Tipping Points

Charles Yang, University of Pennsylvania

The human memory is impressively large and capable of storing detailed linguistic information. These findings raise important questions for the role of the grammar, as what can be stored needn't be computed. But evidence from language acquisition that not only is the grammar necessary, children are extremely adept at identifying the types of linguistic patterns (e.g., exceptions and rules) and learning them differently.

The main thrust of this work is to develop a calculus (Yang 2005), one which builds on the third factor of efficient computation (Chomsky 2005), that weighs in on the balance between storage and computation. We suggest that grammar/rule emerges at certain tipping points, where the number of rule-following items greatly exceeds the number of exceptions. The calculation of the tipping point is supported by psycholinguistic evidence, which suggests that exceptions cause delay in the real time computation of rules. Under very general assumptions, it is possible to derive that a rule/process applicable to N items can tolerate no more than N/ln(N) exceptions before losing productivity, i.e. all items are subject to lexicalized storage.

We present two classes of evidence to showcase tipping points in language.

First, it is well known that 80-90% of English words are primarily stress initial (Cutler & Carter 1987), yet no theories of English metrical stress, or English learning children, treat English as a quantity insensitive system. Transparently, a statistical majority of 80-90% does not guarantee productivity. Through a corpus study of child directed English, we show that the tipping point model accounts for the developmental stages in metrical stress acquisition, while supporting Halle's theory of English stress (1998).

Second, it is instructive to study cases where grammar fails, a most prominent case being the so-called "paradigmatic gaps" (e.g., stride-strode-*strode/?*stridden). Gaps emerge due to the absence of productive process, which arises when exceptions exceed the critical tipping point (N/ln N). We show that several well known inflectional gaps in English, Polish, Spanish and Russian are predictable on purely numerical basis.