

Heads and history
NIGEL VINCENT & KERSTI BÖRJARS
The University of Manchester

Heads come in two kinds: lexical and functional. While the former are treated in a largely uniform way across theoretical frameworks, with the latter, things are different. Functional heads have been reified as a core theoretical construct within Minimalism, where they abound, but have much less presence in LFG and an even more reduced role in HPSG. The difference between the two kinds of heads also plays out in the diachronic domain. Nouns, verbs and adjectives often have consistent historical trajectories over centuries. Many of the nouns of modern English, for example, were also nouns a millenium ago in Old English even if they have undergone extensive phonological and semantic change in the meantime. The diachronic profiles of items that realise functional heads are very different, since typically they start out as full lexical words before undergoing the various changes that fall under the pre-theoretical label of grammaticalization. English *will* is a good case in point, having begun life as a lexical verb meaning 'want' before becoming the temporal/modal marker that it is today (and in some systems being assigned to the class of functional heads). The key question then becomes: how do diachrony and synchrony interact, and in particular how is the historical relation between lexical and functional categories treated, in different grammatical frameworks? In the present paper, we seek to compare and contrast LFG and HPSG as models of (morpho)syntactic change and in turn to compare them with the Minimalist approach to the same dataset.

There has to date been relatively little work from a diachronic perspective within LFG (but see the contributions to Butt & King 2001 for some examples and Börjars & Vincent, in press, for a general overview) and virtually nothing within HPSG. And yet in different ways both approaches have much to offer those who are interested in bringing formal methods into historical linguistics. In particular, their less rigid approach to phrase structure when compared to Minimalism makes them of special interest in this context. We explore here three concepts, two drawn from the HPSG literature and one from LFG, and examine how they play out in the diachronic domain. The first is the distinction between head and marker, where the latter is defined by Sag & Pollard (1994: 45) as "a word that is 'functional' or 'grammatical' as opposed to substantive, in the sense that its semantic content is purely logical in nature (perhaps even vacuous)". Crucially, a marker is not a head. Their interest here is particularly focussed on the status of complementizers, a category which is well know to have a varied range of diachronic sources: (pro)nominal (e.g. English *that*, Korean *kɛs* < 'thing' with finite clauses), prepositional (e.g. Eng *for*, French *à*, *de* with infinitival clauses), and verbal (e.g. Yoruba *kpé*, Uzbek *deb* both deriving from verbs meaning 'say'). Such examples suggest a diachronic trajectory as in (1):

(1) HEAD > MARKER

More recently, in some versions of HPSG the concept 'marker' has been replaced by that of a 'weak head', defined by Abeillé *et al* (2006: 156) as 'a lexical head that shares its syntactic category and other HEAD information with its complement'. Their example concerns the French prepositions *à* 'to' and *de* 'of', for which they distinguish two broad classes of uses, one in which they are characterised as full lexical heads as in (2a) and the other in which they are weak heads as in (2b) (in what follows to keep things simple we use only examples with *à*):

- (2) a. Il est allé à la gare
he be.PRES.3SG go.PSTPRT to the station
'he went to the station'
- b. Il m'a invité à venir demain
he me-have.PRES.3SG invite.PSTPRT to come.INF tomorrow
'he invited me to come tomorrow'

They represent the lexical preposition in much the same way as it would be represented in other frameworks: it is of the category *prep-word* and takes an N-headed (or in other approaches D-headed) complement. The difference between frameworks is rather to be seen in the treatment of the grammaticalized use of the preposition to introduce an infinitive. For Abeillé *et al*, the weak head *à* in (2b) is a head in the sense that it selects a complement, viz the infinitival VP *venir demain*, and it adds a value for the feature MARKING to the phrases it heads, but remains weak in the sense that it inherits the valence list of its complement. What is of interest in the present context is that this latter

use of *à* is the product of a process of change. Its Latin etymon *ad* has a range of lexical meanings and the new grammatical uses develop over several centuries (Adams 2013: Ch XIII). This in turn leads us to replace the diachronic trajectory in (1) with that in (3)

(3) HEAD > WEAK HEAD

On the face of it, such a change looks very similar to the way developments of this kind have been modelled within Minimalism. An item like *à* in (2b) is considered to be a complementizer (Kayne 1999) and located in C or one of the sub-heads of C such as Fin (Rizzi 1997), so that diachronically the shift is from lexical to functional head (Roberts & Roussou 2003). However, Abeillé *et al* (2006, note 12) are at pains to stress that, in their words, “weak heads differ from functional heads in LFG or GB”. In other words a weak head is not a new type of category. As they go on to say: “Although a weak head’s category is underspecified in the lexicon, in any given syntactic context, it has a completely ordinary syntactic category (e.g. N or V). It is important to emphasize that when a weak head inherits a value of type *verb* or *noun*, it does not actually ‘become’ a verb or a noun (i.e., a lexical object of type *noun-word* or *verb-word*).” Rather, it maintains its status as a *prep-word*, which it shares with the full lexical preposition seen in (2a). In other words the change is not a matter of grammatical category but of the manner in which elements of this kind integrate with the other parts of the sentence.

How then can such a change be modelled within LFG, a framework in which the distinction between category and function is built into the basic architecture via the distinction between f-structure and c-structure? The lexical preposition heads a PP with the associated f-structure in (4) (Dalrymple 2001: 151-3):

(4)

$$\left[\text{OBL}_{goal} \left[\begin{array}{ll} \text{PRED} & \text{'gare'} \\ \text{PCASE} & \text{OBL}_{goal} \end{array} \right] \right]$$

On this view, core prepositions like *à* are already in some sense partly grammaticalized even when they take a nominal argument since they do not have their own PRED feature, but rather act as functional co-heads. The change will thus be as in (5);

(5) Lexical head (with own PRED value) > functional head acting as co-head (with loss of PRED)

For the infinitival construction, the choice is between maintaining the prepositional analysis, which entails a c-structure of the form $[_{PP} [_P \hat{a}] [_{VP} \textit{venir demain}]]$ and would imply that diachronically the shift is not in the prepositional head but rather in an expansion of its f-structure to include XCOMP as well as OBL. Alternatively, we have a CP with *à* defined as the value for the COMPFORM feature within its associated f-structure. The latter solution amounts to saying that there has been a diachronic shift at the categorial level, viz:

(6) P > C

The empirical evidence is split. Latin prepositions did not govern infinitives, but there was a construction in which *ad* took a gerund as complement, thus *ad dicendum* ‘towards, for speaking’. The change seems to have involved the loss of the gerund (in this function at least) and its replacement by the infinitive, itself also a verbal noun in origin. While this argues for *ad* and its Romance reflexes having retained the status of prepositions, the fact that there are in the modern languages alternations between prepositional infinitives and finite complements introduced by *que* ‘that’ argues for the shift from P to C.

Whichever solution is in the end adopted, there is a further difference between the use of functional heads in LFG and Minimalism that needs to be emphasised. In the remark quoted above Abeillé *et al* refer to ‘LFG and GB’. While it is true that in the latter, functional heads were for the most part restricted to C, T, I and D, at least one strand of Minimalism, the so-called cartographic approach developed by Cinque and others, takes the further step of decomposing heads like C into a set of subsidiary functional heads such as the Force, Fin, etc (Rizzi 1997). This reflects the desire within that approach to encode all properties in a rigid, binary-branching, universally headed conception of phrase structure, something which in turn follows from the rejection of the conceptual difference between f-structure and c-structure, which is at the core of LFG as a model of natural language.

Another development within LFG is Toivonen's (2003) notion of a non-projecting word. Items in this class are of category X^0 , but do not project to X' or XP, are marked as such in the lexicon and are head-adjoined to an associated and projecting X^0 . Her case study focuses on Swedish particles such as *ihjäl* 'to death' in the string *slagit ihjäl* 'kill, lit. beat to death', where *slagit* is of the category V^0 as is the whole string but where *ihjäl* is a non-projecting P. As she demonstrates, the items that fall within the class of particles belong to a number of different categories — verbal, nominal, adjectival and prepositional — but what they have in common is that they adjoin to another item, to which in effect they cede head status, in this respect showing some parallels with the concept of weak head. What Toivonen does not observe, but which is striking once the diachronic perspective is adopted, is that most if not all the items she categorises as non-projecting in this sense are themselves historically derived from full projecting categories or even phrases. The form *ihjäl* for example is a frozen PP < *i hel* 'in the land of the dead'

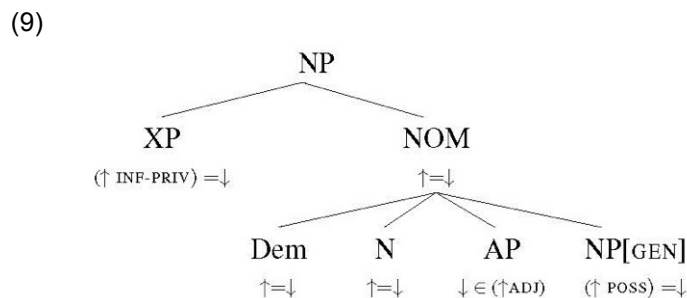
The notion of non-projecting categories is used to account for the historical development of functional category by Börjars, Harries & Vincent (BHV) (2016) (see also Lander & Haegeman 2013). In Old Norse, definiteness marking takes the form of either a nominal affix (7a) or a separate syntactic element (7b):

- (7) a. hestr-inn
horse-DEF
'the horse'
- b. (h)inn stóri hestr
DEF big.WK horse
'the big horse'

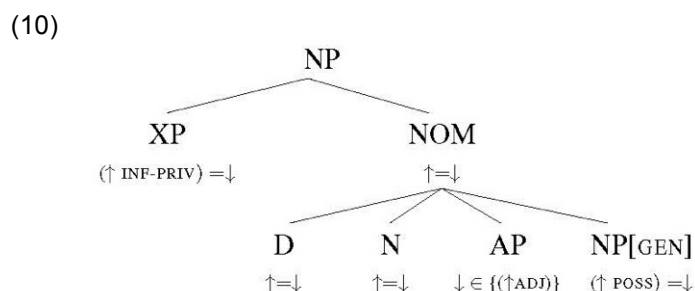
Given the the lack of complementary distribution between definiteness markers, demonstratives, and possessive pronouns and the fact that nouns may be interpreted as definite or indefinite dependent on context and without overt marking, BHV suggest that there is no unified category D at this stage. There is however an initial position, which is privileged in terms of information structure. This is most clearly seen with possessives, which normally follow the noun (*faðir hans* 'his father', *húsfreyja þín* 'your wife') but when overtly contrasted may precede:

- (8) at minn faðir væri eptirbát þíns fōður
COMP 1SG.POSS father was after.boat 2SG.POSS.GEN father.GEN
'that my father trailed in the wake of yours' (Gunnl 9.33)

From such evidence they argue for the phrase structure in (9), where demonstratives and possessives may convey definite meaning but there is no clearly identifiable definiteness marking element:



Moving on from Old Norse to one of its descendants, early Faroese, BHV show that there is now a situation in which a definiteness marker is an obligatory part of the NP, but it has no fixed position and therefore, they suggest, does not project. It is a D^0 head and a non-projecting category in Toivonen's sense. This is represented in (10), with D taking its place as one of the elements dominated by NOM:



In modern Scandinavian languages, by contrast, the fixed initial position has come to be occupied by an (in)definiteness item and they therefore suggest that there is now a full DP. On this account, a non-projecting category provides an intermediate stage between no overt structure — i.e. an NP-language in the sense of Bošković (2008, 2009) — and the full projection emanating from a new functional head. The case of the North Germanic DP demonstrates a different kind of change from our first example, one in which what is at issue is not the changing status of an item that has been historically present throughout the period under consideration but the emergence of a new category, and ultimately its phrasal projection, where previously there had been nothing or only word-internal bound morphology.

We are now in a position to draw a number of conclusions from these case studies:

- a) Neither LFG nor HPSG have any internal principle within their architecture that predicts the direction of change. This is a notable difference when compared to Minimalism, where the fact that grammaticalization changes show a directionality — over time verbs become auxiliaries but auxiliaries do not revert to being full lexical verbs — follows from the fact that Universal Grammar allows raising but not lowering as a derivational operation. We suggest however that it is better to consider the driving force of change to be the external circumstances of language use, but to deploy the devices of formal syntax in order to model such changes as and when they are attested.
- b) Both LFG and HPSG have the resources to effect such modelling, but further comparison over a more diverse set of examples is required to understand better the similarities and differences between constructs like ‘weak heads’, ‘markers’ and ‘non-projecting words’. We believe, however, that the present paper has demonstrated some fruitful avenues for future research within both frameworks.
- c) We have also sought to show how the diachronic domain is a fruitful one within which to compare the descriptive and analytical capacity of the two frameworks.
- d) One respect in which the frameworks differ and thus make different predictions about the nature of change lies in the way the relation between form and function is conceived. Thus, within LFG, since f-structure and c-structure — and indeed potentially other dimensions such as m-structure and i-structure — do not have to be, so to speak, ‘in synch’, it follows that change in one dimension may proceed independently of the other. For example, new f-structures may become associated with existing c-structures, as on one interpretation of the development of prepositional complementizers from Latin to French. Conversely, the history of definiteness in North Germanic suggests that new c-structures may become associated with what is in essence the same f-structure. By contrast, within HPSG, form and function are more closely aligned, especially within the variant which is SBCG. This leads to a view of change more in line with that advocated in Traugott & Trousdale (2013), in which form and function are united within distinct and diachronically evolved constructions.

References

- Abeillé, Anne, Olivier Bonami, Danièle Godard & Jesse Tseng (2006) The syntax of French *à* and *de*: an HPSG analysis. In Patrick Saint-Dizier (ed.) *Dimensions of the Syntax and Semantics of Prepositions*. Dordrecht: Springer, 147-162.
- Adams, J.N. (2013) *Social Variation and the Latin Language*. Cambridge: Cambridge University Press.
- Börjars, Kersti, Pauline Harries & Nigel Vincent (2016) Growing syntax: the development of a DP in North Germanic. *Language* 92: e1-e37.
- Börjars, Kersti & Nigel Vincent (in press) Lexical-Functional Grammar. In Adam Ledgeway & Ian Roberts (eds) *The Cambridge Handbook of Historical Syntax*, Cambridge: Cambridge University Press.
- Bošković, Željko (2008) What will you have, DP or NP?, *Proceedings of NELS 37*: 101-114.
- Bošković, Željko (2009). More on the no-DP analysis of article-less languages. *Studia Linguistica* 63: 187-203.
- Butt, Miriam & Tracy Holloway King (eds) (2001) *Time over Matter. Diachronic Perspectives on Morphosyntax*. Stanford: CSLI.
- Dalrymple, Mary (2001) *Lexical Functional Grammar*. San Diego: Academic Press.
- van Gelderen, Elly (2007) The definiteness cycle in Germanic. *Journal of Germanic Linguistics* 19: 275–308.
- van Gelderen, Elly (2011) *The Linguistic Cycle*. Oxford: Oxford University Press.
- Lander, Eric, and Liliane Haegeman (2013) Old Norse as an NP language, with observations on the common Norse and Northwest Germanic Runic inscriptions. *Transactions of the Philological*

- Society* 112: 279–318.
- Kayne, Richard (1999) Prepositional complementizers as attractors. *Probus* 11:39–73.
- Rizzi, Luigi (1997) The fine structure of the left periphery, in Liliane Haegeman (ed.) *Elements of Grammar*, Dordrecht: Kluwer, 281-337.
- Roberts, Ian & Anna Roussou (2003) *Syntactic Change. A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge University Press.
- Sag, Ivan & Carl Pollard (1994) *Head-Driven Phrase Structure Grammar*. Chicago: Chicago University Press.
- Toivonen, Ida (2003) *Non-projecting words: A case study of Swedish verbal particles*. Dordrecht: Kluwer.
- Traugott, Elizabeth Closs & Graeme Trousdale (2013) *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.