IMPLICATIONS FROM INFANT SPEECH STUDIES ON THE UNIT OF PERCEPTION*

Peter W. Jusczyk
Department of Psychology, University of Oregon, Eugene, Oregon 97403, USA

One of the great overlooked tasks of language acquisition is how the language learner develops a representation of the sound properties of a word that allows for the recognition of words in fluent speech. Thus, while child phonologists have given a great deal of attention to the nature of the representations that underlie the child's earliest productions of speech (e.g. Ferguson, 1986; Macken, 1980; Menn, 1980; Vihman, 1978), and other child language specialists have focused on the structure of the child's semantic categories, much less is known about what sort of representation of the sound structure of words permits their comprehension by the child (cf. Jusczyk, 1985). Yet, one could argue that learning the essential phonemic characteristics that distinguish between one name and another in one's native language is at least as important as distinguishing the boundaries of potential referents for the names. Clearly, to be a fluent speaker-hearer of a language, one needs to develop the appropriate categories for both the sounds and the meanings of words.

What little information is available about the nature of underlying phonetic representation during the early stages of language acquisition comes from two sources: infant speech perception experiments and word learning studies with children during the second year of life. At first glance, the information from these two sources appears to be contradictory. The studies of infant speech perception (see Aslin, Pisoni & Jusczyk, 1983 for a review) indicate an organism with very fine discriminative capacities, capable of distinguishing contrasts that might occur in any language. In contrast, the studies of early word learning (e.g. Shvačkin, 1973; Garnica, 1973) suggest that the process of acquiring phonemic contrasts occurs a step at a time over the course of many months. Although this discrepancy might simply be attributable to the sensitivity of the tasks employed, a closer inspection suggests that the tasks are tapping different kinds of capacities. A typical infant speech perception experiment, employing a procedure like high amplitude sucking (HAS), is a same-different discrimination task. In order to be credited with discriminating some contrast (even between items that differ by a single phonetic feature) the infant only needs to note that the preshift stimulus differs in some way from the postshift stimulus. The infant need not register the

*The research reported here was supported by a research grant to the author N.I.C.H.D. (#HD-15795). The author also whishes to thank Ranka Bijeljac- Babic, Josiane Bertoncini, Patricia Kuhl, Jacques Mehler, Joanne Miller, Deborah Kemler Nelson, Robert Remez and Janet Werker for comments they made on a previous version of this manuscript.
way in which the stimulus pair differ, only that they are not identical. For this purpose, a holistic comparison of the overall similarity of the stimuli may be sufficient for discrimination.

A word-learning type task presents the child with a different type of situation. In such a task, the child is often presented with a number of novel objects, each of which has an unfamiliar name. The child must learn which name goes with which object. Here the child cannot succeed by simply making holistic same-different judgments since the contrast is not simply between a pair of items but among many different ones that differ along a number of dimensions. In order to succeed on such a task, the child must develop some sort of representation of the sound properties of the name that allow it to be picked out and identified from all the other possible names that might be uttered in that situation. Presumably the representation which the child must develop for this purpose is related in some way to the underlying speech perception capacities that allow infants to distinguish speech sounds from one another.¹

In what ways might the infant's speech perception capacities be related to the kinds of representations used to recognize words in fluent speech? At the very least, the basic speech perception capacities provide some sort of psychophysical lower-bound on the kinds of representations that one could form concerning the sound structure of words in a language. But this need not imply that the representations themselves are packed with such fine-grained information about all the potential phonetic distinctions that could arise between words. In fact, given the goal of arriving at a maximally efficient representation of the sound structure of words in a language--one that allows them to be rapidly and accurately identified during fluent speech--it seems likely the ideal representation would be one with only enough detail to distinguish between possible words in a language. This suggests an additional source of constraint upon the lexical representations developed by the language learner, viz. the phonological structure of the target language itself. Knowledge of which sound categories from the inventory of those occurring in a particular language mark out meaningful distinctions, as well as the permissible sequences of these categories (i.e. what phonotactic constraints exist) is useful in developing the most efficient representations of the acoustic characteristics of lexical items. In particular, information about which strings of phonemes constitute legal sequences for forming words in a language provides important information that could be employed in segmenting continuous speech into discrete words. Of course, information about the prosodic patterns, phonemic categories (and the allophones that map into each category) and phonotactic constraints of a particular language is present in the input that the language learner receives. Consequently, such regularities observed in the nature of the input during the course of language acquisition may well influence the way in which speech sounds are represented by infants. This being the case, the infant would be expected to move from a general representation of utterances to a language-specific one.

Thus, in order to understand how the infant becomes a recognizer of fluent speech we need to know something about (1) their underlying sensory capacities, (2) the kinds of perceptual representations they employ, and (3) the way in which knowledge of