More distributed than Distributed Morphology: Rethinking the shape of mental lexicon

As a fundamental component of human language, the mental lexicon is an indispensable part of any linguistic theory, be it in the shape of an all-in-one warehouse of phonological, semantic, and formal information (the traditional view) or as a series of distributed lists (Distributed Morphology, Halle & Marantz 1993 et seq). In this paper, I will argue for an even more distributed Lexicon and defend this argument from conceptual, experimental, and empirical perspectives.

**Conceptual motivation:** Syntactic research today has developed into a sub-atomic (feature) level. This is true for most generative approaches. Lexicalist theories attribute syntactic variation to lexically encoded features and their value setting (e.g. Borer-Chomsky Conjecture, Baker 2008). Neo-Constructionist theories like Exo-Skeletal Syntax (Borer 2005) and Nanosyntax (Starke 2009) explain morpho-syntactic phenomena with functional categories, though these are not encoded in the (Lexicalist) Lexicon. Similarly, Cartographic theories (Rizzi 1997, Cinque 1999, et seq) depict an elaborate left periphery, also in the form of functional categories. While these approaches are successful in numerous aspects, there remains a basic question: where are the functional categories (qua features) stored?

This is not an issue for theories assuming an all-in-one Lexicon, but such a Lexicon is not an ideal model for a feature-based program (hence DM). Nor is it entirely desirable to remove grammatical information from the Lexicon and encode it in syntax, as the featural primitives of such information do need long-term storing, but syntax, as a minimally computational system, ideally does not have storing capacity. In this respect, DM’s Narrow Lexicon (List 1) could be a nice solution, except that it also contains roots, which definitely need huge storing capacity. Thus, DM does not really reduce the size burden of the “syntax-appendix” part of the mental lexicon, for however many Vocabulary Items there are in List 2, there must be as many correspondent entries in List 1.

**Experimental evidence:** Since mental lexicon is part of memory, facts of the brain should provide direct evidence for its shape. Indeed, neurolinguistic studies generally assume a distributed lexicon or even multiple modality-specific lexicons (Ingram 2007). Among others, Barner & Bale’s (2002) research on category-specific aphasia supports the DM hypothesis of underspecified root.

Two recent studies seem to confirm the conceptual considerations discussed above. First, Berwick & Chomsky (2016) suggest that lexical items (or at least their Merge-related features) are stored in the middle temporal cortex, which is separated from the syntax region (frontal cortex) by the superior temporal cortex. This means that syntax indeed does not store lexical information in itself. Second, Franco (2014) notices that logopenic Primary Progressive Aphasia (PPA) patients tend to use a lot of light verbs, whose retrieval is left intact by erosion of the (root-storing) lexicon. Given that light verbs are (semi-)grammatical items, this suggests that functional categories and roots are probably not stored in the same place.

**Proposal:** With the above conceptual motivation and experimental evidence, I tentatively split DM List 1 into two sub-lists, i.e. List 1.1 for morpho-syntactic features, and List 1.2 for roots. Roots are linguistically instantiated concepts (or their “indices”, as in Harley 2014a). List 1.1 can be viewed as a radically reduced version of the traditionally conceived “UG-appendix” Lexicon, though it is not really “appended”, as syntax and lexicon have separate brain locations. Lists 1.2, 2, and 3 hold the various non-grammatical information of the mental lexicon. There is probably some fixed mapping between List 2 (Vocabulary) and List 3 (Encyclopedia) mediated by roots, but the insertion of roots (List 1.1) and that of morpho-syntactic features (List 1.1) are mutually
independent. As such, selection between functional categories and roots is not predetermined in List 1, but possibly regulated by the “stretching” ability of the conceptual component (Borer 2005).

**Further implication:** The above more distributed Lexicon provides us with a new approach to grammaticalization and lexicalization. Grammaticalization amounts to linking a Vocabulary Item in List 2 to some functional category in List 1.1 plus delinking it from a root in List 1.2 (call this process “relinking”). Lexicalization (in its most basic form) is simply root creation in List 1.2, usually accompanied by the creation of a new List 2 entry. A first implication here is that roots are not UG-given, but only emerge in the process of language acquisition. I remain agnostic to the source of List 1.1 entries, but they may be emergent as well (Biberauer & Roberts 2015).

Second, since grammaticalization and lexicalization are different processes, List 1.1 and List 1.2 can be expanded independently. This may be relevant to the typological variation in “analyticity” (Huang 2014). Languages with a more developed List 1.1 (e.g. Nahuatl) are more synthetic, while languages with a (comparatively) more developed List 1.2 (e.g. Chinese) are more analytic.

Third, the two-step relinking provides an account for semi-grammaticalization. For instance, many languages have complex verbs involving a secondary element with both contentful and aspectual contribution, e.g. in German *ein-schlafen “in-sleep→fall asleep*, *ein* both contributes its literal meaning (metaphorically) and adds a resultative reading, which suggests that a functional category (say, Res) is already existent in List 1.1, but still does not have a dedicated List 2 exponent, because the delinking of */ein/* from \(\sqrt{EIN}\) is obstructed due to the substantial contribution of the root concept. By contrast, the relinking process is complete for */ver/* in *verschlafen “PFV-sleep→oversleep*, which is completely bleached and becomes an exponent for the List 1.1 entry [AKTIONSART: PFV]. NB idiomaticity is assigned to the entire complex label at Spell-Out (Harley 2014b).

In conclusion, the further distribution of the DM Narrow Lexicon into List 1.1 (morpho-syntactic features) and List 1.2 (roots) not only fits into neurolinguistic facts and substantially reduces the size of the “syntax-appendix” Lexicon, but also offers a new perspective to a number of empirical issues (e.g. analyticity, semi-grammaticalization).

**References**


