



Comment

Language as a whole – A new framework for linguistic knowledge integration

Comment on “Approaching human language with complex networks” by Cong and Liu

Xinying Chen^{a,b,*}

^a School of International Study, Xi'an Jiaotong University, China

^b Text Technology Lab, Department of Computer Science and Mathematics, Goethe University, Germany

Received 10 July 2014; accepted 17 July 2014

Available online 25 July 2014

Communicated by L. Perlovsky

Researchers have been talking about the language system theoretically for many years [1]. A well accepted assumption is that language is a complex adaptive system [2] which is hierarchical [3] and contains multiple levels along the meaning-form dimension [4]. Over the last decade or so, driven by the availability of digital language data and the popularity of statistical approach, many researchers interested in theoretical questions have started to try to quantitatively describe microscopic linguistic features in a certain level of a language system by using authentic language data. Despite the fruitful findings, one question remains unclear. That is, how does a whole language system look like? For answering this question, network approach, an analysis method emphasizes the macro features of structures, has been introduced into linguistic studies [5]. By analyzing the static and dynamic linguistics networks constructed from authentic language data, many macro and micro linguistic features, such as lexical, syntactic or semantic features have been discovered and successfully applied in linguistic typographical studies so that the huge potential of linguistic networks research has revealed [6].

What is particularly interesting about recent development in this area is that researchers have been able to systematically analyze linguistic features above the sentence level since the network approach break through the limitation of traditional linguistic features annotated based on lexical or sentences level in corpus. So that linguistic networks, as the representations of the whole body of language data, may be a better approach to explore human language systems.

Another exciting potential of network approach, maybe even a more important one, is that it provides a universal framework which can integrate different linguistic studies. The majority of linguistic comparative studies have tended to focus on linguistic features in certain level, such as morphological, syntactic or semantic level. The reason is that there are notable difficulties in across-levels comparison studies since different terms and linguistic features are used to describe language structures in different levels. It is important to solve this problem so we are able to explore the universality and diversity between different language levels as well as the connections between them. Only in this way

DOI of original article: <http://dx.doi.org/10.1016/j.plrev.2014.04.004>.

* Correspondence to: School of International Study, Xi'an Jiaotong University, China.
E-mail address: chenxinying@mail.xjtu.edu.cn.

can we truly understand the essence of language as a system. Under the network framework, the structure features can be observed and compared despite the across-levels differences between linguistic units and their features. As Liu and Cong [7] tried to compare the network structures of different levels of modern Chinese and then find the inner connections between these levels by analyzing different linguistic networks, representing different language levels, constructed from the same body of authentic language data, real across-levels comparative research can be done in similar way for different languages.

The same framework of network approach can also be applied in not only theoretical linguistic research but also in many other different branches or application driven studies. For instance, different features of linguistic networks have been successfully used into language classification studies [8,9] and strong correlations between text quality and complex network features of students' articles were also found by researchers [10]. The network model also can be used to measure, imitate and predicate the language evolution process [11]. Studies using network approach can be widely found in machine translation, text classification, second language learning, child language acquisition and many other linguistic areas. This makes it is possible that to integrate findings from different linguistic branches in the same framework and then make closer connections between different linguistic branches.

As a scientific analysis method, network approach is naturally close to other disciplines such as neuroscience, communication science, genetics, statistics and physics. Linguistic network studies are in favor of bring closer relationships between linguistics and these subjects, especially the connections with neuroscience which is dealing with the neuro-system of the human brain that was recognized as the biological foundation of the human language.

Although the potential of network-based language research seems huge, the reality is that existing studies are dispute in different linguistic branches and they are lacking of strong connections to each other and previous linguistic studies. It depends on the future development that whether the linguistic network research can achieve these potential. But as showed by Cong and Liu [6], it already has had a promising and good start.

References

- [1] Kretzschmar WA. *The linguistics of speech*. Cambridge University Press; 2009.
- [2] Beckner C, Blythe R, Bybee J, Christiansen MH, Croft W, Ellis NC, et al. Language is a complex adaptive system: position paper. *Lang Learn* 2009;59(Supplement 1):1–26.
- [3] Sgall P, Hajičová E, Panevová J. *The meaning of the sentence in its semantic and pragmatic aspects*. Dordrecht: Reidel Publishing Company; 1986.
- [4] Mel'čuk I. *Dependency syntax: theory and practice*. Albany: State University of New York Press; 1988.
- [5] Solé R. Syntax for free? *Nature* 2005;434:289.
- [6] Cong J, Liu H. Approaching human language with complex networks. *Phys Life Rev* 2014;11:598–618 [in this issue].
- [7] Liu H, Cong J. Empirical characterization of modern Chinese as a multi-level system from the complex network approach. *J Chin Linguist* 2014;42:1–38.
- [8] Abramov O, Mehler A. Automatic language classification by means of syntactic dependency networks. *J Quant Linguist* 2011;18(4):291–336.
- [9] Liu H, Cong J. Language clustering with word co-occurrence networks based on parallel texts. *Chin Sci Bull* 2013;58:1139–44.
- [10] Antiqueira L, Nunes MG, Oliveira Jr ON, Costa LDF. Strong correlations between text quality and complex networks features. *Physica A* 2007;373:811–20.
- [11] Liang W, Shi Y, Huang Q. Modeling the Chinese language as an evolving network. *Physica A* 2014;393:268–76.