Chapter 34

HPSG and Lexical Functional Grammar

Stephen Wechsler
The University of Texas

Ash Asudeh
University of Rochester

Here is the abstract: concrete gets you through abstract better than abstract gets you through concrete

1 Introduction

Head-Driven Phrase Structure Grammar is similar in many respects to its cousin framework, Lexical Functional Grammar or LFG (Bresnan et al. 2015; Dalrymple 2001). Both HPSG and LFG are lexicalist frameworks in the sense that they distinguish between the morphological system which creates words, and the syntax proper which combines those fully inflected words into phrases and sentences. Both frameworks assume a lexical theory of argument structure (Müller & Wechsler 2014) in which verbs and other predicators come equipped with valence structures indicating the kinds of complements and other dependents that the word is to be combined with. Both theories treat control (equi) and raising as a lexical property of certain control or raising predicates. Both representational systems are based on unification grammar (Kay 1984a), employing directed graphs that are often represented in the form of recursively embedded feature structures. Phonologically empty nodes of the constituent structure are avoided in both theories, with the gaps appearing in long-distance dependencies as the sole exception in some analyses, and complete elimination of empty categories even in those cases, in others.

At the same time, there are interesting differences. Each theory makes available certain representational resources that the other theory lacks. LFG has out-

Stephen Wechsler & Ash Asudeh. 2020. HPSG and Lexical Functional Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 1241–1263. Berlin: Language Science Press. DOI:??

put filters in the form of *constraining equations*, HPSG does not. HPSG's feature structures are *typed*, those of LFG are not. The feature descriptions (directed graphs) are fully integrated with the phrase structure grammar in the case of HPSG, while in LFG they are intentionally separated in an autonomous level of representation in the form of a *functional structure* or f-structure. These differences lead some linguists to feel that certain types of generalization are more perspicuously stated in one framework than the other. Because LFG's functional structure is autonomous from the constituent structure whose terminal yield gives the order of words in a sentence, that functional structure can instead serve as a representation of the grammatical functions played by various components of a sentence. This makes LFG more amenable to a functionalist motivation, and also provides a standard representation language for capturing the more cross-linguistically invariant properties of syntax. Meanwhile, HPSG is more deeply rooted in phrase structure grammar, and thus provides a clearer representation of the locality conditions that are important for the proper functioning of grammars.

This chapter presents a comparison of the two theories with an emphasis on contrasts between the two. It is organized by grammatical topic.

2 Phrases and Endocentricity

A phrasal node shares certain grammatical features with specific daughters, such as the HEAD features that it shares with the head daughter. In HPSG this is accomplished by means of structure-sharing (reentrancies) in the immediate dominance schemata and other constraints on local sub-trees such as the Head Feature Principle. LFG employs essentially the same mechanism for feature sharing in a local sub-tree but implements it slightly differently. Each node in a phrase structure is paired with a so-called functional structure or *f-structure*, which is formally a set of attribute-value pairs. It is through the f-structure that the nodes of the phrase structure share features. The phrase structure is referred to as *c-structure*, for categorial or constituent structure, in order to distinguish it from f-structure. The grammar, in the form of a standard rewriting system, directly generates only c-structures, not f-structures. Those c-structure rules introduce equations that form a projection function from c-structure to f-structure. For example, the phrase structure grammar in 1 and lexicon in 2 generate the tree in 3.

(1) a. $\text{S} \rightarrow \underset{(\uparrow \text{SUBJ})=\downarrow}{\text{NP}} \quad \underset{\uparrow=\downarrow}{\text{VP}}$

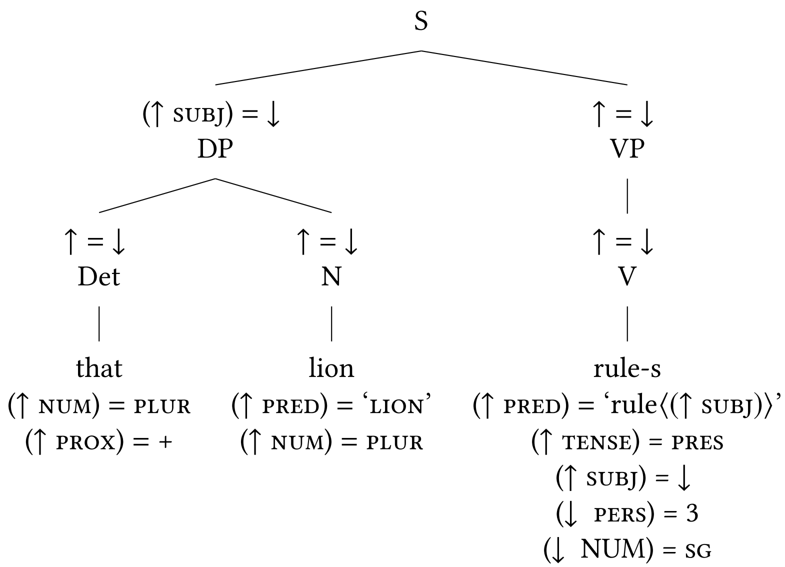
b. NP $\rightarrow$ $\left(\begin{array}{c}\text{Det}\\ \uparrow=\downarrow\end{array}\right)$ $\begin{array}{c}\text{N}\\ \uparrow=\downarrow\end{array}$

c. VP $\rightarrow$ $\begin{array}{c}\text{V}\\ \uparrow=\downarrow\end{array}$ $\left(\begin{array}{c}\text{NP}\\ (\uparrow \textsc{obj})=\downarrow\end{array}\right)$

(2) a. *lion*: N (↑ ‘)’ = ‘lion’
(↑ NUM) = SG

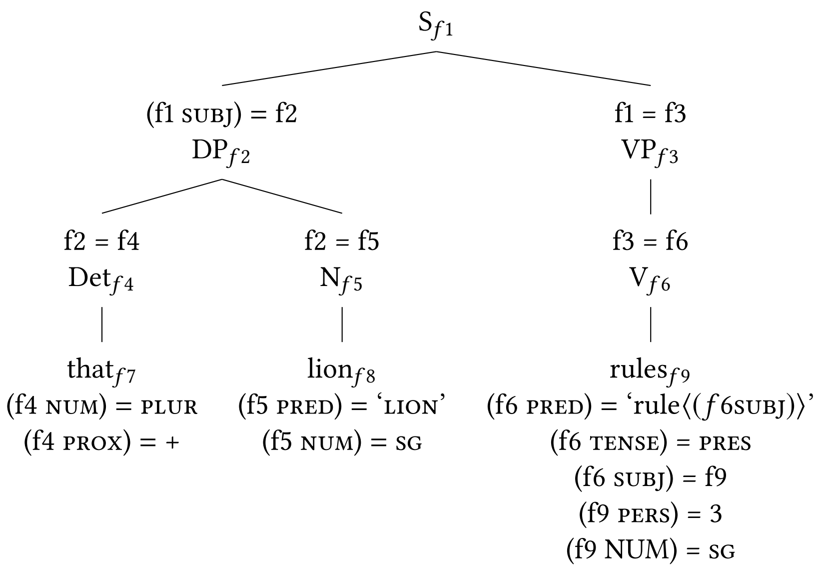
b. *rule*: V (↑ ‘)’ = ‘rule⟨(↑ SUBJ)’
-*s*: *infl* (↑ TENSE) = PRES
(↑ SUBJ) = ↓
(↓ PERS) = 3
(↓ NUM) = SG

(3)



Each node in the c-structure maps to a function, that is, to a set of attribute-value pairs. Within the equations, the up and down arrows are metavariables over function names, interpreted as follows: the up arrow refers to the function to which the mother node maps, and the down arrow refers to the function that its own node maps to. To derive the f-structure from 3, we instantiate the metavariables to specific function names and solve for the function associated with the root node (here, S). In 4 the function names *f1, f2,* etc. are subscripted to the node labels. The arrows have been replaced with those function names.

(4)



Collecting all the equations from this tree and solving for *f1*, we arrive at the f-structure in 5:

(5)
$$\left[\begin{array}{ll} \textsc{subj} & \left[\begin{array}{ll} \textsc{pred} & \textit{`lion'} \\ \textsc{num} & \textit{sg} \\ \textsc{pers} & \textit{3} \\ \textsc{prox} & + \end{array}\right] \\ \textsc{pred} & \textit{`rule}\langle(\uparrow \textsc{subj})\rangle\textit{'} \\ \textsc{tense} & \textit{pres} \end{array}\right]$$

Since the up and down arrows refer to nodes of the local subtree, LFG annotated phrase structure rules like those in 1 can often be directly translated into HPSG immediate dominance schemata and principles constraining local subtrees. By way of illustration, let FS (for *f-structure*) be an HPSG attribute corresponding to the f-structure projection function. Then the LFG rule in 6a (repeated from 1a above) is equivalent to the HPSG rule in 6b:

(6) a. $S \;\rightarrow\; \underset{(\uparrow \textsc{subj})=\downarrow}{DP} \quad \underset{\uparrow=\downarrow}{VP}$

b. $S[\textsc{fs}\ \boxed{1}] \rightarrow\ DP[\textsc{fs}\ \boxed{2}]\quad VP[\textsc{fs}\ \boxed{1}[\textsc{subj}\ \boxed{2}]]$

Let us compare the two representations with respect to heads and dependents.

Taking heads first, the VP node annotated with $\uparrow = \downarrow$ is an *f-structure head*, meaning that the features of the VP are identified with those of the mother S. This effect is equivalent to the tag $\boxed{1}$ in 6b. Hence $\uparrow = \downarrow$ has an effect similar to HPSG's Head Feature Principle. However, in LFG the part of speech categories and their projections such as N, V, Det, NP, VP, DP, etc. belong to the c-structure and not the f-structure. As a consequence those features are not subject to sharing, and any principled correlations between such categories, such as the fact that N is the head of NP, V the head of VP, C as head of CP, and so on, are instead captured in an explicit version of (extended) X-bar theory applying to the c-structure. The LFG based theory of endocentricity is considerably weaker (more permissive) than what is typically found in most transformation based grammars. The version of extended X-bar theory in Bresnan et al. (2015: chapter 6) assumes that all nodes on the right side of the arrow of the phrase structure rule are optional, with many unacceptable partial structures ruled out in the f-structure instead. Also not all structures need to be endocentric (i.e. not all structures have a head daughter in c-structure). The LFG category S shown in 6a is inherently exocentric, lacking a c-structure head, and is used for the analysis of copulaless clauses. (English is also assumed to have endocentric clauses of category IP, where an auxiliary verb of category I (for Inflection) serves as the c-structure head.) S is also used for flat structures in non-configurational clauses found in languages such as Warlpiri.

Functional projections like DP, IP, and CP are typically assumed to form a 'shell' over the lexical projections NP, VP, AP, and PP (plus CP can appear over S). In fact this idea of extending X-bar to functional categories has its origin in the LFG work of the late Yehuda Falk (Falk 1984), from which it then spread to transformational theories. This is formally implemented by having the functional head (such as Det) and its lexical complement (such as NP) be f-structure co-heads. See for example the DP *that lion* in 3, where Det and N are both annotated with $\uparrow = \downarrow$. The DP, Det, and N nodes all map to the same f-structure, namely the subsidiary structure serving as the value of SUBJ (see 5). What makes this unification possible is that function words lack a PRED (for 'predicate') feature that would otherwise indicate a semantic form. Content words such as *lion* have such a feature ([PRED 'LION']), and so if the Det had one as well then they would clash in the f-structure. Note more generally that the f-structure flattens out much of the hierarchical structure of the corresponding c-structure.

Complementation works a little differently in LFG from HPSG. Note that the LFG rule 6a indicates the SUBJ grammatical function on the subject NP node, while the pseudo-HPSG rule 6b indicates the SUBJ function on the VP functor selecting the subject. A consequence of the use of functional equations in LFG is that a grammatical relation such as SUBJ can be locally associated with its formal exponents, whether a configurational position in phrase structure (as in 3), head-marking (agreement), or dependent marking (case). A subject-marking case affix can introduce a so-called 'inside out' functional designator, (SUBJ ↑), which requires that the f-structure of the DP bearing that case ending be the value of a SUBJ attribute (Nordlinger 1998).[1] This aspect of LFG representations makes it convenient for functionalist and typological work on grammatical relations.

3 Valence

In LFG a lexical predicator such as a verb selects its complements via f-structure rather than c-structure. A transitive verb selects a SUBJ and OBJ, which are features of f-structure, but it cannot select for the category 'DP' because such part of speech categories belong only to c-structure. For example the verb stem rule in 2b has a PRED feature whose value contains (↑ SUBJ), which has the effect of requiring a SUBJ function in the f-structure. The f-structure (shown in 5) is built using the defining equations, as described above. Then that f-structure is checked against any *existential constraints* such as the expression (↑ SUBJ), which requires that the f-structure contain a SUBJ feature. That constraint is satisfied, as shown in 5. Moreover, the fact that (↑ SUBJ) appears in the angled brackets means that it expresses a semantic role of the 'rule' relation, hence the SUBJ value is required to contain a PRED feature, which is satisfied by the feature [PRED 'LION'] in 5.

Selection for grammatical relations instead of formal categories enables LFG to capture the flexibility in the expression of a given grammatical relation described at the end of the previous section. As noted there, in many languages the subject can be expressed either as an independent DP phrase as in English, or as a pronominal affix on the verb. As long as the affix introduces a PRED feature and is designated by the grammar as filling the SUBJ relation, then it satisfies the subcategorization requirements imposed by a verb. A more subtle example of flexible expression of grammatical functions can be seen in English constructions where an argument can in principle take the form of either a DP (as in 7a) or a clause (as in 7b) (example 7 is from Bresnan et al. (2015) PAGE).

(7) a. That problem, we talked about for days.
    b. That he was sick, we talked about for days.

[1] For function f, attribute a and value v, $(f\ a) = v$ iff $(a\ v) = f$.

c. We talked about that problem for days.

d. *We talked about that he was sick for days.

The variant of *talk* taking an *about*-PP selects neither a DP nor a clausal complement, but rather an object (OBJ) of an oblique (OBL_{about}) function:

(8) *talk*: V (↑ ‘)’ = ‘talk-about⟨(↑ SUBJ)(↑ OBL_{about}OBJ)⟩’

It is not the verb but the local c-structure environment that conditions the category of that argument: the canonical object position right-adjacent to *about* can only house a DP, while the topic position allows either DP or clause (as seen by comparing 7c and 7d). In LFG the grammatical functions such as SUBJ and OBJ represent equivalence classes across various modes of c-structure expression.

HPSG captures this variability in the expression of arguments in essentially the same way as LFG, despite some terminological differences. The HPSG correspondent of the LFG SUBJ specification is not an HPSG SUBJ valence list item, but rather the item of the ARG-ST list that the (optional) SUBJ item maps to, when it appears. The same applies to complements. The disjunction between DP and clausal expression of an argument is encoded in the ARG-ST; the restriction to DP (observed in examples 7c,d) is encoded on the relevant valence list.

4 Head mobility

The lexical head of a phrase can sometimes appear in an alternative position apparently outside of what would normally be its phrasal projection. Assuming that an English finite auxiliary verbs is the (category I) head of its (IP) clause, then that auxiliary appears outside its clause in a yes/no question:

(9) a. [$_{IP}$ she is mad].

b. Is [$_{IP}$ she GAP mad]?

Transformational grammars capture the systematic relation between these two structures with a head-movement transformation that leaves the source IP structure intact, with a trace replacing the moved head. The landing site of the moved clausal head is often assumed to be C, the complementizer position, as motivated by complementarity between the fronted verb and a lexical complementizer observed most strikingly in German but also found in other languages, including some English constructions such as the following:

(10) a. I wonder whether [$_{IP}$ she is mad].

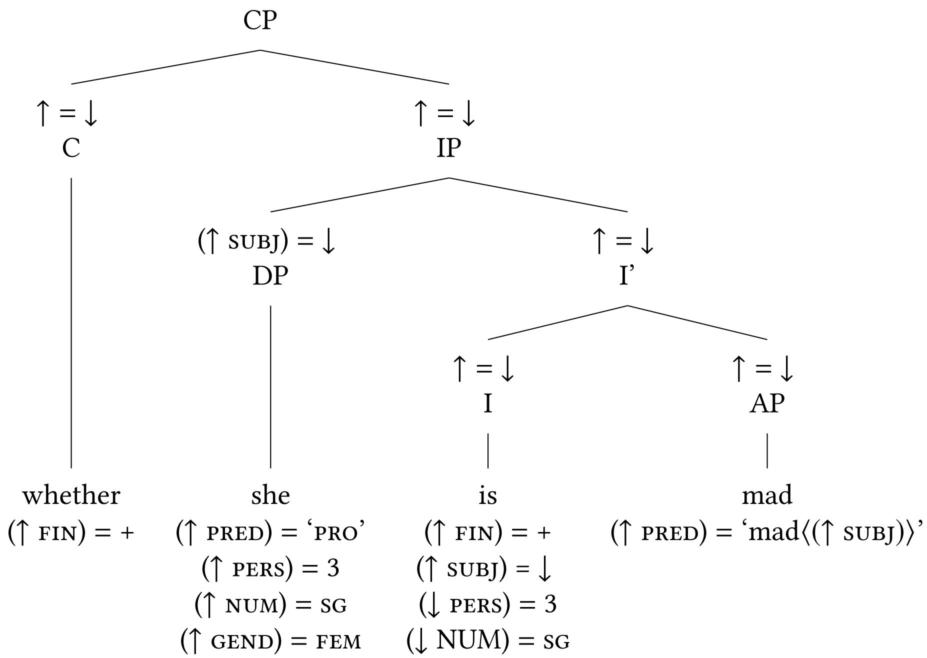
b. I wonder, is [$_{IP}$ she GAP mad]?

c. *I wonder whether is she mad.

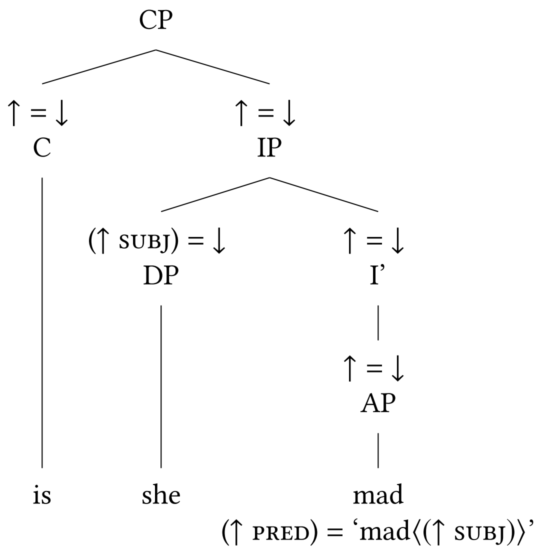
In HPSG the sentences in 9 are treated as displaying two distinct structures generated by the grammar. For example, assuming ternary branching in 9b then the subject DP *she* and predicate AP *mad* would normally be assumed to be sisters of the fronted auxiliary *is*. On that analysis the structure is flattened out so that *she mad* is not a constituent. In fact for English the fronting of *is* can even be seen as a consequence of that flattening: English is a head-initial language so the two dependents *she* and *mad* are expected to follow their selecting head *is*. (SAG REFS)

Although LFG is non-transformational, it can express the intuition behind the I-to-C movement analysis due to the separation of autonomous c- and f-structures. Recall from the above discussion of the DP in 3 that functional heads such as determiners, auxiliaries and complementizers do not introduce new f-structures, but rather map to the same f-structure as their complement phrases. The finite auxiliary can therefore appear in either I or C without affecting the f-structure, as we will see presently. Recall also that c-structure nodes are optional and can be omitted as long as a well-formed f-structure is generated. Comparing the non-terminal structures of 11 and 12, the I node is omitted from the latter structure but otherwise they are identical.

(11)



(12)



(Most of the lexical equations are omitted from 12 for clarity.) Given the ↑ =

↓ annotations, the C, I, and AP nodes (as well as IP and CP) all map to the same f-structure, namely the one shown in 13.

(13)
$$\left[\begin{array}{ll} \text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \textit{`PRO'} \\ \text{NUM} & \textit{SG} \\ \text{PERS} & \textit{3} \\ \text{GEN} & \textit{FEM} \end{array}\right] \\ \text{PRED} & \textit{`mad}\langle(\uparrow \text{SUBJ})\rangle\textit{'} \\ \text{FIN} & + \end{array}\right]$$

The C and I positions are appropriate for markers of clausal grammatical features such as finiteness ([FIN +]), such as auxiliary verbs like is and complementizers like finite *that* and infinitival *for*: *I said that/*for she is present* vs. *I asked for/*that her to be present*. English has a specialized class of auxiliary verbs for marking finiteness from the C position, while in languages like German all finite verbs, including main verbs, can appear in a C position that is unoccupied by a lexical complementizer. Summarizing, the LFG framework enables a theory of head mobility based on the intuition that a clause has multiple head positions where inflectional features of the clause are encoded.

5 Case, agreement, and constraining equations

The basic theory of agreement is the same in LFG and HPSG (see Wechsler 2020, Chapter 6 of this volume): agreement occurs when multiple feature sets arising from distinct elements of a sentence specify information about a single abstract object, so that the information must be mutually consistent (Kay 1984b). The two forms are said to agree when the values imposed by the two constraints are compatible, while ungrammaticality results when they are incompatible. An LFG example is seen in (3), where the noun, determiner and verbal suffix each specify person and/or number features of the same SUBJ value.

The basic mechanism for case marking works in essentially the same way as agreement, in both frameworks: in case marking, distinct elements of a sentence specify case information about a single abstract object, hence that information must be compatible. To account for the contrast in 14a, nominative CASE equations are associated with the pronoun *she* and added to the entry for the verbal agreement suffix *-s*:

(14) a. She/*Her/*You rules.

b. *she*: *Pron* $(\uparrow \textsc{case}) = \textsc{nom}$
$(\uparrow \textsc{pers}) = 3$
$(\uparrow \textsc{num}) = \textsc{sg}$
$(\uparrow \textsc{gend}) = \textsc{fem}$

c. *her*: *Pron* $(\uparrow \textsc{case}) = \textsc{acc}$
$(\uparrow \textsc{pers}) = 3$
$(\uparrow \textsc{num}) = \textsc{sg}$
$(\uparrow \textsc{gend}) = \textsc{fem}$

d. -*s*: *infl* $(\uparrow \textsc{tense}) = \textsc{pres}$
$(\uparrow \textsc{subj}) = \downarrow$
$(\downarrow \textsc{pers}) = 3$
$(\downarrow \text{NUM}) = \textsc{sg}$
$(\downarrow \textsc{case}) = \textsc{nom}$

The variant of 14a with *her* as subject is ruled out due to a clash of CASE features within the value of SUBJ in the f-structure. The variant with you as subject is ruled out due to a clash of PERS features. This mechanism is essentially the same as in HPSG, where it operates via the VALENCE feature.

This account allows for underspecification of both the case assigner and the case bearing element, and of both the controller and target of agreement. In English, for example, gender is marked on some pronouns but not on a verbal affix; and case is marked on the verbal affix but not on nominals, with the exception of the pronouns. But certain case and agreement phenomena do not tolerate underspecification, and for those phenomena LFG offers an account using a *constraining equation*, a mechanism absent from HPSG and indeed ruled out by current principles of HPSG theory. (Some early precursors to HPSG included a special feature value called ANY that functioned much like an LFG constraining equation (Shieber 1986: pp. 36-7), but that device has been eliminated from HPSG.) The functional equations described so far in this chapter function by building the f-structure, as illustrated in 3 and 5 above; such equations are called *defining equations*. A constraining equation has the same syntax as a defining equation, but it functions by checking the completed f-structure for the presence of a feature. An f-structure lacking the feature designated by the constraining equation is ill-formed.

The following lexical entry for *she* is identical to the one in 14b above, except that the CASE equation has been replaced with a constraining equation, notated with $=_c$.

(15) *she*: *Pron* $(\uparrow \text{CASE}) =_c \text{NOM}$
$(\uparrow \text{PERS}) = 3$
$(\uparrow \text{NUM}) = \text{SG}$
$(\uparrow \text{GEND}) = \text{FEM}$

The f-structure is built from the defining equations, after which the SUBJ field is checked for the presence of the [CASE NOM] feature, as indicated by the constraining equation. If this feature has been contributed by the finite verb, as in 14, then the sentence is predicted to be grammatical; if there is no finite verb (and there is no other source of nominative case) then it is ruled out. This predicts the following grammaticality pattern:

(16) Who won the popular vote in the 2016 election?
   a. She did! / *Her did!
   b. *She! / Her!

English nominative pronouns require the presence of a finite verb, here the finite auxiliary *did*. Constraining equations operate as output filters on f-structures and are the primary way to grammatically specify the obligatoriness of a form, especially under the assumption that all daughter nodes are optional in the phrase structure. As described in Section 3 above, obligatory dependents are specified in the lexical form of a predicator using existential constraints like ($\uparrow$ SUBJ) or ($\uparrow$ OBJ). These are equivalent to constraining equations in which the particular value is unspecified, but some value must appear in order for the f-structure to be well-formed.

A constraining equation for case introduced by the case-*assigner* rather than the case-bearing element, predicts that the appropriate case-bearing element must appear. A striking example from Serbo-Croatian is described by Wechsler & Zlatić (2003: p. 134), who give this descriptive generalization:

(17) Serbian/Croatian Dative/Instrumental Case Realization Condition.

If a verb or noun assigns dative or instrumental case to an NP, then that case must be morphologically realized by some element within the NP.

In Serbo-Croatian most common nouns, proper nouns, adjectives, and determiners are inflected for case. An NP in a dative position must contain at least one such item morphologically inflected for dative case, and similarly for instrumental case. The verb *pokloniti* 'give' governs a dative object, such as *ovom studentu* in 18a. But a quantified NP like *ovih pet studenata* 'these five students' is marked for invariant genitive case, and can appear in any case position— except when it fails to satisfy the condition in 17, such as this dative position (Wechsler & Zlatić 2003: p. 125):

(18) a. pokloniti knjige ovom studentu
give.INF books.ACC this.DAT.SG student.DAT.SG
'to give books to this student'

b. * pokloniti knjige [ovih pet studenata]
give.INF books.ACC this.GEN.PL five student.GEN.PL
('to give books to these five students')

Similarly, certain foreign names such as *Miki* and loanwords such as *braon* 'brown/ brunette' are undeclinable, and can appear in any case position, except those ruled out by 17. Thus the dative example in (19)a is unacceptable unless the inflected possessive adjective *mojoj/mojom* 'my' appears. When the possessive adjective realizes the case feature, it is acceptable. In (19)b we contrast the undeclined loan word *braon* 'brown' with the inflected form *lepoj* 'beautiful'. The example is acceptable only with the inflected adjective (Wechsler & Zlatić 2003: p. 134).

(19) a. Divim se *(mojoj) Miki.
admire.1SG REFL my.DAT.SG Miki
'I admire (my) Miki.'

b. Divim se *braon/ lepoj Miki.
admire.1sg REFL brown/ beautiful.DAT.SG Miki
('I admire brunette/ beautiful Miki')

This complex distribution is captured simply by positing that the dative (and instrumental) case assigning equation on verbs and nouns, such as the verbs *pokloniti* and *divim* in the above examples, is a constraining equation:

(20) $(\uparrow \text{OBL}_{dat}\ \text{CASE}) =_c \text{DAT}$

Any item in dative form within the NP, such as *ovom* or *studentu* in (18)a or *mojoj* or *lepoj* in (19), could introduce the [CASE DAT] feature that satisfies this equation, but if none appears then the sentence fails. In contrast, other case-assigning equations (e.g. for nominative, accusative, or genitive case, or for cases assigned by prepositions) are defining equations, which therefore allow the undeclined NPs to appear.

6 Agreement and affixal pronouns

Agreement inflections that include the person feature derive historically from incorporated pronominal affixes. Distinguishing between agreement markers and affixal pronouns can be a subtle and controversial matter. LFG provides a particular formal device for representing this distinction within the f-structure: a true pronoun, whether affixal or free, introduces a semantic form called ‘PRO’, which an agreement inflection does not. For example, Bresnan & Mchombo (1987) argue that the Chichewa (Bantu) object marker (OM) is an incorporated pronoun, while the subject marker (SM) alternates between agreement and incorporated pronoun, as in this example:

(21) Njûchi zi-ná-wá-lum-a a-lenje.
10.bee 10.SM-PST-2.OM-bite-FV 2-hunter
‘The bees bit them, the hunters.’

According to Bresnan & Mchombo (1987) the class 2 object marker *wá* is a pronoun, so the phrase *alenje* ‘the hunters’ is not the true object but rather a postposed topic cataphorically linked to the object marker, with which it anaphorically agrees in noun class. Meanwhile, the class 10 subject marker *zi-* alternates: when an associated subject NP (*njûchi* ‘bees’ in (2)) appears then it is a grammatical agreement marker, but when no subject NP appears then it functions as a pronoun. This is captured in LFG with the simplified lexical entries in (22):

(22) a. *lum*: *V* (↑ ‘)’ = ‘bite⟨(↑ SUBJ)(↑ OBJ)⟩’
b. *wá-*: *Aff* (↑ OBJ GEND) = 2
(↑ OBJ PRED) = ‘PRO’
c. *zi-*: *Aff* (↑ SUBJ GEND) = 10
((↑ SUBJ PRED) = ‘PRO’)

The PRED feature in (22)b is obligatory while that of (22)c is optional, as indicated by the parentheses. These entries interact with the grammar in the following manner. The two grammatical functions governed by the verb in (22)a are SUBJ and OBJ (the *governed* functions are the ones designated in the predicate argument structure of a predicator). According to the principle of completeness, a PRED feature must appear the f-structure for each governed grammatical function that appears within the angle brackets of a predicator. By the uniqueness condition it follows that there must be *exactly one* PRED feature, since a second such feature would cause a clash of values.[2]

[2]Each PRED value is assumed to be unique, so that two ‘PRO’ values cannot unify.

The OM *wá-* introduces the [PRED 'PRO'] into the object field of the f-structure of this sentence; the word *alenje* 'hunters' introduces its own PRED feature with value 'HUNTERS', so it cannot be the true object, and instead is assumed to be in a topic position. Bresnan & Mchombo (1987) note that the OM can be omitted from the sentence, in which case the phrasal object (here, *alenje*) is fixed in the immediately post-verbal position, while that phrase can alternatively be preposed when the OM appears. This is explained by assuming that the post-verbal position is an OBJ position, while the adjoined TOPIC position is more flexible.

The subject *njûchi* can be omitted from 21, yielding a grammatical sentence meaning 'They (some plural entity named with a class 10 noun) bit them, the hunters.' The optional PRED feature equation in (22)b captures this pro-drop property: when the equation appears then a phrase such as *njûchi* cannot appear in the subject position, since this would lead to a clash of PRED values ('PRO' versus 'BEES'); but when the equation is not selected then *njûchi* must appear in the subject position in order for the f-structure to be complete.

The diachronic process in which a pronominal affix is reanalyzed as agreement has been modeled in LFG as the historic loss of the PRED feature, along with the retention of the pronoun's person, number, and gender features (**coppock+wechsler:2010**). The anaphoric agreement of the older pronoun with its antecedent then becomes reanalyzed as grammatical agreement of the inflected verb with an external nominal. Finer transition states can also be modeled in terms of selective feature loss. Clitic doubling can be modeled as optional loss of the PRED feature where some semantic vestiges of the pronominal, such as specificity of reference (SUNER CITE). Selective loss of a phi feature leads to paradigm leveling. Coppock and Wechsler hypothesize that Hungarian 'definite object agreement' resulted from the radical loss of all phi features, leaving only the vestigial definiteness.

Comparing HPSG, a similar analysis of pronouns and agreement is possible but the particular way in which the semantics is grammaticalized makes LFG suitable for capturing this blah blah blah.

7 Lexical mapping

LFG and HPSG both adopt *lexical approaches to argument structure* in the sense of Müller & Wechsler (2014): a verb or other predicator is equipped with a valence structure indicating the grammatical expression of its semantic arguments as syntactic dependents. Both frameworks have complex systems for mapping semantic arguments to syntactic dependents that are designed to capture prevailing semantic regularities within a language and across languages. The respective systems differ greatly in their notation and formal properties but it is unclear whether there are any theoretically interesting differences, such as types of analysis that are available in one but not the other. This section identifies some of the most important analogues across the two systems, namely LFG's Lexical Mapping Theory (Bresnan et al. 2015: chapter 14) and the theory of macro-roles proposed by Davis and Koenig for HPSG (see Wechsler, Koenig & Davis 2020, Chapter 9 of this volume).

In Lexical Mapping Theory, the argument structure is a list of a verb's argument slots, each labeled with a thematic role type such as Agent, Instrument, Recipient, Patient, Location, and so on, in the tradition of Charles Fillmore's *Deep Cases* (**fillmore:1968**; **Fillmore:1977**) and Pāṇini's *kārakas* (**kiparsky+staal:1969**). The ordering is determined by a thematic hierarchy that reflects priority for subject selection.[3]. The thematic role type influences a further classification by the features [±*o*] and [±*r*] that conditions the mapping to syntactic functions (this version is from Bresnan et al. (2015: p. 331)):

(23) **Semantic classification of argument structure roles for function**:

patientlike roles:	$\underset{[-r]}{\theta}$
secondary patientlike roles:	$\underset{[+o]}{\theta}$
other semantic roles:	$\underset{[-o]}{\theta}$

The features [±*r*] (thematically *restricted*) and [±*o*] (*objective*) cross-classify grammatical functions: subject is [−*r*, −*o*], object is [−*r*, +*o*], obliques are [+*r*, −*o*] and restricted objects are [+*r*, +*o*]. A monotonic derivation (where feature values cannot be changed) starts from the argument list with the Intrinsic Classification (*I.C.* in example 24 below), then morpholexical operations such as passivization can suppress a role (not shown), then the thematically highest role (such as the Agent), if [−*o*], is selected as Subject (*Sbj.* in example 24) and then any remaining features receive positive values by default (*Def.* in example 24).

(24) Derivation of *eat* as in *Pam ate a yam.*

[3]The particular ordering proposed in **Bresnan+Kanerva:1989** and Bresnan et al. (2015) is the following: *agent* > *beneficiary* > *experiencer/goal* > *instrument* > *patient/theme* > *locative*

a-structure:	*eat*	<	*ag*	*th*	>
	I.C.		[−*o*]	[−*r*]	
	Sbj.		[−*r*]		
	Def.			[+*o*]	
			|	|	
f-structure:			SUBJ	OBJ	

In the macro-role theory formulated for HPSG, the analogues of [−*o*] and [−*r*] are the macro-roles 'Actor' (ACT) and 'Undergoer' (UND), respectively. The names of these features reflect large general groupings of semantic role types, but there is not a unique semantic entailment such as 'agency' or 'affectedness' associated with each of them. 'Actor' (ACT) and 'Undergoer' (UND), name whatever semantic roles map to the subject and object, respectively, of a transitive verb.[4] On the semantic side they are disjunctively defined: *x* is the Actor and *y* is the Undergoer iff '*x* causes a change in *y*, or *x* has a notion of *y*, or …' (quoted from Wechsler, Koenig & Davis 2020, Chapter 9 of this volume, example (15)). Such disjunctive definitions are the HPSG analogues of the Lexical Mapping Theory 'semantic classifications' shown in (23) above. In the HPSG macro-role system, linking constraints dictate that the ACT argument maps to the first element of ARG-ST, and that the UND argument maps to some nominal element of ARG-ST (25 and 26 are from Wechsler, Koenig & Davis 2020, Chapter 9 of this volume, examples (17) and (18)):

(25) $\begin{bmatrix} \text{CONTENT|KEY} & \begin{bmatrix}\text{ACT} & \boxed{1}\end{bmatrix} \\ \text{ARG-ST} & \left\langle \text{NP}_{\boxed{1}}, \ldots \right\rangle \end{bmatrix}$

(26) $\begin{bmatrix} \text{CONTENT|KEY} & \begin{bmatrix}\text{UND} & \boxed{2}\end{bmatrix} \\ \text{ARG-ST} & \left\langle \ldots, \text{NP}_{\boxed{2}}, \ldots \right\rangle \end{bmatrix}$

The first element of ARG-ST maps to the subject of an active voice verb, so (25)-(26) imply that the subject is the ACT if there is one, and otherwise it is the UND. Similarly, in Lexical Mapping Theory as described above, the subject is the [−*o*] highest argument, if there is one, and otherwise it is the [−*r*] argument. In this simple example we can see how the two systems accomplish exactly the same thing. A careful examination of more complex examples might point up theoretical differences, but it seems more likely that the two systems can express virtually the same set of mappings.

[4]Note, for example, that within this system the 'Undergoer' argument of the English verb *undergo*, as in *John underwent an operation*, is the object— and not the subject, as one might expect if being an 'Undergoer' involved actually undergoing something.

In LFG the *argument structure* (or *a-structure*) contains the predicator and its argument roles classified and ordered by thematic role type and further classified by Intrinsic Classification. It is considered a distinct level of representation, along with c-structure and f-structure. As a consequence the grammar can make reference to the initial item in a-structure, such as the agent (*ag*) in 24, which is considered the 'most prominent' role and often called the *a-subject* ('argument structure subject') in LFG parlance. To derive the passive voice mapping, the a-subject is suppressed in a morpholexical operation that crucially takes place before the subject is selected:

(27) Derivation of *eaten* as in *A yam was eaten (by Pam).*

$$
\begin{array}{llcccc}
\text{a-structure:} & \textit{eat} & < & \textit{ag} & \textit{th} & > \\
 & \text{I.C.} & & [-o] & [-r] & \\
 & \text{passive} & & \varnothing & & \\
 & \text{Sbj.} & & & [-o] & \\
 & & & & | & \\
\text{f-structure:} & & & & \text{SUBJ} &
\end{array}
$$

(The optional *by*-phrase is considered to be an adjunct referring to the passivized a-subject.) Note that the passive alternation is *not* captured by a procedural rule that replaces one grammatical relation (such as OBJ with another (such as SUBJ). The mapping from word strings to f-structures in LFG is monotonic, in the sense that information cannot be destroyed or changed. As a result the mapping between internal and external structures is said to be transparent in the sense that the grammatical relations of parts of the sentence are preserved in the whole (for discussion of this point, see Bresnan et al. (2015: chapter 5)). In early versions of LFG, monotonicity was assumed for the syntax proper, while destructive procedures were permitted in the lexicon. This was canonized in the *Principle of Direct Syntactic Encoding*, according to which all grammatical relation changes are lexical (Bresnan 1982). At that time an LFG passive rule operated on fully specified predicate argument structures, replacing OBJ with SUBJ, and SUBJ with an OBL_{by} or an existentially bound variable. The advent of LMT brought monotonicity to the lexicon as well.

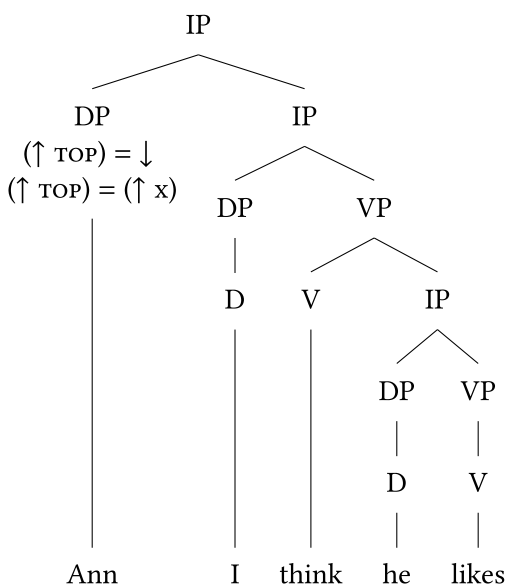
8 Long distance dependencies

In LFG a long distance dependency is modeled as a reentrancy in the f-structure, a structure lacking from HPSG. The HPSG theory of long distance dependencies is based on that of GPSG and uses the percolation of a 'slash' feature through the constituent structure. But the two frameworks are essentially very similar, both working by decomposing a long distance dependency into a series of local dependencies. As we will see, there are nevertheless some minor differences with respect to what hypothetical extraction patterns can be expressed.

Both frameworks allow either a 'gap' or 'gapless' account: regarding LFG see Bresnan et al. (2015) for gap and Dalrymple (2001) for gapless; regarding HPSG see Pollard & Sag (1994) for gap and Sag et al. (2003) for gapless. Gaps have been motivated by the (controversial) claim that the linear position of an 'empty category' matters for the purpose of weak crossover and other binding phenomena (Bresnan et al. 2015: chapter 9). In this section I compare gapless accounts.

LFG has two grammaticalized discourse functions, top ('topic') and foc ('focus'). A sentence with a left-adjoined topic position is depicted in (28). The topic phrase *Ann* serves as the object of the verb *like* within the clausal complement of *think*. This dependency is encoded in the second equation annotating the topic node, where the variable *x* ranges over strings of attributes representing grammatical functions such as subj, obj, obl, or compl. These strings describe paths through the f-structure. In this example *x* is resolved to the string compl obj, so this equation has the effect of adding to the f-structure in (30) the curved line representing token identity.

(28)



(29)

(30)
$$\begin{bmatrix} \textsc{top} & [\text{“Ann”}] \\ \textsc{subj} & [\text{“I”}] \\ \textsc{pred} & \text{‘}\textit{think}\ \langle (f\ \textsc{subj})(f\ \textsc{compl}) \rangle\text{’} \\ \textsc{compl} & \begin{bmatrix} \textsc{subj} & [\text{“he”}] \\ \textsc{pred} & \text{‘}\textit{like}\ \langle (g\ \textsc{subj})(g\ \textsc{obj}) \rangle\text{’} \\ \textsc{obj} & \end{bmatrix} \end{bmatrix}$$

HPSG accounts are broadly similar. One HPSG version relaxes the requirement that the arguments specified in the lexical entry of a verb or other predicator must in its valence lists. Verbal arguments are represented by elements of the ARG-ST list, so the list for the verb *like* contains two DPs, one each for the subject and object. In a sentence with no extraction, those ARG-ST list items map to the valence lists, the first item appearing in SUBJ and any remaining ones in COMPS. To allow for extraction, one of those ARG-ST list items is permitted to appear on the SLASH list instead. The SLASH list item is then passed up the tree by means of strictly local constraints, until it is bound by the topicalized phrase.

The LFG dependency is expressed in the f-structure, not the c-structure. Bresnan et al. (2015: chapter 2) note that this allows for category mismatches between the phrases serving as filler and those in the canonical, unextracted position. This is illustrated in example (7) above. The lexical entry for talk (about) in (8) selects an OBL_{about} OBJ but does not specify the part of speech category of that argument; meanwhile, the phrase structure rules dictate that the top position allows (at least) DP and CP, while the position right adjacent to about can only house a DP. Category mismatches pose a problem for transformational theories that assume the Projection Principle, since the ‘moved’ constituent should satisfy the conditions imposed on the phrase in its source position. But HPSG seems to permit essentially the same analysis as the LFG analysis just sketched. In HPSG, the preposition *about* would have a disjunctive DP/CP on its ARG-ST list item, but only DP is selected for its COMPS list item; and the topic position would allow either DP or CP.

Constraints on extraction such as accessibility conditions and island constraints can be captured in LFG by placing constraints on the attribute string x (Dalrymple 2001). If subjects are not accessible to extraction then we stipulate that SUBJ cannot be the final attribute in x; if subjects are islands then we stipulate that SUBJ cannot be a non-final attribute in x. If that attribute path is the only place such constraints can be stated then this would make the interesting (but false; see presently) prediction that the theory of extraction cannot distinguish between constituents that map to the same f-structure. For example, as noted in Section

4 function words like determiners and their contentful phrases like NP are usually assumed to be f-structure co-heads, so the DP *the lion* maps to the same f-structure as its NP daughter *lion* (see diagram (3). This predicts that if the DP can be extracted then so can the NP, but of course that is not true:

(31) a. The lion, I think she saw.
   b. * Lion, I think she saw the.

These two extractions would involve the same attribute path, namely COMPL OBJ. In fact LFG theory does not assume that path constraints exhaust the possibilities for expressing extraction conditions. The 'manner of speaking' verbs in ??a

(32) Who did Chris think/*whisper that David saw?

Dalrymple (2001) notes that "There is no reason to assume that the grammatical function of the sentential complements of these two verbs differs" and instead proposes that verbs place a boolean feature [LDD±] on their clausal complements; nonbridge verbs like whisper assigning [LDD − −] and other verbs assigning [LDD+]. Then the extraction path is then subject an *off-path constraint* stating that any COMPL in the path cannot contain the feature [LDD+].

9 Control and raising

Raising and control (equi) words are treated in virtually the same way in LFG and HPSG: a subject control word (such as *hope* in (33)) specifies that its subject is (also) the subject of its predicate complement.

(33) Pam hopes to visit Fred.
   a. *hope*: $(\uparrow \text{PRED}) = \text{`hope}\langle(\uparrow \text{SUBJ})(\uparrow \text{XCOMP})\rangle\text{'}$
   $(\uparrow \text{SUBJ}) = (\uparrow \text{XCOMP SUBJ})$
   b. *hope*: $[\ \text{ARG-ST}\ \langle \boxed{1}\text{NP}, \text{VP}[\textit{inf}; \text{SUBJ}\ \langle \boxed{1} \rangle]\rangle]$

The LFG entry for hope in (33a) contains the grammatical function XCOMP ('open complement'), the function reserved for predicate complements such as the infinitival phrase *to win the race*. The control equation specifies that its subject is equivalent to the subject of the verb *hope*; and the tag $\boxed{1}$ in (33b) plays the same role in the simplified HPSG entry in (33b). One interesting difference is that the HPSG representation allows only for control (or raising) of subjects and not complements. More precisely, it allows for control of the outermost or final dependent to be combined with the verbal projection. This is because of the list cancellation regime that operates with valence lists. The expression ‘VP’ in (33b) represents an item with an empty COMPS list. In a simple English clause the verb combines with its complement phrases to form a VP constituent, which which the subject is then combined to form a clause. Assuming the same order of combination in the control structure, it is not possible to control the complement of a structure that contains a structural subject, as in (34a):

(34) a. * Fred hopes Pam to visit.
b. Fred hopes to be visited by Pam.
c. ?! (↑ SUBJ) = (↑ XCOMP OBJ)

The intended meaning would be that of (34b). The passive voice is needed in order to make the intended controllee (*Fred*) the subject of *visit* and thus available to be controlled. This restriction to subjects follows from the HPSG theory, while in LFG it follows only if control equations are defined so as to exclude equations like the one in (34c).

10 Semantics

11 Conclusion

Blah blah blah

Abbreviations

Acknowledgements

References

Bresnan, Joan (ed.). 1982. *The mental representation of grammatical relations* (MIT Press Series on Cognitive Theory and Mental Representation). Cambridge, MA/ London: MIT Press.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Mark Wechsler. 2015. *Lexical-functional syntax*. 2nd edn. Oxford: Blackwell Publishers Ltd. DOI:10.1002/978111910566

Bresnan, Joan & Sam Mchombo. 1987. Topic, pronoun, and agreement in Chicheŵa. *Language* 63. 741–82. DOI:10.2307/415717

Dalrymple, Mary. 2001. *Lexical Functional Grammar* (Syntax and Semantics 34). New York, NY: Academic Press.

Falk, Yehuda N. 1984. The English auxiliary system: A Lexical-Functional analysis. *Language* 60(3). 483–509.

Kay, Martin. 1984a. Functional Unification Grammar: A formalism for machine translation. In Yorick Wilks (ed.), *Proceedings of the 10th International Conference on Computational Linguistics and 22nd Annual Meeting of the Association for Computational Linguistics*, 75–78. Stanford University, California: Association for Computational Linguistics.

Kay, Martin. 1984b. *Functional unification grammar: a formalism for machine translation.* Association for Computational Linguistics. 75–78.

Müller, Stefan & Stephen Mark Wechsler. 2014. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76. DOI:10.1515/tl-2014-0001

Nordlinger, Rachel. 1998. *Constructive case: Evidence from Australia* (Dissertations in Linguistics). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction.* 2nd edn. (CSLI Lecture Notes 152). Stanford, CA: CSLI Publications.

Shieber, Stuart M. 1986. *An introduction to unification-based approaches to grammar* (CSLI Lecture Notes 4). Stanford, CA: CSLI Publications.

Wechsler, Stephen. 2020. Agreement. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 161–184. Berlin: Language Science Press. DOI:??

Wechsler, Stephen Mark & Larisa Zlatić. 2003. *The many faces of agreement* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

Wechsler, Stephen, Jean-Pierre Koenig & Anthony Davis. 2020. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 255–303. Berlin: Language Science Press. DOI:??