

On the Productivity of Linguistic Processes

A current trend in linguistic theorizing returns to the construction-specific approach in the 1960's (Goldberg 1995, Culicover 1999, Goldberg & Jackendoff 2004, etc.) To use Culicover's terminology, language contains "nuts": periphery constructions in the sense of Chomsky (1981) that are restricted to specific lexical items but nevertheless interact with both general and language-specific principles. There is simultaneous work in language development that stresses the nonproductivity in children's early words and grammar, while suggesting that language acquisition is a piecemeal inductive process (Tomasello 2000, 2003). In light of these findings, the proponents of this line of work argue that the core syntax system be abandoned, along with the overarching parameters intended to cover a broad range of syntactic facts.

This paper is a preliminary report on how one might take up the "nuts" challenge. On the one hand, there are good theoretical reasons for not abandoning the parameter system that sharply delimits the hypothesis space of the learner: to wit, the formal results on unconstrained learning have been overwhelmingly negative (Kearns & Vazirani 1994). (See Fodor 2001 for discussion.) On the other, the "nuts" challenge is real: those who wish to retain the core parameter system must articulate a theory that keeps the core parameter values separate from the more idiosyncratic periphery constructions. In other words, a parameter value must be able to tolerate a certain amount of exceptions: e.g., an American child should not conclude from "___ seems nice" that English is a pro-drop language.

Our point of departure is another sort of "nuts" in language, namely, the case of past tense in English morpho-phonology. Few doubt that there is a "core" rule of adding -d for inflection (Pinker 1999). Moreover, recent work (Yang 2000, Albright & Hayes 2003) has shown the irregulars to form mini classes of regularity that resemble the "nuts" phenomena in syntactic investigations. Note that the postulation of the -d rule must tolerate over 100 irregular verbs that are exceptions. By contrast, no one learns that, for instance, inflects verbs ending in *-ow* to *-ew* in past tense (along the lines of *blow-blew*, *know-knew*, *grow-grew*). In the latter case, apparently, the child has seen too many exceptions to the *ow->ew* "rule" and decided that enough is enough.

This paper addresses the question when "enough is enough". That is, there is a maximum number of exceptions that a rule can tolerate; once this threshold is crossed, the rule becomes strictly lexical, i.e., non-productive. Similarly, if a syntactic regularity is not applicable to a sufficient large number of constructions, it would be relegated to the status of nuts and its productivity limited to a specific lexical context; otherwise, it will be built into the core syntax system.

Our instantiation of this insight is both formal and empirical. We conjecture that the organization of the lexicon is governed by a principle of economy, which minimizes the time complexity of lexical access and retrieval. This readily leads to a computational model of the lexicon which captures the well-known frequency effects in lexical processing. Moreover, the data structure of the model is amendable to the so-called online algorithms in computer sciences. We describe a theorem which informs the learner the learner when a rule is productive and when it is unproductive and restricted to a fixed list of lexical items. (See Appendix.) Extensive evidence from lexical corpora will be presented to support the theoretical results. While we have not derived formal results for the productivity of syntactic constructions, the spirit of the analysis carries over, and we will report simulation results using realistic child-directed data from CHILDES.

APPENDIX

Let N be the total number of words that follow certain structural description (e.g., nouns ending in sibilants). Let M be the number of those that follow a potential rule R in certain context: (e.g., adding -iz in plural forms). And let $N-M$ be the number of “nuts” that do not follow rule R : in this example, the number of nouns that end in sibilants but do not add -iz in plurals, e.g., “fish”, “tooth”, etc.

Assume a list processing model of the lexicon which arranges the words according to their frequencies: those more frequent are placed on higher positions of the list, and are thus accessed faster than those less frequent items lower on the list. This naturally captures the frequency effects in lexical processing. Moreover, the adjusting of the list with respect of word frequency can be carried with little computational effort--hence psychological plausibility--while maintaining near optimal efficiency (Rivest 1976).

If R is productive, then the $(N-M)$ exceptions must be listed before the M items that do follow R ; write the expected retrieval time under the lexicon with R as $T(N, N-M)$. On the other hand, if R is not productive, then all N items must be listed exhaustively: write the expected time retrieval in this case to be $T(N, N)$.

Conjecture: If $T(N, N-M) < T(N, N)$, then R is productive (one rule plus $N-M$ nuts)
otherwise, R is unproductive (all N words are treated as nuts)

Under the reasonable Zipfian assumptions of word frequency, we can solve M as a function of N : a rule can tolerate a surprising small number of nuts.

References

- Albright, A. & Hayes, B. (2003). Rules vs. analogy in English past tenses. *Cognition*, 90: 119-161.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Foris.
- Culicover, P. (1999). *Syntactic nuts*. Oxford U. Press.
- Fodor, J. D. (2001). Parameters and the periphery: reflections on *Syntactic Nuts*. *J. Linguistics*. 37, 367-392.
- Goldberg, A. (1995). *Constructions*. U. Chicago Press.
- Goldberg, A. & Jackendoff, R. (2004). The English resultatives as a family of constructions. *Language*. 532-568.
- Kearns, M. & Vazirani, U. (1994). *An introduction to computational learning theory*. MIT Press.
- Pinker, S. (1999). *Words and rules*. Basic Books.
- Rivest, R. (1976). On self-organizing sequential search heuristics. *C. ACM*. 19 (2), 63-67.
- Tomasello, M. (2000). Do young children have adult syntactic competence. *Cognition*, 74, 209-253.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard U. Press.
- Yang, C. (2000) *Knowledge and learning in natural language*. PhD thesis, MIT. Published in 2002, Oxford U. Press.