Available online at www.sciencedirect.com

ScienceDirect

Physics of Life Reviews 21 (2017) 215–217

PHYSICS of LIFE reviews

www.elsevier.com/locate/plrev

Comment

Projectivity is the mathematical code of syntax
Comment on "Dependency distance: A new perspective on syntactic patterns in natural languages" by Haitao Liu et al.

Anat Ninio

Dept. of Psychology, The Hebrew University of Jerusalem, Israel

Received 15 June 2017; accepted 16 June 2017
Available online 21 June 2017
Communicated by L. Perlovsky

Keywords: Syntax; Dependency distance; Projectivity

The thought-provoking proposal by Liu, Xu and Liang [1] to integrate cognitive considerations involved in Dependency distance minimization (DDM) in syntactic theory will, I am sure, be much discussed in the literature. The authors are to be congratulated for a most serious, thorough and mature presentation. I am sure that DDM will find its rightful place alongside other cognitive factors that are considered to significantly affect the structure and evolution of language such as frequency with its serious impact on grammaticalization [2].

To round off the picture, I would like to add some thoughts on projectivity, a phenomenon said by Liu et al. to be influenced by DDM. I am confident that they are right and that DDM indeed influenced the emergence of projectivity during the evolution of language. No doubt there could have been other solutions to the problem projectivity solves.

This problem stems from the fact that sentences possess two structures. On the one hand, as Liu et al. remind us in Saussure's name, a sentence is always produced or received linearly, one word after another. In graph-theoretical terms [3], the uttered sentence is a one-dimensional object, its structure **a linear section** or part of a line.

The second structure a sentence has is a two-dimensional object, a **directed rooted tree**. In terms of Dependency Grammar, the relation of *government* (or its inverse, *dependency*) is an asymmetrical relation between pairs of words of a sentence [4,5]. It is a function, namely, a many-one mapping relation which is directional (or anti-symmetrical), anti-reflexive, and anti-transitive. The application of this ordering relation to the words of a well-formed sentence creates a structure represented by a directional graph of the tree type; it is a two-dimensional abstract object, totally connected; it has a single root and no cyclicity [3].

According to dependency theory, the transferal from the two-dimensional tree-structure of a sentence to its one-dimensional phonetic chain-structure is regulated by a single rule of the grammar, that of adjacency or projectivity [6,7]: A dependent must appear in a sentence immediately adjacent to its head except that the two may be separated by dependent(s) of either words. This rule is applied recursively, so that if the inserted dependent has a dependent of

DOI of original article: http://dx.doi.org/10.1016/j.plrev.2017.03.002.
E-mail address: Anat.Ninio@huji.ac.il.

http://dx.doi.org/10.1016/j.plrev.2017.06.021

its own, the latter may in turn be inserted between its own head and *the head's* head. The result is that the deeper the dependency relation between a pair of words on the tree (the further away from the root), the closer are the words in the linear sentence. In other words, the degree of proximity of two grammatically related words in the sentence-chain encodes the relative subordination of the relevant dependency relation, in comparison to other relations either word participates in.

This linguistic definition of adjacency lends itself to mathematical formalization. As the term projectivity implies, mathematically speaking, the planar graph which is the tree projects to the one-dimensional object which is the line, namely, it undergoes a projective transformation from two dimensions to one dimension. The transformation folds and packages the connected nodes of the two-dimensional tree in such a way that the depth of the edges relative to the root is translated to distance; the deepest the edge on the tree, the smallest the distance. This makes it possible for a hearer to reconstruct the structure of the tree from that of the line by applying the projectivity transformation the other way round, to translate dependency distances to relative depth in the tree.

This transformation is of course reminiscent of drawing in perspective: the further an object from the viewer in the three-dimensional space represented by the picture, the smaller it is drawn in the picture-plane. Drawing in perspective is the visual representation of the mathematical transformation of projective geometry whereby a three-dimensional object is projected onto a two-dimensional plane. This method was invented by Renaissance painters. It was a mathematician of the 15th century, Leon Battista Alberti, who first formulated the mathematics of showing distance in painting based on planar projections. Eventually, this lead to the development of analytic geometry [8]. Most relevantly to our topic, Alberti was a cryptographer. A method for encoding an n-dimensional object in an $n-1$ dimensional one in a way it can be reconstructed is a kind of encryption, enabled by the mathematics of projective geometry.

It follows that the grammatical rule of adjacency or projectivity is nothing less than the expected mathematical solution for a problem of connecting between a two-dimensional and a one-dimensional representation of the syntactic structure of the sentence. It might be the only viable systematic method of packaging the hierarchical structure of the tree into the string so that it can be unpacked back into the tree. There are two arguments for such a view. First, projectivity is pervasive, ubiquitous, and probably universal. The great majority of languages and the great majority of sentences within these languages follow it [9–12]. Although it is claimed that there are some syntactic patterns with long-distance or crossing-dependency patterns such as topicalization, wh-fronting and extraposition, in actuality there are elegant theoretical solutions that do away with the discontinuity and long-distance dependencies in such patterns [13,14].

The second argument for the robustness of the projectivity method comes from Natural Language Processing. It appears that parsing algorithms allowing unrestricted non-projective structures are unsolvable in terms of time [15–17]. Such problems are called in the parser literature computationally intractable, NP-hard, or NP-complete. Taking parsers as models of human parsing, we might say that attempts to resolve the structure of sentences without using the projectivity principle are doomed to fail as our cognitive resources are insufficient for alternative algorithms.

To conclude, projectivity appears to be not so much a side-effect of DDM as a mathematical requisite for a method to encode a two-dimensional tree in a one-dimensional sentence-string in a way that makes reconstruction possible. It is an intriguing idea that language evolved to make use of this method so that it fits our cognitive requirements for a minimal dependency distance in the sentence. It is interesting to speculate that dependency itself is an outcome of the same process. Let us hope that Liu et al. continue to explore this fascinating topic.

References

[1] Liu H, Xu C, Liang J. Dependency distance: a new perspective on syntactic patterns in natural languages. Phys Life Rev 2017;21:171–93. http://dx.doi.org/10.1016/j.plrev.2017.03.002 [in this issue].
[2] Bybee J, Perkins R, Pagliuca W. The evolution of grammar: tense, aspect and modality in the languages of the world. Chicago: University of Chicago Press; 1994.
[3] Tutte WT. Graph theory. Menlo Park, CA: Addison-Wesley; 1984.
[4] Mel'cuk IA. Studies in dependency syntax. Ann Arbor: Karoma; 1979.
[5] Hudson R. Word grammar. Oxford: Basil Blackwell; 1984.
[6] Mel'cuk IA. Dependency syntax: theory and practice. Albany, NY: State University of New York Press; 1988.
[7] Robinson JJ. Dependency structures and transformational rules. Language 1970;46:259–85.
[8] Andersen K. The geometry of an art: the history of the mathematical theory of perspective from Alberti to Monge. New York: Springer; 2007.

[9] Nivre J, Nilsson J. Pseudo-projective dependency parsing. In: Proceedings of the 43rd annual meeting of the association for computational linguistics; 2005. p. 99–106.
[10] Futrell R, Mahowald K, Gibson E. Large-scale evidence for dependency length minimization in 37 languages. Proc Natl Acad Sci USA 2015;112(33):10336–41.
[11] Ferrer i Cancho R. Why do syntactic links not cross? Europhys Lett 2006;76(6):1228.
[12] Kuhlmann M. Mildly non-projective dependency grammar. Comput Linguist 2013;39(2):507–14.
[13] Hudson RA. Language networks: the new word grammar. Oxford University Press; 2007.
[14] Osborne T, Putnam M, Groß T. Catenae: introducing a novel unit of syntactic analysis. Syntax 2012;15(4):354–96.
[15] McDonald R, Nivre J. Characterizing the errors of data-driven dependency parsing models. In: Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning. Association for Computational Linguistics; June 2007. p. 122–31.
[16] McDonald R, Satta G. On the complexity of non-projective data-driven dependency parsing. In: Tenth international conference on parsing technologies; 2007. p. 121–32. Available from: http://dl.acm.org/citation.cfm?id=1621426.
[17] Neuhaus P, Bröker N. The complexity of recognition of linguistically adequate dependency grammars. In: 35th annual meeting of the association for computational linguistics; 1997. p. 337–43.