Human Infant Speech Perception and Language Acquisition: Rules vs. Statistics

Janet Werker (UBC),
Gary Marcus (NYU),
Jacques Mehler (University of Trieste),
Helen Neville (University of Oregon),
Nuria Sebastian-Galles (Universitat de Barcelona)

March 18–20, 2004

During the first year of life, infants become exquisitely tuned to the properties of the native language. This includes sensitivity to the consonant and vowel sounds of the native language, the acceptable sequences of these sounds, the rhythmical properties of the native language, and the multiple cues which signal boundaries between words, phrases, and sentences. Perceptual sensitivity to these properties is essential for word learning, acquisition of grammar, and language acquisition more generally. The predominant theoretical approach guiding this work has, until recently, stemmed from the joint contribution of Chomsky—arguing for a rule based, symbol manipulation, foundation for language acquisition rooted in targeted evolution, and Lenneberg—arguing that language is supported by special areas in the brain. In the past several years, however, a number of studies have documented the role of statistical learning in accounting for the changes in perceptual sensitivity seen in infancy. In this workshop, we explored and debated the extent to which statistical learning vs. rule learning might help account for both perceptual tuning, and subsequent mapping of tuned perceptual categories to the task of language acquisition.

The first evening of the workshop, Wednesday evening, Richard Aslin, a cognitive scientist from the University of Rochester, provided a broad overview of work in statistical learning in infancy. He began with an explication of the original work showing that infants can use transitional probabilities (TPs) to segment words from a stream of speech. In this work, infants were presented with continuous strings of syllables in which the only distinguishing information for positing a "word" was the TP between syllables. Among the strings were three syllable items that had TPs between syllables of 1.0, and three syllable items that had TPs of .33. Following a 3–4 minute familiarization period, infants of 7 months had used these TPs to pull out "words", choosing to listen longer (as indicated by looking time) to "part words" (syllable strings with TPs of only .33) over "words" (strings with TPs of 1.0). Aslin presented a number of updates.

Two kinds of studies generated the most discussion in the debates which began the next morning. 1) those studies comparing the learning of transitional probabilities for linguistic stimuli to learning the transitional probabilities for non-linguistic stimuli such as tones and for visual images. These studies are of interest because they challenge the notion that language learning relies on specialized learning mechanisms, and raise the alternative possibility that generally available learning mechanisms are used in the service of language acquisition. 2) those studies examining the conditions under which infants can learn non-adjacent dependencies. Whereas adjacent TPs are important in word segmentation (and also in learning some aspects of morphology), non-adjacent TPs are needed if this learning mechanism can contribute to our understanding of how other aspects of morphology as well as syntax are acquired. A number of studies indicate that the learning of non-adjacent dependencies is much more difficult than is learning of adjacent dependencies. Success requires one of a number of manipulations; the insertion of a brief (even undetectable) silent interval between the syllables, high variability of the non-crtierial intervening syllables, and/or another linguistic cue such as syllable stress. Vigorous debate centered on the question of why these types of manipulations facilitate learning. Is it because they simplify the pattern detection process that results from the use of statistical regularities? Or is it because these manipulations turn the task from one of statistical learning into one of rule learning? This debate centers on the question of whether language has a specialized rule-based foundation which enables symbol manipulation.

Two other statistical learning mechanisms were discussed. One was how learning of distributional regularities could account for the building of language appropriate phonetic categories. It was shown recently (by members of our group) that infants of 6-8 months can use distributional regularities to change their phonetic categories. If presented all members of an 8 step continuum, but with more instances of stimuli 2 and 7 in one group (the bimodal group) and more of 4 and 5 in the other group (the monomodal group), infants in the bimodal group divide the continuum into two perceptual categories whereas infants in the monomodal group collapse it into one. New data on the types of information this distributional learning might generalize to, and the ages at which infants might have this mechanism available, were discussed. The other statistical mechanism discussed was associative probabilities. New data were presented showing that learners track the probabilities of word-object associations almost perfectly. Using fMRI data, possible neural systems underlying this learning were discussed, and will be compared the neural mechanisms used in learning other probabilistic associations.

Computational Modelling was another focus of the workshop. A number of researchers in the world are using a variety of computational models to account for and predict the regularities seen in language acquisition. The unique new approach from Gary Marcus in our group is to use computational techniques to formally model the newest findings in developmental neurobiology that might help explain how a brain can become organized for complex cognitive tasks such as language acquisition. The goal behind this work is again to address the fundamental theoretical controversies outlined at the beginning of this report that underly research in language acquisition.

The theoretically guided presentations and debates described above, were informed throughout the meeting with new data in a number of areas which then spawned their own advances in theory. One content area focused on phonetic and phonological perception in bilingual infants and adults, and the link between perceptual changes in infancy and subsequent word learning. This led to the presentation of a new theoretical framework - “shallow” (perceptual) vs. “deep” (functional linguistic use) learning.

A second content area focused on new data showing the different languages young infants can discriminate based on rhythmical properties. Included in the presentation were new steps in the quantification of languages based on rhythmical properties (% vowel and Delta consonant in the words of the language). The novel theoretical contribution here is whether these rhythmical properties are correlated with, and predictive of, the underlying word order of the language. If so, this would be another way in which learning the acoustic properties of the language could help bootstrap its acquisition.

The research presented relied on both behavioural and neuroimaging techniques. One portion of the workshop therefore involved presentation of new team-based advances in the use of these techniques. The two most novel advances discussed were 1) the improvement of optical imaging (near infrared spectroscopy) for imaging the infant brain while listening the language, and 2) the development of a new ERP (electrophysiological, recording of electrical activity over the brain) signature of word segmentation.

The opportunity to bring so many researchers together from so many different fields was essential to the success of this workshop. Moreover, with the help from BIRS, we ensured broad graduate student and postdoctoral fellow involvement as well. The background infrastructure support provided by BIRS was so appreciated, and helped make the meeting a resounding success.