

Dimensions of variation in the expression of functional features: modelling definiteness in LFG

Kersti Börjars & John Payne
The University of Manchester

It has long been recognised that the traditional distinction word–clitic–affix is not sufficient to capture the full range of variation in the exponence of functional features cross-linguistically. Zwicky (1987) and others showed that the properties used to distinguish these categories do not always pattern as neatly as assumed by the three-way distinction, and he introduced the “hybrid category” PHRASAL AFFIX. By analysing the variation in expression of case, Plank (1991) showed further complexities in the patterning of properties. In more recent work Spencer and Luís (2012a) divide the characteristics used to distinguish between words and bound elements into two groups: formal and placement characteristics. They argue that clitics lie at the intersection of these two; canonical clitics have the formal properties of affixes, but the positional properties of (function) words. They recognise that many elements will not show the canonical behaviour, but an element can slide along the scale of form properties towards the affix end without simultaneously acquiring word positional properties and hence become a non-canonical clitic.

In this paper, we will consider the expression of definiteness cross-linguistically and argue that the range of variation in expression is even broader than has been recognised in this literature. The markers range from syntactically independent projecting words to purely prosodic processes and display variation in a number of discrete dimensions. We show that the LFG architecture enables these different dimensions to be modelled independently, and that the unnecessary problems created by over-simplistic categorisation (e.g. into “clitic” vs “affix”) evaporate as a consequence.

By functional definiteness marker (FDM), we understand definiteness markers which by their presence are sufficient to induce a definite interpretation of a noun phrase, i.e. in LFG terms they will map to a [DEF +] feature in f-structure. We consider only elements that do not carry functional features beyond definiteness (though they will frequently also carry agreement features). This entails that we only include definiteness markers which are distinguishable from demonstratives. We also distinguish FDMs from elements such as the [DEF ±] marker on Swedish adjectives, which must be modelled as a purely morphological feature in the sense that it does not contribute to f-structure.

We distinguish (at least) the following dimensions of variation:

(i) *Prosodic vs segmental* Does the FDM have solely prosodic instantiation, e.g. the Iron dialect of Ossete (Abaev, 1949) where stress shifts leftwards, or solely segmental instantiation (as in the majority of cases), or a mixture of these (languages like Tongan (Poser, 1985)). If definiteness is associated with stress placement, there has to be a mapping between a functional feature [DEF +] and p(rosodic)-structure.

(ii) *Syntactic vs morphological* Is the FDM introduced as a separate word by the c-structure rules, or is it introduced as a morphological component of another word? At the extreme syntactic end are elements such as the Danish definiteness marker, which can surface in isolation from a head noun in non-elliptic constructions. At the other end, we have FDMs which are clearly morphologically integrated with their hosts, as evidenced for instance by arbitrary irregularities. Syntactic FDMs will be represented by an independent category D. FDMs which are a morphological component of another category (typically N or A) will be modelled by an inside-out designator which associates the feature [DEF +] with the f-structure of the noun phrase.

(iii) *Prosodic independence vs dependence* English *the* is syntactic, like the Danish marker, and hence found under a D in c-structure, but unlike the Danish element, it is prosodically weak. Hence it needs a host and cannot occur independently. It should be noted that this prosodic dependence is distinct from degree of morpho-phonological integration (see (viii) below).

(iv) *Single or multiple instantiation* Does one sole element introduce the FDM, or can a single noun-phrase contain multiple instances of such elements, as with definiteness agreement in Arabic? We assume that more generally, principles of economy cause avoidance of multiple exponence of any functional feature, in this case the FDM, but that languages may be uneconomical in this sense and require more than one marker. Frequently, a historical explanation for the multiple exponence can be found.

(v) *Paradigmatic contrast with other determiners* Does the element which introduces the FDM occur

in complementary distribution with other determiner-like elements, e.g. demonstratives or possessor phrases, or is it happy to co-occur with these? The same principle of economy assumed for single or multiple exponence would be at work here.

(vi) *Standard versus special placement* Do the words which introduce FDMS occupy the positions which such words normally hold in the absence of definiteness marking, or are the words which introduce FDMS subject to special positioning by virtue of their definiteness? For example, in Bulgarian, the phrase which introduces the FDM must be left-most. Rules which associate a defined position for the element, whatever its category, which introduces the functional feature [DEF] will handle these special placements. Such a rule may refer to c-structure, p-structure or i(nformation)-structure units in determining placement.

(vii) *Wide scope in coordinations* Does the FDM need to be introduced separately on individual coordinates, or can it have wide scope over a coordination? This characteristic has been assumed in the literature to be one that distinguished between morphological and syntactic status (see for instance Zwicky and Pullum (1983)). However, we show that there is no such correlation.

In this paper, we will show how the analysis of any of these elements requires an architecture in which there is no assumption of a one-to-one mapping between dimensions. The Bulgarian definiteness marker *-ta* serves as an illustrative example. The element is clearly a bound element, triggering arbitrary stem alternations (see Halpern (1995) and Spencer and Luís (2012b)). On the other hand, it requires special positioning, the correct generalisation being that the word or phrase which contains it must be leftmost in the noun phrase, and that if the leftmost element is a phrase, it occurs on the head of that phrase (whether or not this itself is the leftmost word in the noun phrase). For Anderson (2005), this special positioning would require a phrasal affix analysis, but this is contradicted by the morphophonological integration of the marker. In our analysis, Bulgarian will contain a syntactic rule which requires the element which introduces the feature [DEF +] to occur left-most. This might be the head noun, in which case the noun itself must contain the *-ta* marker. Or it might be an adjective phrase, in which case the adjective which heads this phrase will contain the *-ta* marker. In either case, the word containing the *-ta* marker will project a [DEF +] feature into f-structure.

The general conclusion is then that the properties discussed here, and familiar from the literature on the distinction between ‘affix’ and ‘clitic’, do not define two or three or even four categories. At the very best, there may be some relatively common clusterings which could be abstracted into categories, but we are not even confident that a careful statistical analysis of the distribution of the properties would warrant this.

References

- Abaev, V I. 1949. *Osetinskij jazyk i fol'klor, vol. 1*. Moskva-Leningrad: Akademija Nauk SSSR.
- Anderson, Stephen R. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.
- Halpern, Aaron. 1995. *On the placement and morphology of clitics*. Stanford: CSLI Publication.
- Plank, Frans. 1991. From cases to adpositions. *EUROTYP Working Papers. Theme 7: Noun phrase structure* 13.
- Poser, William J. 1985. Cliticization to np in lexical phonology. In *Proceedings of WCCFL IV*, ed. Jeffrey Goldberg, Susannah MacKay, and Michael T Wescoat, 262–272. Stanford Linguistics Association.
- Spencer, Andrew, and Ana Luís. 2012a. Canonical clitics. In *Canonical morphology and syntax*, ed. Dunstan Brown, Marina Chumakina, and Grev G Corbett, 123–150. Oxford University Press.
- Spencer, Andrew, and Ana Luís. 2012b. *Clitics: an introduction*. Cambridge: Cambridge University Press.
- Zwicky, Arnold M. 1987. Suppressing the Zs. *Journal of Linguistics* 23:133–148.
- Zwicky, Arnold M, and Geoffrey K Pullum. 1983. Cliticization vs. inflection: English n't.