

Analyzing language development from a network approach

Jinyun Ke

English Language Institute, University of Michigan

Keywords

Abstract

Recently network analyses have seen a rapid growth in various areas. It has been shown that analyzing systems as networks will provide new insights into old questions (Barabási 2002; Buchanan 2002). Inspired by these developments, we apply network analyses in the study of language acquisition. Network measures provide us new quantitative indices to examine the development in individual children, as well as individual differences among children and the adult-child interaction in the course of development. We have worked on some English data from the Manchester corpora in CHILDES (Ke & Yao 2005). Figure 1 illustrates some findings from the analyses of the networks at different MLU stages of four children in the corpora: while the networks of children all approach to those of the adults (a large network size and a large connectivity), different children go along different paths: some children are fast in learning new words, but slow in exploiting the words; while some others learn new words slowly, but tend to apply the known words more actively. While most mothers' networks have a high correlation with those of their children, there are different interaction strategies in talking to children: while some mothers show simultaneous growth in the complexity of their speech along their children's development, some appears to be more stable in their speech style (in terms of vocabulary size and word combinations), such as Carl's mother shown in Figure 1.

While traditional corpora analyses have paid a lot of attention to the frequency of words (Ellis 2002), network analyses show that the frequency of words may not be well correlated with other properties of the words, such as degree and centrality. As word frequency mostly depends on the properties of the individual words (such as semantic salience and/or grammatical function), degree and centrality reflect better the global structure of the lexical network. Moreover, these measures require little syntactic analyses of the data, and some of syntactic structure or words' syntactic features may emerge from the global structure of the networks. For example, we find in the English data that the articles "the" and "a" shift from "hub" to "authority" during children's development, while they stably appear as "authority" in adult's networks.

In this paper, we will apply similar analyses on some French acquisition data from CHILDES and from the Lyon's DDL research group, and make comparison with the English data. In particular, we will examine the development of articles in French, to see if "le", "la", "les" have the similar shifting process as English's "the" and "a". Also we will explore the adult-children interaction with more complex network measures.

References:

Barabási, A-L. (2002). *Linked: the New Science of Networks*. Cambridge, Mass. Perseus Pub.

Buchanan, M. (2002). *Nexus: Small Worlds and the Groundbreaking Science of Networks*. W. W. Norton & Company.

Ellis, N. C. (2002). Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143-188.

Ke, Jinyun and Yao Yao. (2005). A study on language development from a network perspective. Under review.

