

[published online as: Hudson, Richard. 2017. Cross-language diversity, head-direction and grammars. Comment on “Dependency distance: a new perspective on syntactic patterns in natural languages” by Haitao Liu et al. doi:<http://dx.doi.org/10.1016/j.plrev.2017.06.005>.]

Cross-language diversity, head-direction and grammars. Comment on “Dependency distance: a new perspective on syntactic patterns in natural languages” by Haitao Liu et al.

Richard Hudson

Department of Linguistics, University College London, Gower Street, London WC1E 7EE

r.hudson@ucl.ac.uk

This paper (Liu, Xu & Liang 2017) – referred to below as ‘LXL’ – is an excellent example of cross-disciplinary work which brings together three very different disciplines, each with its different methods: quantitative computational linguistics (exploring big data), psycholinguistics (using experiments with human subjects) and theoretical linguistics (building models based on language descriptions). The measured unit is the dependency between two words, as defined by theoretical linguistics, and the question is how the length of this dependency affects the choices made by writers, as revealed in big data from a wide range of languages.

LXL’s main claim is that, as speakers and writers, we tend to minimize dependency distance; the comparisons with random structures provide convincing evidence for this claim. LXL also offer an explanation for this tendency: that it is the result of way in which grammars have evolved. As many theoretical linguists claim, this evolution is guided by a multiplicity of functional pressures, and LXL establish that minimization of dependency distance is one of these pressures. They also relate this tendency to the relative processing difficulties encountered by readers and recorded in psycholinguistic experiments, and caused, ultimately, by limitations on working memory.

1. Cross-language diversity

However, the paper also raises an important new question: even if all the languages studied show the same general tendency to minimize dependency distance, why do they produce such different figures for actual dependency distance? Figure 4 in the paper shows mean distances for large corpora in twenty languages ranging from about 2 in Rumanian and Japanese to nearly 4 in German, Hungarian and Chinese. Another study (which LXL cite) shows similar variation (albeit measured differently) between Arabic, with 16.4 for 12-word sentences and Ancient Greek with 26.3 (Futrell, Mahowald & Gibson 2015). If human working-memory capacity is the same across languages, this variation demands an explanation.

A possible explanation, of course, is that working-memory capacity is not, in fact, the same for all speakers. This is, in fact, true; for example, psychological experiments have shown that working-memory capacity is strongly related to intelligence (Engle, Kane & Tuholski 1999) and to age (Miyake & Shah 1999). But, so far as I know, nobody has suggested that this capacity might vary with language, with (say) Chinese speakers having twice the capacity of (say) Japanese speakers. Before we accept this conclusion we need to explore alternatives very carefully.

One alternative explanation is that different language communities demand different degrees of readability. After all, we can’t assume that everything in a written corpus is easy to read. On the contrary, we all know that some writing, especially academic writing, is very hard to read; moreover,

we can be fairly certain that this difficulty is in part due precisely to the parameter measured in this paper: dependency distance. Consequently, we can't take the existence of a sentence in a written corpus as evidence that it is easy to read. What we need, therefore, is a measure of readability which is independent of the focus variable, dependency distance. This is the territory of psycholinguistics, where there is a massive literature on other influences, from word length to the prevalence of specific constructions such as passivization. Fortunately, there are also automatic procedures for measuring readability which could be applied to the corpora under analysis (Pirnay-Dummer 2016), though those developed for English will need to be adapted to other languages. The research question, therefore, is whether dependency distance correlates with other, independent, measures of readability. If it does, then we shall have the interesting conclusion that different communities demand, or allow, different degrees of readability – surely an interesting conclusion in its own right. But if dependency distance does not correlate with readability, then we have to look for a different explanation for the variation between languages.

2. Head-final or head-initial?

Another relevant research question that LXL raise is why head-final languages are so much more common than head-initial languages. This is an area where the same team have done significant work (Liu 2010), showing that directionality can be explored as a continuous parameter using big data, but the fact remains that when typologists apply the traditional classification based on the order of a verb relative to its subject and object (Tomlin 1986:22), they find that about 45% of languages are head-final (SOV or OSV) while only 12% are head-initial (VSO or VOS). The remaining languages (43%) choose the English-type head-medial pattern (SVO or OVS), whose popularity LXL explain convincingly by pointing out that it allows a word to have two dependents adjacent to it; but dependency distance doesn't distinguish head-final and head-initial orders. The differences here suggest that it is harder to process a dependency initiated by the head than one initiated by the dependent; but if that's so, future work on big data should distinguish the two, rather than ignoring the direction of the dependency.

3. Grammars

The last issue that I would like to raise is the balance between grammars and grammar-users. As LXL observe, the pressure to minimize dependency distance can change grammars, with the popularity of head-medial order as a classic example, but it is also only one such functional pressure, so grammars can in fact prevent grammar-users from minimizing dependency distance. For example, in English there is an absolutely rigid rule for the order of indirect and direct objects: indirect before direct; so we can say *She gave her husband a present* but it's absolutely ungrammatical to say **She gave a present her husband* (though, of course, it's fine with *to*: *She gave a present to her husband*.) Moreover, there are a few verbs that don't allow *to* as an alternative; one such is the verb *ask*, which allows *She asked her husband a question* but not **She asked a question (to) her husband*. Whatever the reason for these restrictions, they are part of English grammar, so the grammar-user has very little freedom for minimizing dependency distance in examples like *She asked the man who came to mend the dishwasher which had broken down a question* (with a very long dependency from *asked* to *a question*).

Engle, Randall, Michael Kane & Stephen Tuholski. 1999. Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and

- functions of the prefrontal cortex. *Models of working memory. Mechanisms of active maintenance and executive control*, 102–134. Cambridge University Press.
- Futrell, Richard, Kyle Mahowald & Edward Gibson. 2015. Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences* 112(33). 10336–10341.
- Liu, Haitao. 2010. Dependency direction as a means of word-order typology: A method based on dependency treebanks. *Lingua* 120. 1567–1578.
- Liu, Haitao, Chunshan Xu & Junying Liang. 2017. Dependency distance: a new perspective on syntactic patterns in natural languages. *Physics of Life Reviews*. This issue.
- Miyake, Akira & Priti Shah. 1999. Toward unified theories of working memory. Emerging general consensus, unresolved theoretical issues and future research directions. In Akira Miyake & Priti Shah (eds.), *Models of Working Memory. Mechanisms of active maintenance and executive control.*, 442–481. Cambridge: Cambridge University Press.
- Pirnay-Dummer, Pablo. 2016. Linguistic analysis tools. *Handbook of Writing Research.*, 427–442. New York: Guilford.
- Tomlin, Russell. 1986. *Basic Word Order: Functional Principles*. London: Croom Helm.