# Viewing The Development of Speech Perception As An Innately Guided Learning Process

Peter W. Jusczyk and Josiane Bertoncini
*Language and Speech* 1988 31: 217
DOI: 10.1177/002383098803100301

# VIEWING THE DEVELOPMENT OF SPEECH PERCEPTION AS AN INNATELY GUIDED LEARNING PROCESS*

PETER W. JUSCZYK
*Laboratoire de Sciences Cognitives et Psycholinguistique*
*C.N.R.S. and E.H.E.S.S., Paris*
*and*
*Department of Psychology, University of Oregon*

and
JOSIANE BERTONCINI
*Laboratoire de Sciences Cognitives et Psycholinguistique*
*C.N.R.S. and E.H.E.S.S., Paris*

*(Received May 24, 1988; accepted August 9, 1988)*

The present paper examines the issue of how speech sounds may be treated by infants as special signals relative to other kinds of acoustic stimuli. Consideration is given to the view that the mechanisms underlying the infant's perception of speech are specialized for this purpose. Some of the difficulties of providing definitive evidence for or against this position are noted. Then the thesis is advanced that special processing of speech may lie in the inherent salience which such sounds have for infants. In particular, it is suggested that the development of speech perception follows the course of an innately guided learning process. One key assumption of this view is that speech sounds may be more apt to attract attention or have a higher priority for further processing than other types of acoustic signals. Recent evidence from a number of new paradigms for studying infant speech perception is reviewed in light of this position. The paper concludes with a discussion of how findings concerning the development of speech perception during the first year of life fit with the innately guided learning view.

*Key words:* speech perception, infants, development, innateness

## INTRODUCTION

To what extent are there specialized processing mechanisms associated with the infant's perception of speech sounds? This kind of question has been of concern to researchers ever since the first studies of infant speech perception capacities were under-

* Preparation of this manuscript was supported in part by an NICHHD grant (#HD-15795) to the first author. We would like to thank Mike Posner, Asher Cohen, and Ann Marie Jusczyk for comments that they made on an earlier version of the paper. In addition, we wish to extend our gratitude to Jean-Pierre Changeux, Jacques Mehler, and Juan Segui for the fruitful discussions which we had with them on this topic.

Correspondence should be addressed to Peter W. Jusczyk, Department of Psychology, University of Oregon, Eugene, OR 97403, U.S.A.

taken (e.g., Eimas, Siqueland, Jusczyk, and Vigorito, 1971). Today, despite a considerable accumulation of data, we find that the question is far from being answered. In part, this is because the debate about special speech perception capacities in infants has tended to focus almost exclusively on the nature of the mechanisms that mediate the discrimination of phonetic contrasts. For reasons that we outline below, we believe that the available experimental methods and the data derived from them will not permit a clear answer to the question as to whether phonetic or general auditory mechanisms are responsible for the infant's discrimination of speech contrasts during the first few months of life. At the same time, we also believe that there is a growing body of evidence that suggests that infants are specially prepared for speech perception in ways that go beyond the simple discrimination of speech contrasts. Such special preparation is critical in allowing the infant listening to the steady stream of sounds in the environment to identify those that will permit him or her to communicate with other humans. For this purpose, the infant needs not only to distinguish human utterances from other sounds, but also to categorize correctly the sounds that belong to a particular language. In so doing, the infant has to cope with such facts as: different talkers vary in the way they pronounce the same thing, some talkers speak different languages, utterances are not unstructured wholes but contain elements that repeatedly appear in other utterances, etc. Given that infants somehow arrive at the beginnings of understanding and speaking language toward the end of the first year of life (for a recent review see Gleitman, Gleitman, Landau, and Wanner, in press), it is hard to see how this could occur unless the infant is prepared to extract critical cues that permit a correct classification of the speech input. We would argue that the way in which speech perception develops during infancy is indicative of what Gould and Marler (1987) have referred to as an "innately guided learning process". Before elaborating this view, let us briefly consider why it is so difficult to resolve the issue of specialization at the level of mechanisms underlying infant speech discrimination.

## INFANT SPEECH DISCRIMINATION STUDIES AND THEIR IMPLICATIONS FOR SPECIAL MECHANISMS

Early studies in the field focused on the infant's ability to discriminate different speech sound contrasts. For example, Eimas and his colleagues examined various types of phonetic distinctions such as voicing (Eimas *et al.*, 1971), place of articulation (Eimas, 1974), and manner of articulation (Eimas, 1975; Eimas and Miller, 1980b). Such studies helped to establish that the early speech perception capacities were not limited to a particular type of speech contrast, but generally to the types of contrasts that occur in many different languages. In addition, studies involving infants of different cultures (e.g., Lasky, Syrdal-Lasky, and Klein, 1975; Streeter, 1976) indicated that the initial perceptual categories are apparently quite similar across cultures.

Given that much of the research focused on capacities for distinguishing phonetic segments from each other, it was perhaps only natural that these capacities became the center of discussions about any innate specialization for perceiving speech sounds.

Consequently, when it was demonstrated that strong parallels exist in the way that infants perceive certain nonspeech sound contrasts (Jusczyk, Pisoni, Walley, and Murray, 1980; Jusczyk, Pisoni, Reed, Fernald, and Myers, 1983), doubts arose as to whether the underlying perceptual capacities for discriminating phonetic segments were truly specific to speech. These doubts were reinforced by the demonstration that nonhuman mammals, such as the chinchilla and the rhesus monkey, give some evidence of showing similar perceptual categories for phonetic segments (Kuhl and Miller, 1975, 1978; Kuhl and Padden, 1983; Waters and Wilson, 1976). Hence, some have argued that general auditory processing mechanisms could account for the kinds of discriminative abilities that infants exhibit in speech perception experiments (e.g., Jusczyk, 1981; Kuhl, 1978).[1]

Although the data from animals and the nonspeech studies with infants raise serious questions about the existence of a set of specialized speech perception capacities, they do not completely rule out such an account. The discovery that animals and humans arrive at the same categories for some types of information does not mandate that they use the same mechanisms for this purpose. In other words, there is more than one way to arrive at the same categorization of a set of stimulus materials (for further discussion of this point see Jusczyk, 1986b; Kuhl, 1986).

Similar difficulties arise in conjunction with the interpretation of the data from the infant perception studies. One problem is the lack of precision that exists in specifying the nature of the infant's categories. Because of the limited behavioral response repertoire of infants less than 4 months of age, the available procedures are often restricted to measures of discrimination and preference. Such measures provide some information about perceptual categories because they indicate which sorts of distinctions infants

[1] The notion that the underlying perceptual categories are specific to speech sounds has a number of interesting implications with respect to their form. Consider for a moment what the nature of these categories would have to be in order to account for what we know about the discrimination abilities of human infants. First, the categories themselves would have to be something other than phonemic ones in order to account for the fact that a newborn infant could, in principle, be learning any language and not just the predominant one spoken in his or her immediate environment. For this reason, the categories cannot be specific to a particular language. Any phonetic pre-category would have to have a broad range and be capable of being re-tuned to correspond to the precise category existent in what is going to be the infant's native language (e.g., see Aslin, 1981). It is not at all obvious that such retuning requires a speech-specific mechanism rather than a general auditory one. The latter account (e.g., see Jusczyk, 1985a, 1986c) holds that the initial precategories are not phonetic and that they only become phonetic as the child begins to develop a lexicon (for recognizing as well as producing, words). By this view, the finding that infants six months old and younger discriminate certain foreign language contrasts, but 10–12 months old do not (e.g., Werker and Tees, 1984; Werker and Lalonde, 1988) is attributable to a changeover to a phonetic mode of perception in the older group. The younger age group succeeds in discriminating the contrasts because they rely on general auditory processing mechanisms. In contrast, the speech-specific mechanism position (e.g., Eimas, 1982) views the developmental change as the result of losing phonetic categories that are not employed in the native language that the infant is in the process of acquiring.

can and cannot make. To the extent that two sounds are indistinguishable for infants, they must necessarily belong to the same perceptual category. However, perceptually distinguishable objects or events may also be treated as belonging to the same perceptual category. Thus, in order to accurately describe perceptual categories, one needs to determine not only which distinctions are made but also which similarities are perceived. Furthermore, an additional limitation of the procedures available for testing very young infants is that they are reliable only for group data and do not permit statements to be made about the capacities of an individual infant (Jusczyk, 1985b). Large stimulus contrasts must be used in testing groups of subjects to allow for the possibility of inter-subject variability in categorizing the stimuli. Hence, when examining the discrimination of voice onset time differences (e.g., Eimas *et al.,* 1971) or analogous timing differences for nonspeech sounds (e.g., Jusczyk *et al.,* 1980), the items in the stimulus pairs differ by 20 to 30 msec. This makes it difficult to obtain information about the precise loci of perceptual category boundaries. Thus, the fine details of whether there is a strict match between speech and nonspeech categories are often lacking.

Even when strong parallels are noted in the perception of speech and nonspeech stimuli, it could be argued that, rather than reflecting generalized auditory processing mechanisms, the speech perception mechanisms have simply been fooled into processing the nonspeech sounds. This suggestion is not without merit. Ethologists have observed that complex behaviors, so-called fixed action patterns, can be triggered by artificial "superstimuli" as well as by the typical sign stimuli that an organism is likely to encounter in its natural environment. The demonstration of triggering by a superstimulus does not undermine the view that the mechanisms responsible for producing the action pattern serve a specific behavioral function. Indeed, if Mattingly (1972) is correct in asserting that speech is a sign stimulus for humans, one could legitimately ask whether certain nonspeech sounds constitute superstimuli for speech mechanisms.

## INNATELY GUIDED LEARNING PROCESSES

Do the problems noted above mean that the question of whether there are innately given, specialized processing mechanisms associated with speech perception is unanswerable? Not necessarily. However, it does mean that it may be more fruitful to look for an answer beyond the capacities for distinguishing speech contrasts. This is also suggested by what we know about innately guided learning processes – the notion that many organisms are "preprogrammed" by information in their genetic makeup to learn particular things and to learn them in particular ways (Gould and Marler, 1987). An interesting aspect of such processes is the speed with which relatively complex patterns of behavior seem to unfold in the presence of the appropriate input occurring during a sensitive period of development. A well-known example is the "following behavior" exhibited by goslings shortly after hatching. Similarly, song learning in birds is held to be the outcome of an innately guided learning process (e.g., Marler and Peters, 1981). Certain species, such as the white-crowned sparrow, appear to be preprogrammed to learn their own species-typical song more readily than those of another species. Thus, song input from other species has little effect on the final song form of isolated white-crowned sparrows. Moreover, during their sensitive period, these isolates are capable of selecting and learning their own species-typical song from a tape-recorded medley of songs from various species.

Hence, certain kinds of signals may be especially salient for the developing organism in comparison to other signals. The organism appears to be specially prepared to respond to such signals. However, this preparation need not take the form of a special set of perceptual mechanisms at the level required for the discrimination of one type of signal from another. Instead, specialized processes in the organism may be manifested as biases associated with the type of information that is sought out from the environment, is most apt to receive further processing, has the highest priority for memory encoding, etc. Clearly, the signals that qualify for such special treatment must meet some particular physical description. Yet, within the narrow range of salient signals the same basic perceptual mechanisms may be employed as for signals that fall outside this range. To date, in the study of speech perception, the issue of what characteristics an acoustic signal must have for it to be treated as a speech signal is far from resolved. In general, mature listeners are able to extract coherent messages from speech signals under conditions that involve considerable distortions or masking by noise (Acton, 1970; Dirks and Wilson, 1969; Duquesnoy, 1983; Gelfand, Piper, and Silman, 1986; Liberman and Studdert-Kennedy, 1978; Plomp and Mimpen, 1979; for a review see Jusczyk, 1986a). For example, studies conducted by Remez and his colleagues (e.g., Remez, Rubin, Pisoni, and Carrell, 1981; Remez, Rubin, Nygaard, and Howell, 1987) indicate that even frequency-modulated sine waves lacking fundamental frequency information can be perceived as speech under certain circumstances.

We propose that an innately guided learning process exists for speech perception. This notion implies that the range of signals to which the system responds is restricted and that only limited experience with appropriate stimuli is necessary to fix the parameters of the perceptual system to provide for the most efficient means for recognizing utterances in the native language. Learning of the sound properties of the native language occurs rapidly because the system is structured to be sensitive to certain distributional properties of the input, and not to others. The innately guided learning position can be contrasted with an incremental learning hypothesis that postulates a very flexible system that responds to a wide range of inputs and gradually arrives at a perceptual classification that directly reflects the distributional properties of the input. In the latter system, there are no inherent constraints on which distributional properties can affect the resultant perceptual categories. In effect, the categories mirror the input.

What might it mean to consider the development of speech perception as the outcome of an innately guided learning process? Essentially, it would involve reorientating the discussion concerning specialization in speech perception. Instead of focusing so closely on the nature of the mechanisms underlying the infant's discrimination of speech sounds, more consideration would be given to the ensemble of processes that are engaged when the infant listens to speech. In other words, to what extent are speech sounds more likely to engage the infant's attentional processes than other signals? What type of information in the speech signal does the infant encode? Is there any indication that speech signals are more highly encoded, better remembered or more easily accessed by the infant than other acoustic signals? Is the range of input signals which can influence the eventual phonemic categorization of speech restricted in any principled ways or do the eventual categories simply reflect some central tendency of the distribution of signals encountered? In what ways are developmental changes in speech perception tied to other aspects of language acquisition?

In what follows, we will examine the available infant data for signs that the development of speech perception can be viewed as the result of an innately guided learning process. Hence, we shall look for evidence that speech may have a special status for infants *vis-a-vis* other kinds of acoustic signals. We will also endeavor to point out areas in which further information is required to help resolve this issue. Finally, we shall indicate what our view implies about the way in which speech perception develops.

## IMPLICATIONS FROM PREVIOUS STUDIES FOR WHETHER SPEECH HAS A SPECIAL STATUS FOR INFANTS

For the most part, the studies that have attempted to address the issue of whether speech is special for infants are ones which have focused on the ability to discriminate individual syllables (e.g., Eimas; 1974) or nonspeech analogs of syllables (e.g., Jusczyk *et al.,* 1980). However, many other aspects of early speech perception abilities have been studied over the years (e.g., see Aslin, Pisoni, and Jusczyk, 1983; Kuhl, 1987, for recent reviews). Although such studies did not usually focus on the possible special status of speech sounds for infants, many of these investigations do provide evidence that bear upon the issue. Because space limitations prohibit an exhaustive review of all the literature, we have selected a few areas that we believe are pertinent to the issue at hand.

### *Recognition of the mother's voice by young infants*

One characteristic of behaviors that seem to be products of innately guided learning is that they tend to have important adaptive significance for the organism. For example, bees are able to learn very rapidly which flower-like objects are likely to be the source of food (Meal and Erber, 1978). In other instances, innately guided learning appears to be involved in the way in which offspring come to recognize their parents (e.g., Tinbergen, 1969). In the realm of human behaviors, language plays a central role in communication. Not only is it crucial to acquiring information outside direct experience, but it is also one of the primary means of conveying one's needs and desires. It is only natural, then, that language behaviors may be part of the process by which human infants establish close contacts with their primary caretakers. In fact, in this domain there are signs that human infants are precocious in the facility with which they come to recognize their own mother's voice.

Mehler, Bertoncini, Barrière, and Jassik-Gerschenfeld (1978) investigated the ability of infants between two and four weeks of age to recognize their own mother's voice. They found that when the speech consisted of natural utterances (such as the kind that would normally be directed at an infant), the infants manifested a strong preference for the utterances produced by their own mothers (see also Mills and Meluish, 1974). On the other hand, when unnatural speech samples (produced by having the mothers read a text backwards) were used, the preference for the own mother's voice disappeared. The latter results suggest that the infants' preference is based on something other than the gross acoustic properties of the mother's voice (e.g., breathiness, characteristic voice pitch, spectral characteristics of syllables, etc.). Rather, the preference for the mother's voice occurred when the utterances were structured as communications that an infant might typically receive.

The age of the infants that Mehler *et al.* tested allows the infant considerable time for learning to recognize the mother's voice. However, there are indications that a two to four week period of familiarization is not necessary for recognition to occur. DeCasper and Fifer (1980) demonstrated that newborn infants can learn to adjust their sucking rates so as to listen to their own mother's voice in preference to the voice of another infant's mother. Hence, the preference for the mother's voice apparently occurs rapidly, with very little postnatal experience. In this sense, it is quite consistent with what we would expect of a product of an innately guided learning process.

*Discrimination of the native language from other languages*

Studies showing that the young infant can recognize his or her mother's voice demonstrate that at least some speech-related behaviors carry some special significance for the infant. They may be an indication that speech sounds are salient signals for infants. However, although such sounds may readily attract the infant's attention, it does not necessarily follow that speech signals undergo any sort of detailed structural analysis. Instead, the attention-getting function of the mother's speech might only be a means of engaging the infant in social interaction. Hence, further evidence is required in order to show that the infant focuses on the structural properties of the utterances directed at him or her. One indication that the infant does detect such structural features in the speech of the mother would be if the infant could distinguish utterances in the mother's native language from those in another language. Moreover, if the infant were able to do this even when the utterances were spoken by someone other than his or her mother, this would suggest that the infant has picked up some general information about the structural features characteristic of utterances in the native language.

A series of studies by Mehler, Jusczyk, Lambertz, Halsted, Bertoncini, and Amiel-Tison (1988) demonstrates that infants as young as four days old do have the capacity to distinguish utterances in their mother's native language from those in another language. To prevent infants from discriminating utterances on the basis of differences in the vocal characteristics of different talkers, Mehler *et al.* used utterances produced by a fluent bilingual speaker of the two languages. However, they also wanted to ensure that infants were responding on the basis of some general properties of utterances belonging to a particular language and not simply because they discriminated an isolated utterance in one language from one in another language. For this reason, they presented infants with a series of different utterances from one language and then followed these with either more utterances from the same language (control condition) or new utterances from a different language (experimental condition). The results indicated that newborn French infants discriminated French from Russian utterances. Moreover, they appeared to prefer the French utterances. The language discrimination apparently depends upon some familiarity with at least one of the languages, even if this familiarity comes only in the first few days of life. Thus, infants born in France who came from non-French-speaking homes did not give evidence of distinguishing French from Russian utterances. Nor did French infants display any tendency to distinguish utterances in one foreign language, English, from those in another, Italian.

To verify that the pattern of results obtained with French newborns was attributable to their familiarity with the native language and not to the fact that one language pair was simply less discriminable than the other, Mehler *et al.* also tested two-month-old American infants on the contrasts. They found the complementary pattern of results for this group, namely the American infants discriminated the English and Italian utterances, but not the French and Russian ones. Thus, the overall pattern of results argues that it is familiarity with one of the two languages that accounts for the discrimination results. Evidently, the infant learns very rapidly to detect the characteristic features of utterances in the native language. The results of several additional experiments conducted by Mehler *et al.* suggest that the critical information that infants are extracting may lie in the prosodic characteristics of the utterances. Thus, when the speech samples were low-pass filtered to eliminate all distinctive speech information except for the prosody, infants still distinguished the utterances in the native language from those in the foreign language.

Thus, the language discrimination studies provide some indication that speech stimuli do more than simply capture the infant's attention. Rather, infants are picking up information about important structural features of utterances in the native language. Once again, it is striking that so little postnatal exposure is necessary for this to occur. The characteristic rhythm and prosody of the language are apparently very salient features for infants.

### *Bimodal speech perception*

One indication that an organism may be predisposed to give special treatment to certain environmental signals would be if it had a greater facility for integrating information from different modalities for these types of signals as opposed to others. Such a predisposition would imply that the underlying perceptual systems were structured in some way that makes the integration of the different sources of information more likely.

At first glance, speech perception might not seem to be an area wherein cross-modal matching would have a high priority. It is certainly the case that speech can be perceived in the absence of a clear view of the talker's face. Thus, blind individuals are able to perceive speech, and sighted listeners are capable of perceiving speech over telephone lines. This is not to say that visual input may not be useful or even important under conditions that make hearing difficult. Indeed, we know that trained observers (Campbell and Dodd, 1979), and even some untrained ones (Summerfield, 1988), have some success in "lip-reading". Moreover, in recent years a number of studies with adults have shown that conflicting visual cues may even override auditory cues in phonetic perception (Massaro and Cohen, 1983; McGurk and MacDonald, 1976; Summerfield, 1979). Of course, with adults, it is always possible that the integration of visual and auditory information comes about through a long apprenticeship in viewing other people speaking.

However, there are indications that some type of visual-auditory integration of speech information may be occurring in infants four-to-six months of age. Specifically, Kuhl and Meltzoff (1984) presented infants with two side by side videos of a woman's face pronouncing the vowels [i] and [a]. The facial motions were synchronized with each other and with an auditory stimulus which consisted of either of the two vowels. Kuhl and Meltzoff found that the infants looked longer at the face articulating the vowel identical to the one that they were hearing than they did at the face articulating the other vowel. In a critical control experiment, Kuhl and Meltzoff replaced the speech sounds by pure tone stimuli matched to each syllable in terms of amplitude and durational characteristics. The infants did not display any significant preferences under these conditions, suggesting that the link may be one between mouth shape and a resulting speech sound. MacKain, Studdert-Kennedy, Spieker, and Stern (1983) report a similar finding for a match between the visual and auditory information distinguishing the bisyllables [zuzi] and [vava]. Interestingly, though, in the latter study significant audio-visual matches occurred only when infants were looking to their right sides – a finding that MacKain *et al.* suggest may be related to lateralization for language in the left hemisphere.

Therefore, there are indications that by four-to-six months of age infants show a tendency to integrate auditory and visual sources of information for some speech contrasts. Just how extensive this ability is, is unknown at present. Nevertheless, the existing data are certainly consistent with the view that the infant's perceptual system may be set up to give speech sounds especially detailed processing.

### *Child-directed speech as an attractive signal*

If an organism is innately predisposed for processing certain kinds of signals, then one might expect that such signals would have a higher priority for engaging attention. There are some indications that speech, particularly child-directed speech ("Motherese"), may be salient in this way for infants at some point during the first few months of life. So, four-month olds, but not two-month olds, display a preference for listening to a female voice as opposed to silence or white noise (Colombo and Bundy 1981). Still, it could be argued that the alternatives to speech in this case (i.e., silence or white noise) are rather uninteresting and patternless. Studies involving more complex nonspeech alternatives have been undertaken with older infants. For example, Friedlander and Wisdom (1971) reported that infants between nine and 18 months of age manifest significant preferences for nursery rhymes over electronic noise. In another study, Glenn, Cunningham, and Joyce (1981) took precautions to match a speech stimulus to a nonspeech stimulus on dimensions such as pitch, rhythm, and amplitude. They found that nine-month-old infants preferred to listen to an unaccompanied female voice singing a nursery rhyme as opposed to a solo musical instrument playing the same piece.

Thus, by nine months of age there is some suggestion that speech sounds may be more apt to hold the infant's attention than some other kinds of complex acoustic signals. Whether this preference is present from birth, or arises only when the infant is at the point of learning his or her first words, cannot be determined until younger infants are tested in such a situation. There are some indications within the realm of speech sounds themselves that some types of speech are more likely to capture the infant's attention. In particular, Fernald (1985) found that when given a choice between Motherese and adult-directed speech, four-month-old infants display a significant tendency to select Motherese. Subsequent research (Fernald and Kuhl, 1987) suggests that the fundamental frequency ($F_0$) characteristics of Motherese may account for the preferences that the four-month-olds show. Specifically, Fernald and Kuhl conducted a series of experiments with Motherese and adult-directed speech in which $F_0$, amplitude, and rhythmic characteristics were varied independently. They found that the samples derived from Motherese were preferred to those from adult-directed speech only if they maintained the distinctive pitch characteristics of Motherese (e.g., higher mean $F_0$, wider pitch excursions, and an expanded $F_0$ range).[2]

*Perceptual segmentation of fluent speech*

The finding that Motherese may be a potent attractor of the infant's attention again raises the question of whether this leads to a more detailed processing. As we noted earlier in another context, just because speech sounds turn out to be good elicitors of the infant's attention does not entail that their structural properties will be analyzed. Indeed, Fernald and Kuhl (1987) have suggested that part of the appeal of Motherese may lie in its affective expressive power, particularly in communicating positive affect towards the child. Were it the case that attention to Motherese is relevant only to establishing social relationships, then such data would not bear on the question of whether speech perception is the result of an innately guided learning process. However, there are indications that, aside from whatever role it may play in the establishment of social relationships, the infant's attention to Motherese may be an important step toward language acquisition. Indeed, this may be an area in which speech perception capacities interact with mechanisms involved in other aspects of language acquisition.

In addition to its distinctive $F_0$ characteristics, it has been noted that, relative to adult-directed speech, Motherese tends to be slower, more carefully enunciated and contain more pauses (Broen, 1972; Garnica, 1977; Morgan, 1986; Snow, 1972). Thus, some observers have been led to speculate whether these prosodic properties of Motherese may provide the language learner with acoustic markers of important structural units in

2 As an anonymous reviewer noted, the findings for Motherese raise an interesting question as to why infants may prefer it when it is so different from natural adult-directed speech. We agree with this reviewer's comment that Motherese may be a "superstimulus" for the infants and that it serves the function of marking a number of properties of the input that must be learned in order to acquire the native language. Although eventually it is the adult-directed patterns that the child must master, the exaggerated marking in Motherese may serve to call attention to crucial portions of the input. A case in point is discussed in the section on perceptual segmentation of fluent speech.

language such as clauses and phrases (e.g., Hirsh-Pasek, Kemler Nelson, Jusczyk, Wright Cassidy, Druss, and Kennedy, 1987; Kemler Nelson, Hirsh-Pasek, Jusczyk, and Wright Cassidy, in press; Morgan, 1986; Morgan, Meier, and Newport, 1987). In English, changes in $F_0$, pauses, and segmental lengthening have been shown to be associated with clause and phrase boundaries (e.g., Cooper and Paccia-Cooper, 1980; Garnica, 1977; Klatt, 1976; Martin, 1970; Nakatani and Dukes, 1977). Thus, the suggestion is that the prosody of Motherese may exaggerate such features and render them more detectable for the language learner. In this way, infants may perceptually segment speech into units that correspond, more or less, to important grammatical units.

Hirsh-Pasek *et al.* (1987) sought to determine whether infants treat clauses as though they are perceptual units. For this purpose, they presented six- and nine-month-olds with pairs of English utterances that were segmented artificially by the inclusion of 1 second pauses. For each utterance pair, the pauses in one utterance were inserted in locations corresponding to clause boundaries; for the other member of the pair, the pauses occurred in the middle of clauses. The rationale for the pause placement was the following: To the extent that there are prosodic markers of clauses and infants detect these, then pauses occurring at the clause boundaries will be perceived as less disruptive than those that interrupt the clauses. Hirsh-Pasek *et al.* found that infants displayed a significant preference for listening to those items in which the pauses occurred at the clause boundaries. Hence, these results suggest that, by six months of age, infants are sensitive to the prosodic markers of clausal units. Moreover, in a subsequent investigation, Kemler Nelson *et al.* (in press) examined whether adult-directed speech was equally effective as Motherese in providing infants with cues to clausal units. Their results indicated that this is not the case. Only with Motherese was there a significant preference for the pair members with the pauses inserted at the clause boundaries.

More recently, Jusczyk, Hirsh-Pasek, Kemler Nelson, Kennedy, Woodward, and Piwoz (submitted) have explored a comparable situation involving certain phrasal units in English. Specifically, they investigated whether infants detect cues to major phrasal units such as Subject Noun and Predicate Verb phrases. Once again infants gave evidence of detecting the prosodic markers to the phrasal units. In addition, there was some indication that the markers were more prominent in Motherese than in adult-directed speech. However, in contrast to the investigations involving clausal units, sensitivity to the prosodic markers of phrasal units was observed for nine-month-olds, but not for six-month-olds. Hence, the developmental trend seems to be one of differentiation (i.e., both age groups gave evidence of detecting cues to clausal units, whereas only the nine-month-olds showed sensitivity to the markers of subclausal units like phrases).

Taken together, the results of these studies imply that, at least by six months of age, infants are sensitive to some of the structural features of speech conveyed in Motherese. In particular, they appear to be sensitive to the packaging of information into units such as clauses. This is an indication that they are treating such signals as more than just an invitation to social engagement. Thus, there is evidence that infants are analyzing the input at least to the degree needed to segment it perceptually into clausal units. Moreover, this type of perceptual segmentation may be a necessary step towards the way in which the infant comes to learn the structural properties that hold among utterences in the native language. Because to begin to determine the regularities that hold within and among the units in a particular language, the learner must first be able to identify the relevant units themselves. Thus, it is in this sense that sensitivity to the perceptual organization of the speech signal may play a role in triggering a complex pattern of behaviors that ultimately yield a structural description of utterances. Such an outcome fits well with the innately guided learning view and is clearly consistent with what we know about how complex behavior patterns are triggered in other species.

*Perceptual representations of speech sounds*

The position that we have taken here is one which essentially places specialized processing of speech sounds at a point beyond the infant's discriminative capacities: Speech sounds are special by virtue of their capacity to capture attention and to engage further processing. Another important issue concerns the way in which speech information is retained relative to other kinds of acoustic signals. Thus, special treatment for speech sounds may not show up so much in the processes associated with perceiving the signal, but in the degree to which such information is retained. One of the great feats associated with language acquisition is the rapid spurt of growth in the development of a vocabulary. New words seem to be acquired on the basis of very little exposure and many words get added over a relatively short period of time. It is not obvious that there is anything comparable to this in other domains of knowledge. Is it possible that the origins of such behavior could be traced back to an initial predisposition for encoding speech sounds? For example, one way in which speech sounds might receive more detailed processing relative to other types of acoustic signals could be in the kinds of detail that are retained in perceptual representations of speech. Accordingly, the infant may be better at encoding the relations between various components of the speech signal than at encoding comparable aspects of nonspeech signals.

Obtaining information about the nature of the infant's perceptual representations is difficult given the limited methods available. However, in recent years some progress has been made in this domain. For example, Miller and Eimas (1979; Eimas and Miller, 1981) have conducted a series of studies exploring two- to three-month-olds' capacity to detect recombinations of various speech features. Infants tested with the high amplitude sucking procedure were presented with an alternating series of two syllables like [ba] and [dæ]. Following habituation to these items, the infants were shifted to a new pair that contained the same consonants and vowels but which were combined differently as [bæ] and [da]. The infants successfully detected this kind of recombination of the phones. Similarly, in subsequent experiments Miller and Eimas (1979) reported that infants in this age range were sensitive to reorderings of individual phonetic features and to recombinations involving segmental and suprasegmental cues (Eimas and Miller, 1981). Since in each instance the same features were always present, the only way to discriminate the sounds was by attending to the relations among the features present. Hence, these results show that infants are registering more about speech sounds than simply a loose agglomeration of acoustic features. Rather, these features are related together in a way that appears to preserve the inherent structure of the syllables.

In a slightly different vein, Jusczyk and his colleagues (e.g., Bertoncini, Bijeljac-Babic, Jusczyk, Kennedy, and Mehler, 1988; Jusczyk and Derrah, 1987; Jusczyk, Kennedy, and Jusczyk, in preparation) have investigated the kind of information that is present in young infant's representations of syllables. They modified the high amplitude sucking procedure in such a way that required infants to use a perceptual representation in order to detect the presence of a new item. Specifically, the infants were presented with a randomized series of different syllables that shared some common element (e.g., [bi], [ba], [bo], [bɚ]). Because the syllables in this set all differed from each other, it was not possible for the infants to detect the presence of a new item simply by comparing two successive items. Rather, in order to succeed in the task the infant must encode information about each syllable that serves to distinguish it from the other items present in the original set. After the infants had been habituated to the original stimulus set, a new syllable was introduced. The item either shared the common element with the other members of the set (e.g., [bu] for the set described above) or it did not (e.g., [du]). Jusczyk and Derrah (1987) reasoned that if two-month-old infants encoded the syllables as a sequence consisting of a particular consonant followed by a particular vowel, then they might recognize that the [bu] stimulus shared a common segment with the other stimuli. In this case, the infants might display greater sucking rates for the more novel change – the one that did not include a common element. The infants did not respond in this way. Instead, the infants' representations preserved the detail necessary to detect both types of changes. In a subsequent study, Bertoncini *et al.* (1988) found evidence that the amount of detail present in the infant's representation of syllables appears to increase during the first two months of life. Thus, new born infants' representations did not appear to be sufficiently detailed to differentiate between syllables when only consonantal information was changed. In contrast, with changes involving vowel differences, they performed as well as two-month-olds.

More recently, Jusczyk *et al.* (in preparation) began exploring the extent to which infants are able to retain details about their perceptual representations for syllables. They introduced delay periods filled with distracting visual materials between the presentation of the test sequence and the original stimulus set. They found that, under these conditions, two-month-old infants still retained sufficient information to permit them to detect very subtle changes in the composition of the test set. In fact, there was no indication of any measurable decrease in performance with the 2-minute delay period.

So, beyond their ability to make fine discriminations between speech sounds, there is some indication that during the first few months of life infants are able to encode rather detailed information about speech sounds into perceptual representations. In addition, it appears that they do retain important details about speech sounds for at least brief intervals thereafter. This type of behavior is certainly consistent with the view that speech is not only highly salient for infants, but also that it is a good candidate for further processing.

## WHERE THE GAPS LIE

The areas of research reviewed above provide evidence that speech is a salient signal for infants. In addition, there are suggestions from the language discrimination research, the bimodal perception experiments, and the segmentation and representation studies, that considerable processing of the speech signal does occur. In this sense, the data are certainly consistent with the view that speech perception may be the result of an innately guided learning process. However, as noted earlier, most of the studies that we have considered here were not designed explicitly to investigate whether speech has a special status in this regard relative to other types of auditory signals. Consequently, relevant information concerning similarities and differences in the processing of speech and nonspeech signals in similar situations is often lacking. In addition, other relevant data have not been collected yet. Let us begin with a consideration of where some of the gaps lie in the previous investigations.

*Recognition of the mother's voice*

The primary question here has to do with whether recognition of the mother's voice has to do with speech perception *per se.* At present, studies in this area have always used speech stimuli. This is natural since most vocalizing in humans does involve speech. However, one can ask whether recognition of the maternal voice can occur outside of a typical speech context. The Mehler *et al.* (1978) study, wherein infants failed to show recognition of the mother's voice under conditions in which prosody was disrupted, seems to suggest that a context wherein coherent sentences are presented may be necessary. Nevertheless, it could be argued that the backwards text reading situation they used was highly unnatural and unlikely to fully engage the infant's attention. The kind of comparison that is called for is one that involves a more natural set of nonspeech noises such as humming or perhaps even the types of nonspeech vocalizations used in playing with infants.

*Preference for native language utterances*

Mehler *et al.* (1988) found that newborn infants suck at higher rates when listening to utterances in the native language as opposed to ones in a foreign language. It cannot be denied that they have picked up at least something that characterizes utterances in the native language. However, a question can be posed about the preference. Does it mean that infants prefer to listen to speech sounds over other types of sounds occurring in the natural environment? As we noted in the section on Motherese, there have been some attempts to examine preferences for speech relative to other kinds of interesting acoustic signals (e.g., Friedlander and Wisdom, 1971; Glenn *et al.,* 1981). However, these studies were done with much older infants. What is necessary is a comparison involving newborn infants' preferences for speech and environmental sounds.

*Bimodal speech perception*

The demonstrations that four- to six-month-old infants match visual and auditory input relating to the production of speech sounds are impressive. Moreover, given the history of theorizing concerning the underlying mechanisms for speech perception (e.g., Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967; Ojemann, 1983), the results carry added importance because they suggest the possibility of an early link between production and perception. However, it should be noted that there is considerable evidence in infants at this age for visual-auditory bimodal perception of nonspeech events as well (e.g., Bahrick, 1983; Spelke, 1979). Hence, the demonstration that speech events are perceived bimodally is insufficient in itself to establish claims about special treatment that speech sounds receive. One direction in which to pursue this problem further would be to determine whether infants are somehow more sensitive to cross-modal matches between visual and auditory input for speech than they are for other types of events. For example, bimodal studies in infants have explored the role of temporal synchrony in establishing correspondences between visual and auditory events (Allen, Walker, Symonds, and Marcell, 1977; Bahrick, 1987; Dodd, 1979; Spelke, Born, and Chu, 1983). To date, close comparisons of infants' sensitivity to temporal synchrony for speech versus nonspeech events have not been undertaken. If infants were to show more sensitivity to temporal synchrony between visual and auditory input for speech than they do for nonspeech (e.g., they are better able to detect temporal mismatches), then this might be an indication that a separate system computes the correspondences for speech.

*Perceptual segmentation*

The research on perceptual segmentation of speech (Hirsh-Pasek *et al.,* 1987; Kemler Nelson *et al.,* in press; Jusczyk *et al.,* submitted) demonstrates that the speech signal is not an undifferentiated whole for infants. Rather, there are indications that the infant is sensitive to the presence of prosodic features that serve to mark utterances as structured into different segments. Nevertheless, the fact that infants perceive the speech signal as structured into different segments is not an indication that speech is special in this regard. Comparable studies have yet to be done with complex nonspeech signals. It may be that infants are also sensitive to indices of segments when listening to highly patterned acoustic signals such as music. If so, then a further question would be whether the same markers of perceptual segments occur across all types of acoustic signals, speech or nonspeech. If it turned out that the acoustic markers that the infant attends to when listening to speech are different from those in listening to other kinds of nonspeech signals, then this would suggest that speech sounds undergo different processing.

*Perceptual representations*

In this domain, like many of the other areas already discussed, it is necessary to extend the paradigms employed for studying speech to other types of acoustic signals. Thus, one would like to know the extent to which the infant is able to encode detailed information about various types of nonspeech signals under conditions similar to those that have been studied for speech. One starting point may be to substitute nonspeech analogs for the various types of speech sounds used in studies like those performed by Jusczyk and his colleagues (Bertoncini *et al.,* 1988; Jusczyk and Derrah, 1987; Jusczyk *et al.,* in preparation). In this way, one might be able to ensure that the range of variation in the signals is similar to that employed in the speech studies. However, a case can also be made for studying other, less artificial, types of sounds that the infant is likely to encounter in the natural environment (e.g., music, animal noises, etc.). Such sounds may be more apt to engage the infant's attention which may, in turn, lead to better encoding. A disadvantage of using such signals is that it may be difficult to ensure that the range of variability among the various sounds is comparable to that of the speech sounds. Assuming that these problems could be surmounted, one would like to know whether speech sounds are privileged relative to other sounds in either the detail of the encoding or the facility with which this type of information is retained.

*Other possible indications of special treatment for speech*

As noted earlier, aside from the existence of any special capacities for perceiving speech contrasts, there are a number of ways in which speech sounds may receive special processing relative to other kinds of acoustic signals. One possibility is that speech processing may be more robust than nonspeech processing. Hence, speech signals may be more easily detected in noisy environments, they may be better able to withstand distortion of the signal, and they may be more easy to follow as background noise levels increase. At present, there simply are no data available with infants concerning these matters.

Another way in which speech may receive special treatment is by undergoing more extensive processing than other kinds of signals. This may happen simply because the salience of the sounds makes them more likely to be processed longer (the implication being that other salient signals might undergo a similar degree of processing). Alternatively, to return to an old argument in the field, there may be some special perceptual mechanisms that are engaged in processing speech sounds. The mechanisms in this case might be ones that deal with problems that are apparently uniquely associated with perceiving speech. One speculation is that this might include the normalization mechanisms that help to cope with the variability introduced in the speech signal by changes in talkers, speaking rate, intonational patterns, etc. Thus, we know from the work of Kuhl (1979, 1983) that by six months of age infants are able to ignore variability in the speech signal introduced by changes in talker and intonation. Similarly, there is evidence that infants of from two to three months of age give evidence of compensating for variability associated with changes in speaking rate (Eimas and Miller, 1980a; Miller and Eimas, 1983). However, as was the case for discussions surrounding the existence of special mechanisms for categorical perception of speech sounds by infants (see Jusczyk, 1981, for a review), the data concerning special normalization mechanisms for speech sounds are equivocal. Thus, there is evidence that certain nonhuman species display some ability to compensate for talker differences (e.g., Baru, 1975). In addition, Trehub (1987; Trehub, Thorpe, and Morrongiello, 1987) has demonstrated that infants are capable of extracting some property common to transposed musical melodies. Moreover, Jusczyk *et al.* (1983) found that infants listening to nonspeech contrasts made similar kinds of adjustments for rate changes as those that Eimas and Miller (1980a) had observed for speech sounds. Thus, the fact that parallel sorts of compensatory mechanisms exist in other domains of auditory perception raises questions concerning the specialness of those underlying normalization in speech. In order to demonstrate the existence of any such special mechanisms, it will be necessary to show that the ones operating on speech sounds behave in a significantly different way than those operating in other domains.

## CONCLUDING REMARKS

As we have argued, we believe that the issue of whether the development of speech perception is an outcome of an innately guided learning process extends considerably beyond the existence of a set of specialized perceptual mechanisms. Rather, the notion of an innately guided learning process suggests that the infant is primed in certain ways to seek out some types of signals as opposed to others. Such salient signals are apt to have a high priority for further processing and memory storage. In short, they have an inherent importance for the organism. Moreover, the processing that such salient signals receive leads the infant to focus on the kind of information relevant to certain aspects of language acquisition. Thus, the newborn infant seems ready to keep utterances in the native language separate from those in a foreign language (Mehler *et al.,* 1988). Similarly, at least by six months of age, infants are sensitive to acoustic correlates of clausal units (Hirsh-Pasek *et al.,* 1987). The latter ability is critical for language acquisition because determining the grammatical relations that hold among words requires grouping them according to what clause they appear in.

Finally, let us consider what developmental course might be taken by an innately guided learning process for speech perception. Clearly, the starting point for such a process must be a state which is sufficiently general to allow the infant to learn the sound structure of any natural language. The endpoint of the process is to arrive at a state in which the individual is most efficiently prepared to segment and recognize utterances in the native language. This suggests that speech perception capacities should move toward states that take advantage of the structural regularities that hold among the phonetic segments used by speakers of the native language. In other words, all else being equal, the goal is to arrive at a set of speech perception capacities that are tuned to the rules that govern the production of sound sequences in the native language (these rules would include not only which types of segments are permissible, but also how they may be combined). Thus, the development of speech perception capacities must be from general to specific and an innately guided learning process must provide the means for this. One way for this to occur is by what is sometimes called "parameter setting" (e.g., Chomsky, 1981; Hyams, 1986). The infant is innately prewired with broad categories that may develop in one of several different directions. Certain kinds of input from the environment lead the organism to set the parameters of these categories to develop in one of these directions.

A tendency for increasing experience with a native language to lead to greater specificity in speech perception capacities has been noted at a number of different levels. First, representations of the information available in speech sounds appear to become more detailed over the first few months of life (e.g., Bertoncini *et al.,* 1988). Second, there is evidence that the perception of speech contrasts during the first year of life moves from a general ability to distinguish contrasts that might appear in any language towards one increasingly focused on those contrasts that actually appear in the native language. Thus, studies with young infants demonstrate that they are capable of distinguishing contrasts that do not occur in the native language environment (e.g., Streeter, 1976; Trehub, 1976). However, by 10 to 12 months of age, infants show a diminished sensitivity to certain foreign language contrasts (Werker and Tees, 1984; Werker and Lalonde, 1988; but see Best, McRoberts, and Sithole, 1988, for an instance in which sensitivity is maintained). Third, at a level of more global units, the infant appears to become increasingly sensitive to acoustic correlates of the structural properties of the native language. In fact, the developmental pattern here is very reminiscent of what has been observed at an older age for phonetic contrasts. At four and a half months of age, American infants detect acoustic correlates of clausal units in the native language, English, as well as in a foreign language, Polish (Jusczyk, Hirsh-Pasek, Kemler Nelson, and Schomberg, in preparation). However, by six months of age, only the correlates for the English contrasts are detected by the American infants. Accompanying this drop in sensitivity to units in a foreign language, there appears to be an increased sensitivity to correlates of important units in the native language. Thus, the Jusczyk *et al.* (submitted) study found that American infants between six and nine months of age develop sensitivity to subclausal units in English, such as phrases. In short, the infants are homing in on those particular features that they need to attend to so as to master the language spoken in their immediate environment. The fact that they discover these features so early on, given all those potentially available in the input, suggests that an innately guided learning process is at work.

## REFERENCES


ACTON, B.W. (1970). Speech intelligibility in a background noise and noise-induced hearing loss. *Ergonomics*, **13**, 546–554.

ALLEN, T.W., WALKER, K., SYMONDS, L., and MARCELL, M. (1977). Intrasensory and intersensory perception of temporal sequences during infancy. *Developmental Psychology*, **13**, 225–229.

ASLIN, R.N. (1981). Experiential influences and sensitive periods in perceptual development: A unified model. In R.N. Aslin, J.R. Alberts, and M.R. Petersen (eds.), *Development of Perception: Psychobiological Perspectives (Vol. 2). The Visual System* (pp. 219–255). New York: Academic Press.

ASLIN, R.N. PISONI, D.B., and JUSCZYK, P.W. (1983). Auditory development and speech perception in infancy. In M. Haith and J. Campos (eds.), *Handbook of Child Psychology, Vol. 2. Infancy and Developmental Psychobiology* (pp. 573–689). New York: Wiley.

BAHRICK, L.E. (1983). Infants' perception of substance and temporal synchrony in multimodal events. *Infant Behavior and Development*, **6**, 429–451.

BAHRICK, L.E. (1987). Infants' intermodal perception of two levels of temporal structure in natural events. *Infant Behavior and Development*, **10**, 387–416.

BARU, A.V. (1975). Discrimination of synthesized vowels [a] and [i] with varying parameters in dog. In G. Fant and M.A.A. Tatham (eds.), *Auditory Analysis and the Perception of Speech* (pp. 91–101). London: Academic Press.

BERTONCINI, J., BIJELJAC-BABIC, R., JUSCZYK, P.W., KENNEDY, L.J., and MEHLER, J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General*, **117**, 21–33.

BEST, C.T., MCROBERTS, G.W., and SITHOLE, N.M. (1988). Examination of the perceptual reorganization for speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 345–360.

BROEN, P. (1972). *The Verbal Environment of the Language Learning Child.* ASHA Monograph, 17.

CAMPBELL, R., and DODD, B. (1979). Hearing by eye. *Quarterly Journal of Experimental Psychology*, **32**, 85–99.

CHOMSKY, N. (1981). *Lectures on Government and Binding.* Dordrecht: Foris.

COLOMBO, J., and BUNDY, R.S. (1981). A method for the measurement of infant auditory selectivity. *Infant Behavior and Development*, **4**, 219–223.

COOPER, W.E., and PACCIA-COOPER, J. (1980). *Syntax and Speech.* Cambridge, MA: Harvard University Press.

DECASPER, A.J., and FIFER, W.P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, **208**, 1174–1176.

DIRKS, D.D., and WILSON, R.H. (1969). The effect of spatially separated sound sources on speech intelligibility. *Journal of Speech and Hearing Research*, **12**, 5–38.

DODD, B. (1979). Lip reading in infants: Attention to speech presented in- and out-of-synchrony. *Cognitive Psychology*, **11**, 478–484.

DUQUESNOY, A.J. (1983). The intelligibility of sentences in quiet and in noise in aged listeners. *Journal of the Acoustical Society of America*, **74**, 1136–1144.

EIMAS, P.D. (1974). Auditory and linguistic processing of cues for place of articulation by infants. *Perception and Psychophysics*, **16**, 513–521.

EIMAS, P.D. (1975). Auditory and phonetic coding of the cues for speech: Discrimination of the [r–l] distinction by young infants. *Perception and Psychophysics*, **18**, 341–347.

EIMAS, P.D. (1982). Speech perception: A view of the initial state and perceptual mechanisms. In J. Mehler, E.C.T. Walker, and M. Garrett (eds.), *Perspectives on Mental Representation* (pp. 339–360). Hillsdale, NJ: Erlbaum.

EIMAS, P.D., and MILLER, J.L. (1980a). Contextual effects in infant speech perception. *Science*, **209**, 1140–1141.

EIMAS, P.D., and MILLER, J.L. (1980b). Discrimination of the information for manner of articulation by young infants. *Infant Behavior and Development*, **3**, 367–375.

EIMAS, P.D., and MILLER, J.L. (1981). Organization in the perception of segmental and suprasegmental information by infants. *Infant Behavior and Development*, **4**, 395–399.

EIMAS, P.D., SIQUELAND, E.R., JUSCZYK, P.W., and VIGORITO, J. (1971). Speech perception in early infancy. *Science*, **171**, 304–306.

FERNALD, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, **8**, 181–195.

FERNALD, A., and KUHL, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, **10**, 279–283.

FRIEDLANDER, B.Z., and WISDOM, S.S. (1971). *Preverbal Infants' Selective Operant Responses for Different Levels of Auditory Complexity and Language Redundancy.* Paper presented at Annual General Meeting of the Eastern Psychological Association, New York.

GARNICA, O.K. (1977). Some prosodic and paralinguistic features of speech to young children. In C. Snow and C.A. Ferguson (eds.), *Talking to Children: Language Input and Acquisition* (pp. 63–88). Cambridge: Cambridge University Press.

GELFAND, S.A., PIPER, N., and SILMAN, S. (1986). Consonant recognition in quiet and in noise with aging among normal hearing listeners. *Journal of the Acoustical Society of America*, **80**, 1589–1598.

GLEITMAN, L.R., GLEITMAN, H., LANDAU, B., and WANNER, E. (in press). Where the learning begins: Initial representations for language learning. In F. Newmeyer (ed.), *The Cambridge Linguistic Survey.* Cambridge: Cambridge University Press.

GLENN, S.M., CUNNINGHAM, C.C., and JOYCE, P.F. (1981). A study of auditory preferences in nonhandicapped infants and infants with Down's Syndrome. *Child Development*, **52**, 1303–1307.

GOULD, J.L., and MARLER, P. (1987). Learning by instinct. *Scientific American*, **256**, 62–73.

HIRSH-PASEK, K., KEMLER NELSON, D.G., JUSCZYK, P.W., WRIGHT CASSIDY, K., DRUSS, B., and KENNEDY, L. (1987). Clauses are perceptual units for infants. *Cognition*, **26**, 269–286.

HYAMS, N.M. (1986). *Language Acquisition and the Theory of Parameters*. Dordrecht: Reidel.

JUSCZYK, P.W. (1981). The processing of speech and nonspeech sounds by infants: Some implications. In R.N. Aslin, J.R. Alberts, and M.R. Petersen (eds.), *Development of Perception: Psychobiological Perspectives (Vol. 2). The Visual System* (pp. 191–217). New York: Academic Press.

JUSCZYK, P.W. (1985a). On characterizing the development of speech perception. In J. Mehler and R. Fox (eds.), *Neonate Cognition: Beyond the Blooming, Buzzing Confusion* (pp. 199–229). Hillsdale, NJ: Erlbaum.

JUSCZYK, P.W. (1985b). The high amplitude sucking procedure as a methodological tool in speech perception research. In G. Gottlieb and N.A. Krasnegor (eds.), *Measurement of Audition and Vision in the First Year of Postnatal Life: A Methodological Overview* (pp. 195–222). Norwood, NJ: Ablex.

JUSCZYK, P.W. (1986a). A review of speech perception research. In K. Boff, L. Kaufman, and J. Thomas (eds.), *Handbook of Perception and Human Performance* (pp. 27:1–57). New York: Wiley.

JUSCZYK, P.W. (1986b). Some further reflections on how speech perception develops. In J. Perkell and D.H. Klatt (eds.), *Invariance and Variability in Speech Processes* (pp. 33–35). Hillsdale, NJ: Erlbaum.

JUSCZYK, P.W. (1986c). Towards a model for the development of speech perception. In J. Perkell and D.H. Klatt (eds.), *Invariance and Variability in Speech Processes* (pp. 1–19). Hillsdale, NJ: Erlbaum.

JUSCZYK, P.W., and DERRAH, C. (1987). Representation of speech sounds by young infants. *Developmental Psychology*, **23**, 648–654.

JUSCZYK, P.W., HIRSH-PASEK, K., KEMLER NELSON, D.G., KENNEDY, L.J., WOODWARD, A., and PIWOZ, J. (submitted). Perception of acoustic correlates to major phrasal units by young infants.

JUSCZYK, P.W., HIRSH-PASEK, K., KEMLER NELSON, D.G., and SCHOMBERG, T. (in preparation). Perception of acoustic correlates to clausal units in a foreign language by American infants.

JUSCZYK, P.W., KENNEDY, L.J., and JUSCZYK, A.M. (in preparation). Young infants' memory for information in speech syllables.

JUSCZYK, P.W., PISONI, D.B., REED, M., FERNALD, A., and MYERS, M. (1983). Infants' discrimination of the duration of a rapid spectrum change in nonspeech signals. *Science*, **222**, 175–177.

JUSCZYK, P.W., PISONI, D.B., WALLEY, A., and MURRAY, J. (1980). Discrimination of the relative onset of two-component tones by infants. *Journal of the Acoustical Society of America*, **67**, 262–270.

KEMLER NELSON, D.G., HIRSH-PASEK, K., JUSCZYK, P.W., and WRIGHT CASSIDY, K. (in press). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*.

KLATT, D.H. (1976). Linguistic uses of segment duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, **59**, 1208–1221.

KUHL, P.K. (1978). Predispositions for the perception of speech-sound categories: A species-specific phenomenon? In F.D. Minifie and L.L. Lloyd (eds.), *Communicative and Cognitive Abilities: Early Behavioral Assessment* (pp. 229–255). Baltimore: University Park Press.

KUHL, P.K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America*, **66**, 1668–1679.

KUHL, P.K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development*, **6**, 263–285.

KUHL, P.K. (1986). Reflections on infants' perception and representation of speech. In J. Perkell and

D.H. Klatt (eds.), *Invariance and Variability in Speech Processes* (pp. 19–30). Hillsdale, NJ: Erlbaum.

KUHL, P.K. (1987). Perception of speech and sound in early infancy. In P. Salapatek and L. Cohen (eds.), *Handbook of Infant Perception, Vol. 2* (pp. 275–381). New York: Academic Press.

KUHL, P.K., and MELTZOFF, A.N. (1984). The intermodal representation of speech in infants. *Infant Behavior and Development*, **7**, 361–381.

KUHL, P.K., and MILLER, J.D. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science*, **190**, 69–72.

KUHL, P.K., and MILLER, J.D. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America*, **63**, 905–917.

KUHL, P.K., and PADDEN, D.M. (1983). Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *Journal of the Acoustical Society of America*, **73**, 1003–1010.

LASKY, R.E., SYRDAL-LASKY, A.K., and KLEIN, R.E. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology*, **20**, 215–225.

LIBERMAN, A.M., COOPER, F.S., SHANKWEILER, D.P., and STUDDERT-KENNEDY, M. (1967). Perception of the speech code. *Psychological Review*, **74**, 431–461.

LIBERMAN, A.M., and STUDDERT-KENNEDY, M. (1978). Phonetic perception. In R. Held, H. Leibowitz, and H.-L. Teuber (eds.), *Handbook of Sensory Physiology. Vol. 3: Perception* (pp. 143–178). New York: Springer Verlag.

MACKAIN, K., STUDDERT-KENNEDY, M., SPIEKER, S., and STERN, D. (1983). Infant intermodal perception is a left-hemisphere function. *Science*, **219**, 1347–1349.

MARLER, P., and PETERS, S. (1981). Birdsong and speech: Evidence for special processing. In P.D. Eimas and J.L. Miller (eds.), *Perspectives on the Study of Speech* (pp. 75–112). Hillsdale, NJ: Erlbaum.

MARTIN, J.G. (1970). On judging pauses in spontaneous speech. *Journal of Verbal Learning and Verbal Behavior*, **9**, 75–78.

MASSARO, D.W., and COHEN, M.M. (1983). Phonological constraints in speech perception. *Perception and Psychophysics*, **34**, 338–348.

MATTINGLY, I.G. (1972). Speech cues and sign stimuli. *American Scientist*, **60**, 327–337.

MCGURK, H., and MACDONALD, J. (1976). Hearing lips and seeing voices. *Nature*, **264**, 746–748.

MEAL, R., and ERBER, J. (1978). Learning and memory in bees. *Scientific American*, **239**, 102–110.

MEHLER, J., BERTONCINI, J., BARRIÈRE, M., and JASSIK-GERSCHENFELD, D. (1978). Infant recognition of mother's voice. *Perception*, **7**, 491–497.

MEHLER, J., JUSCZYK, P.W., LAMBERTZ, G., HALSTED, N., BERTONCINI, J., and AMIEL-TISON, C. (1988). A precursor of language acquisition in young infants. *Cognition*, **29**, 143–178.

MILLER, J.L., and EIMAS, P.D. (1979). Organization in infant speech perception. *Canadian Journal of Psychology*, **33**, 353–367.

MILLER, J.L., and EIMAS, P.D. (1983). Studies on the categorization of speech by infants. *Cognition*, **13**, 135–165.

MILLS, M., and MELUISH, E. (1974). Recognition of the mother's voice in early infancy. *Nature*, **252**, 123–124.

MORGAN, J.L. (1986). *From Simple Input to Complex Grammar.* Cambridge, MA: MIT Press.

MORGAN, J.L., MEIER, R.P., and NEWPORT, E.L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, **19**, 498–550.

NAKATANI, L., and DUKES, K. (1977). Locus of segmental cues for word juncture. *Journal of the Acoustical Society of America*, **62**, 714–719.

OJEMANN, G. (1983). Brain organization for language from the perspective of electrical stimulation mapping. *The Behavioral and Brain Sciences*, **6**, 189–230.

PLOMP, R., and MIMPEN, A.M. (1979). Speech-reception threshold for sentences as a function of age and noise. *Journal of the Acoustical Society of America,* **66**, 1333–1342.

REMEZ, R.E., RUBIN, P.E., NYGAARD, L.C., and HOWELL, W.A. (1987). Perceptual normalization of vowels produced by sinusoidal voices. *Journal of Experimental Psychology: Human Perception and Performance,* **13**, 40–61.

REMEZ, R.E., RUBIN, P.E., PISONI, D.B., and CARRELL, T.D. (1981). Speech perception without traditional speech cues. *Science,* **212**, 947.

SNOW, C.E. (1972). Mothers' speech to children learning language. *Child Development,* **43**, 549–565.

SPELKE, E.S. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology,* **15**, 626–636.

SPELKE, E.S., BORN, W.S., and CHU, F. (1983). Perception of moving, sounding objects by four-month-old infants. *Perception,* **12**, 719–732.

STREETER, L.A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature,* **259**, 39–41.

SUMMERFIELD, Q. (1979). Use of visual information for phonetic perception. *Phonetica,* **36**, 314–331.

SUMMERFIELD, Q. (1988). *Visual Perception of Linguistic Gestures.* Paper presented June 6th at conference on "Modularity and the Motor Theory of Speech Perception". New Haven, Connecticut.

TINBERGEN, N. (1969). *The Study of Instinct.* Oxford: Oxford University Press.

TREHUB, S.E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development,* **47**, 466–472.

TREHUB, S.E. (1987). Infants' perception of musical patterns. *Perception and Psychophysics,* **41**, 635–641.

TREHUB, S.E., THORPE, L.A., and MORRONGIELLO, B. (1987). Organizational processes in infants' perception of auditory patterns. *Child Development,* **58**, 741–749.

WATERS, R.S., and WILSON, W.A. (1976). Speech perception by rhesus monkeys: The voicing distinction in synthesized labial and velar stop consonants. *Perception and Psychophysics,* **19**, 285–289.

WERKER, J.F., and LALONDE, C.E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology,* **24**, 672–683.

WERKER, J.F., and TEES, R.C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development,* **27**, 49–63.