A Simple Model of Encoding and Judgment about Non-Adjacent Dependencies

Pablo Gómez (pgomez1@condor.depaul.edu)
Trisha Hinojosa (tripp@condor.depaul.edu)
Department of Psychology, 2219 N Kenmore
Chicago, IL 60614 USA

Studies with adult and infants have shown that subjects can learn fairly complex probabilistic relationships. Researchers have used statistical learning as a laboratory to explore issues like word segmentation (Saffran, Aslin & Newport, 1996) and the acquisition of grammar (Morgan, Meier & Newport, 1987). Statistical learning has become a scenario for an argument between the two competing views about language acquisition: the view that assumes that humans have some innate ability to acquire grammar (cf., Chomsky, 1965); and the view that claims that statistical learning is based on the same learning mechanisms (e.g., distributed supervised learning) as other domains (see Seidenberg, 1997).

Of particular relevance to our research program are the studies that focus on learning of relationships between non-adjacent speech elements. Newport & Aslin (2004) and R. Gomez (2002) have shown that, only under some special circumstances, participants learn relationships between non-adjacent speech elements (e.g., syllables and words). Here, we present the first version of a model that can account for that data. The model uses a simple encoding process, and a decision mechanism inspired in signal detection theory (Green & Swets, 1966). Our mode supports the notion that very simple mechanisms are enough to explain non-adjacent dependency learning without resorting to special language learning modules.

R. Gomez (2002) has shown that adults and infants can learn non-adjacent regularities when the set size of the intermediate element is large (24 elements), but not when the size set is small (e.g., 2). In a follow up study, she showed that participants could learn non-adjacent dependencies if the intermediate element set size was 1. In these studies, words from an invented language were used in utterances of the form $a_1\ X_{1\ to\ N}\ b_1$, where the dependency was between elements (words) a and b, and the set size of the intermediate element was N.

Newport & Aslin (2004) showed that participants could learn non-adjacent dependencies between letters, but not between syllables. For their experiment with syllables, the stimulus had the form $CV1_1\ CV2_{1\ to\ 4}\ CV3_1$, where the dependency was between the consonant-vowel syllables CV1 and CV3. For the experiment with letters, their experiment had the form $C_1V_{1\ to\ 2}\ C_2V_{3\ to\ 4}\ C_3V_{4\ to\ 5}$, where the dependency was between the consonants C_1, C_2 and C_3.

Description of the Model

The model assumes that subjects use a minimalist approach when they encode the training stimuli. If in their subjective estimation, the adjacent (first order) relationships are informative about the rules to form the artificial language, they will tend not to encode the nonadjacent (second order) relationships.

How informative the first order relationship (say, between the first and second elements in Gomez's studies) is can be determined by a very simple computation:

$$I_{X,b}=p(X_j|a_i)\ (1-p(X_j|a_i)), \qquad (1)$$

where $I_{X,b}$ is a measure of how informative the first order relationship is, and $p(X_j|a_i)$ is the estimated conditional probability of element X_j given element a_i. This measure of informativeness can be thought of as the probability to encode the next order of (non-adjacent) relationship.

The grammaticality judgments are based on familiarity (cf. Signal Detection Theory) at the order of relationship that the learner estimated as informative using Equation 1.

This simple model can account for the u-shaped pattern of accuracy that Gomez found as a function of set size in the intermediate component. In addition, it accounts for the difference between the syllable and letter conditions found by Newport and Aslin; this, because the first order relationship between consonants and vowels had some level of subjective informativeness in the syllable condition.

Acknowledgments

Funded partially by a grant of the URC-DePaul University.

References

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Gomez, R. (2002). Variability and detection of invariant structure. *Psychological Science, 13,* 431-436.

Green, D. M. and Swets, J. A. Signal detection theory and psychophysics. New York: Wiley.

Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of intonational and morphological marking of phrases to the acquisition of language. *Cognitive Psychology, 19*, 498-550.

Newport, E. L., & Aslin, R. N. (2004). Learning at distance. I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48,* 127-162.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926-1929.

Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science, 275,* 1599-1603.