴───────┐       ┌───────────┴────────────┐
          I               C   (↑ subj)=↓                 I′
          │               │   ↓ ∈ (↑ topic)               │
      pročitala          li       NP                      │
      read.PST           Q         │                  (↑ obj)=↓
                                  ona                     NP
                                she.NOM                    │
                                                           N
                                                           │
                                                        knigu
                                                       book.ACC
```
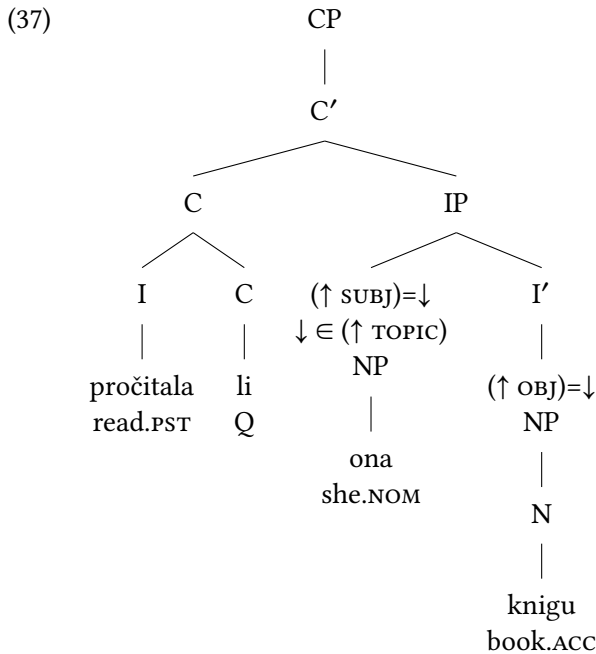
This solution avoids the issue of putting a nonmaximal projection in specifier position, and is independently motivated by the absence of the focus-presupposi-
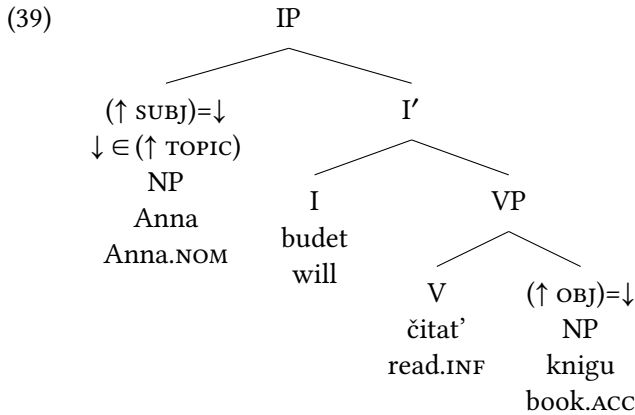
tion articulation that appears when there is an XP in the specifier position. Formally, it could also be applied to Warlpiri, but without the additional motivation, insofar as is now known, and the problem of insertion of the auxiliary into combinations of verb and preverb would remain unsolved.

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. IP $\longrightarrow$ XP*, I′
                    ($\uparrow$ GF)=$\downarrow$ $\uparrow$=$\downarrow$
                    $\downarrow \in (\uparrow$ DF)
       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)

```
                          IP
              ┌───────────┴───────────┐
        (↑ SUBJ)=↓                    I′
        ↓ ∈ (↑ TOPIC)        ┌────────┴────────┐
           NP                I                 VP
          Anna             budet          ┌────┴────┐
        Anna.NOM            will          V      (↑ OBJ)=↓
                                        čitat'      NP
                                       read.INF    knigu
                                                 book.ACC
```
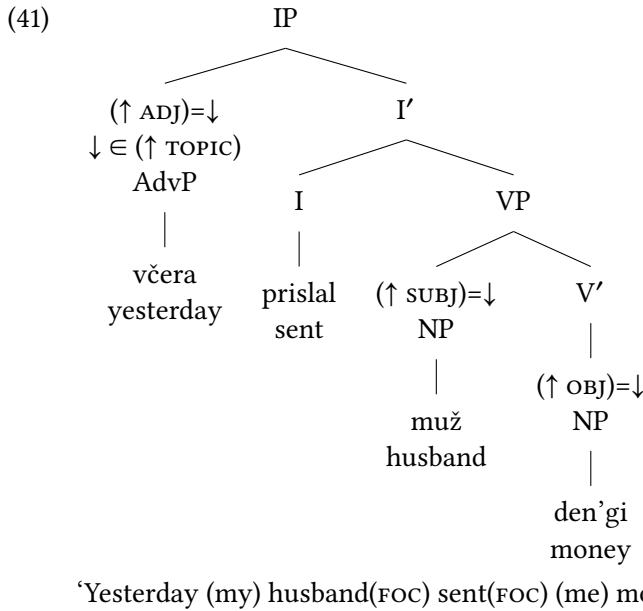
This leads on to the structure of VP.

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):[24]

(40)   a. VP ⟶   (XP)   V′
                  (↑ SUBJ)=↓   ↑=↓
       b. V′ ⟶   V   XP*
                   ↑=↓   (↑ GF)=↓

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):

---

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

(41)

```
                          IP
              ┌───────────┴───────────┐
        (↑ ADJ)=↓                     I′
        ↓ ∈ (↑ TOPIC)          ┌──────┴──────┐
           AdvP                I            VP
            │                  │        ┌────┴────┐
          včera             prislal  (↑ SUBJ)=↓   V′
         yesterday           sent      NP         │
                                        │    (↑ OBJ)=↓
                                       muž     NP
                                     husband    │
                                              den'gi
                                              money
```

'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)[25] resembles Warlpiri in arguably having only one functional projection,[26] 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).

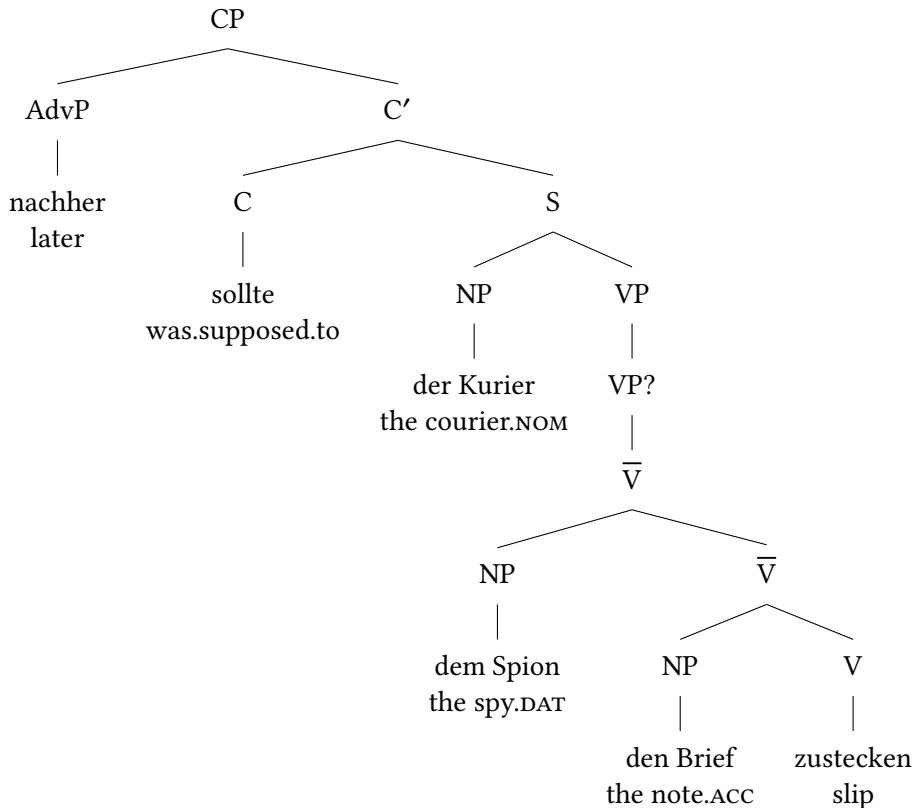[25] A revised version of her 1996 Stanford dissertation of the same title.

[26] 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′",[27] 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 an NP VP structure. A sample main clause with the auxiliary *sollte* is:

(43)   German

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

b.

```
                        CP
             ┌──────────┴──────────┐
           AdvP                    C′
            │            ┌─────────┴─────────┐
          nachher        C                   S
          later          │           ┌───────┴───────┐
                       sollte        NP              VP
                  was.supposed.to    │               │
                                  der Kurier         VP?
                                 the courier.NOM      │
                                                      V̄
                                              ┌───────┴───────┐
                                             NP               V̄
                                              │        ┌───────┴───────┐
                                           dem Spion  NP              V
                                          the spy.DAT  │               │
                                                    den Brief       zustecken
                                                   the note.ACC       slip
```

---

[27]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 $\overline{\text{V}}$ 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', which also has other meanings) appear syntactically to be plausibly taken to be the complement of the auxiliary *sollen* in C, whose VP appears 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 annotated '↑=↓', 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
that the courier.NOM later   the spy.DAT the note.ACC slip
sollte
was.supposed.to
'that the courier was supposed to slip the note to the spy later'

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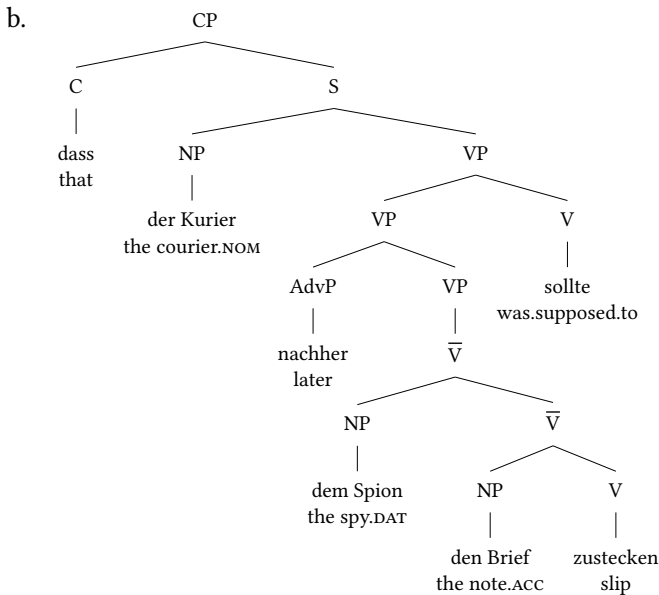
Here we see the complementizer occupying C, and the VP complement to the auxiliary *sollte*, at the end.

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:[28]

(46) Korean

  a. Cwuni-*g*a   kong-ul *j*aypalli *j*a*b*-a.
     Cwuni-NOM ball-ACC quickly catch-PRS
     'Cwuni catches balls quickly.'

---

[28]Choi uses underlining to indicate the non-phonemic voicing, rather than different segmental phonetic symbols.
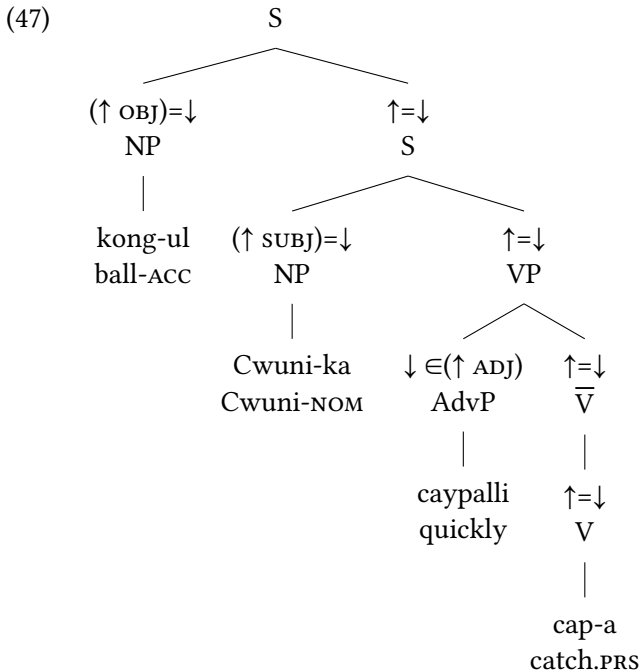
b. Kong-ul Cwuni-*ga* caypalli *ja*n-a.
 ball-ACC Cwuni-NOM quickly catch-PRS
 'Balls Cwuni catches quickly.'

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 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)

```
                            S
              ┌─────────────┴─────────────┐
        (↑ OBJ)=↓                        ↑=↓
           NP                             S
            │               ┌─────────────┴─────────────┐
        kong-ul        (↑ SUBJ)=↓                       ↑=↓
        ball-ACC          NP                            VP
                           │                   ┌─────────┴────────┐
                       Cwuni-ka          ↓ ∈(↑ ADJ)             ↑=↓
                       Cwuni-NOM            AdvP                 V̅
                                             │                   │
                                         caypalli               ↑=↓
                                         quickly                 V
                                                                 │
                                                               cap-a
                                                             catch.PRS
```

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 V̅ is a difficult one, which the new phrase structure theory of Lowe & Lovestrand (2020) might allow us to eliminate and thereby solve, but this 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 preceding 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 non-configurational, 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 'non-projecting words' of Toivonen (2001), discussed in Belyaev 2023 [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 the older LFG literature assumed, (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).[29] 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$_\theta$; 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:

---

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

(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 a 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 gen-
eral 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 and
details of the particular sentence. I suggest that this involves a shift from a tra-
ditional 'syntactic' notion of sentence-hood to a more 'semantic one', since glue
semantics (Asudeh 2023 [this volume]) can connect these multiple superficial
syntactic structures to a single semantic type. Further reductions in the complex-
ity 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':[30]

(53)

```
              IP
            /    \
        NP        VP
         |         |
       Mary       V⁰
                   |
                 swims
```
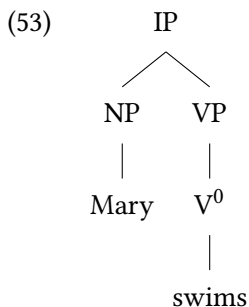
(Bresnan et al. 2016: 120)

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 $\uparrow = \downarrow$ equation and so will 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:
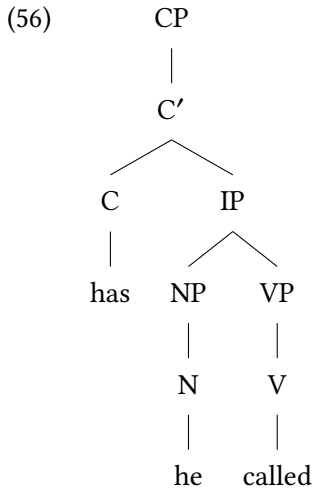
(54)   Has he called?

(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:

---

[30]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.

(56)

```
              CP
              │
              C′
             ╱  ╲
            C    IP
            │   ╱  ╲
           has NP   VP
                │    │
                N    V
                │    │
               he  called
```

The IP provides the location of the subject, which has no alternative locations in English finite clauses.[31] The conditional clause in (55b) has the same structure, with the preposed auxiliary replacing the overt complementizer that appears in (55a).

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.
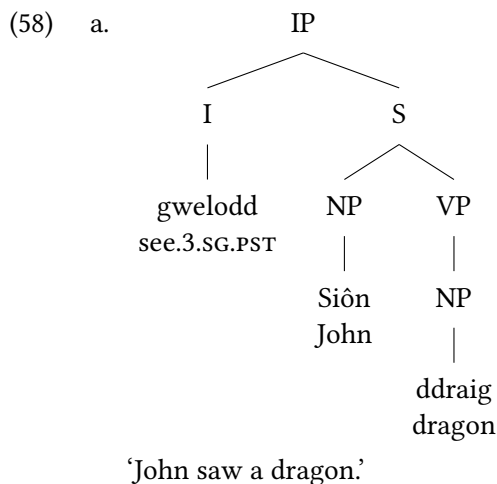
---

[31]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.
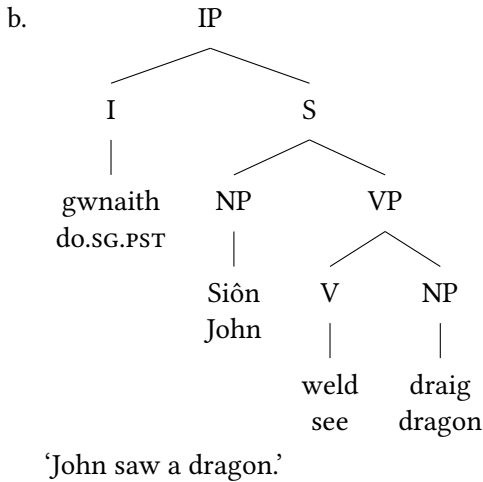
5.1 Welsh

The basic structure for a main clause is:

(57)

```
              IP
              │
              I′
           ┌──┴──┐
           I      S
           │     ┌─┴─┐
        Aux/V   NP   VP
```
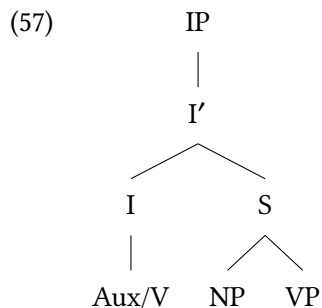
If there is no auxiliary, the finite verb appears at the front of the sentence, in the I position, as shown in (58a) below, very similarly to King's (1995) analysis of Russian. But if there is an auxiliary, the auxiliary appears in I, the verb initially in V of VP, again similarly to Russian, as shown 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.

```
                    IP
                ┌───┴───┐
                I        S
                │      ┌──┴──┐
             gwelodd  NP    VP
          see.3.SG.PST │     │
                      Siôn   NP
                      John   │
                           ddraig
                           dragon
```

'John saw a dragon.'

b.

```
                        IP
                   ┌────┴────┐
                   I         S
                   │      ┌──┴──┐
                gwnaith   NP    VP
                do.SG.PST │   ┌──┴──┐
                         Siôn V    NP
                         John │     │
                             weld  draig
                             see   dragon
```

'John saw a dragon.'

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)[32] 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

---

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

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.[33] 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 'auxiliaries'.

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 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):

---

[33]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.

(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 accommodated by including an appropriate constraining equation in the c-structure rule:

(63)   NP ⟶   S
                         ↑=↓
               (↓ VFORM) =$_c$ ING

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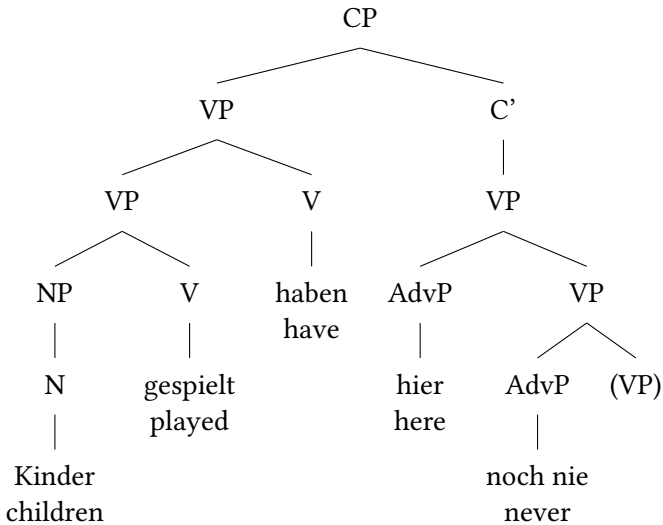
(66)   German (Berman 2003: 36)
       Kinder   gespielt haben hier noch nie.
       children played   have   here still   never
       'Children have never played here.'

Her solution was to introduce all arguments, including subjects, under VP in nested VPs, her VP not being clearly distinct from Choi's $\overline{\text{V}}$, with a complex verb at the bottom, as discussed in a considerable amount of literature in different frameworks.[34]

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

---

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

(67)   German

```
                                    CP
                          ┌─────────┴─────────┐
                         VP                   C'
                    ┌─────┴─────┐             │
                   VP           V             VP
              ┌─────┴─────┐     │        ┌─────┴─────┐
             NP           V   haben     AdvP         VP
              │           │    have      │      ┌─────┴─────┐
              N         gespielt        hier   AdvP       (VP)
              │          played         here    │
           Kinder                             noch nie
          children                             never
```

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 'scram-
bling', 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 non-configurational, 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 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 non-configurational 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. This is that nonconfigurational S is what ensues diachronically when a language makes such extensive use of discourse-conditioned preposing that the learner gives up trying to analyse it, but instead returns as c-structure something that is essentially a list of fragments, similar to what the XLE system (Forst & King 2023 [this volume]) does when it can't find a parse. But the syntax does find an f-structure for the c-structure fragment list appearing under S. The idea of non-configurational 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 non-configurational S has no category feature value.

Acknowledgements

References

Akmajian, Adrian. 1984. Sentence types and the form-function fit. *Natural Language & Linguistic Theory* 2(1). 1–23. DOI: 10.1007/bf00233711.

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. 2023. Complex predicates. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*, 267–308. Berlin: Language Science Press. DOI: 10.5281/zenodo.10185946.

Asudeh, Ash. 2023. Glue semantics. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*, 651–697. Berlin: Language Science Press. DOI: 10.5281/zenodo.10185964.

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. 2023. Core concepts of LFG. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*, 23–96. Berlin: Language Science Press. DOI: 10.5281/zenodo.10185936.

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 1982b, 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.

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. *Linguistic Inquiry* 18(2). 173–237.

Dalrymple, Mary, Ronald M. Kaplan, John T. III Maxwell & 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. 2023. Computational implementations and applications. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*, 1083–1123. Berlin: Language Science Press. DOI: 10.5281/zenodo.10185986.

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. Unpublished manuscript, University of Stuttgart.

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, Reinhart & 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. *X̄ 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. 2023. LFG, Optimality Theory and learnability of languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*, 961–1032. Berlin: Language Science Press. DOI: 10.5281/zenodo.10185980.

Laughren, Mary. 1989. The configurationality parameter and Warlpiri. In László Marácz & Pieter Muysken (eds.), *Configurationality*, 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. DOI: 10.1093/oso/9780198851158.001.0001.

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. 2023. LFG and Australian languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*, 1253–1287. Berlin: Language Science Press. DOI: 10.5281/zenodo.10185998.

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*, 262–294. 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. 2023. Information structure. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*, 823–853. Berlin: Language Science Press. DOI: 10.5281/zenodo.10185972.

Zaenen, Annie & Ronald M. Kaplan. 1995. Formal devices for linguistic general-
izations: 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).