Annu. Rev. Psychol. 1999. 50:509–35

INFLUENCES ON INFANT SPEECH PROCESSING: Toward a New Synthesis

Janet F. Werker and Richard C. Tees

Department of Psychology, University of British Columbia, Vancouver, British Columbia, V6T 1Z4, Canada; e-mail: Jwerker@cortex.psych.ubc.ca

KEY WORDS: cross-language, perception, development, models, word-learning

ABSTRACT

To comprehend and produce language, we must be able to recognize the sound patterns of our language and the rules for how these sounds "map on" to meaning. Human infants are born with a remarkable array of perceptual sensitivities that allow them to detect the basic properties that are common to the world's languages. During the first year of life, these sensitivities undergo modification reflecting an exquisite tuning to just that phonological information that is needed to map sound to meaning in the native language. We review this transition from language-general to language-specific perceptual sensitivity that occurs during the first year of life and consider whether the changes propel the child into word learning. To account for the broad-based initial sensitivities and subsequent reorganizations, we offer an integrated transactional framework based on the notion of a specialized perceptual-motor system that has evolved to serve human speech, but which functions in concert with other developing abilities. In so doing, we highlight the links between infant speech perception, babbling, and word learning.

CONTENTS

0084-6570/99/0201-0509$08.00

INTRODUCTION

What is remarkable about an infant's ability to successfully process the sounds of speech? The act of perceiving auditory speech represents a difficult and complex computational task. In written speech, each individual letter has its own form; there are spaces between words, and punctuation marks are used to indicate divisions into phrases, sentences, and paragraphs. Although we also think of spoken speech as linear and composed of discrete elements, the acoustic wave form shows no clear boundaries between individual phonemes (a basic sound unit, or phone, used in a language to distinguish one word from another) or between individual syllables, or even words. Moreover, the acoustic cues that signal the beginning of a new phrase or sentence are only probabilistic at best. Yet to perceive (and eventually produce) one's native language, it is essential that an infant not only successfully isolate and segment the individual units in the stream of speech, but that she also represent in some way the information that specifies the regularities among various productions of the same phoneme or word and ignore irrelevant variations.

During the past 30 years, researchers have focused on trying to understand how infants begin to solve the complex computational task of speech processing (for a review of common methods used to test infant speech perception, see Werker et al 1998a). We know that infants begin life with a number of perceptual-motor biases that allow them to "break into" the stream of speech, pull out and represent its units, and eventually map sound to meaning. As well, it is apparent that acoustic and phonological cues in spoken speech provide probabilistic information as to the boundaries of linguistic units, and that infants are well designed to detect and utilize this probabilistic information. Finally, researchers are beginning to identify the developmental achievements that allow infants to use the different kinds of information in the stream of speech in their elaboration of language-specific sensitivities.

WHAT LINGUISTIC SENSITIVITIES DOES THE YOUNG INFANT BRING TO SPEECH PROCESSING?

Sensitivity to Consonants and Vowels

Newborns begin life with a remarkable sensitivity to the acoustic cues that signify different basic elements of speech. For example, they are able to discriminate fine phonetic differences between consonants in syllables such as /ba/ versus /da/ or /ba/ versus /pa/ (Eimas et al 1971). The phonetic differences infants can discriminate most easily in consonants are those that actually occur in one or more of the world's languages. When presented with a series of 8–12 stimuli, with "anchors" synthesized to incorporate the critical acoustic information representing a consonant-vowel sound pair, and with intermediate stimuli to represent equal-step changes in the acoustical differences that differentiate the two consonants , 2- to 3-month-old infants show evidence of discrimination at the same places along the continuum as do adult speakers. For instance, they show evidence of discrimination at precisely those places on the continuum where adults shift their labeling from, for example, /ba/ to /da/ to /ga/ (Eimas 1974) (Figure 1). That young infants detect some equal-sized acoustic changes more readily than others, and that the changes they are able to discriminate "map on" to those used in the world's languages, reveals an initial set of perceptual sensitivities that enables the infant to begin to process the most fundamental information in human speech.

The ability to discriminate one vowel sound from another is also evident in young infants (Trehub 1973). Moreover, early on, infants show a cohesive internal structure to their vowel categories. For example, infants aged 2–4 months are able to perceive vowel identity across a variety of contexts, treating as equivalent the vowel /i/ as spoken by a man, woman, or child in either a rising or falling intonation contour, and distinguish it from the vowel /a/ spoken in these varying contexts (Kuhl 1979). This perception of isolated steady state vowels also seems to involve what Kuhl (1993) labeled a "magnet effect." Presented initially with "good" exemplars of a vowel category and then required to detect occurrences of "poor" instances, both 6-month-old infants and adults showed poorer vowel discrimination than when tested with "poor" exemplars prior to "good" ones (Grieser & Kuhl 1989).

Infants also show a sensitivity to visual information in speech. As adults, our speech percepts represent a combination of what we see and hear and under some circumstances are actually influenced more by what we see (Green

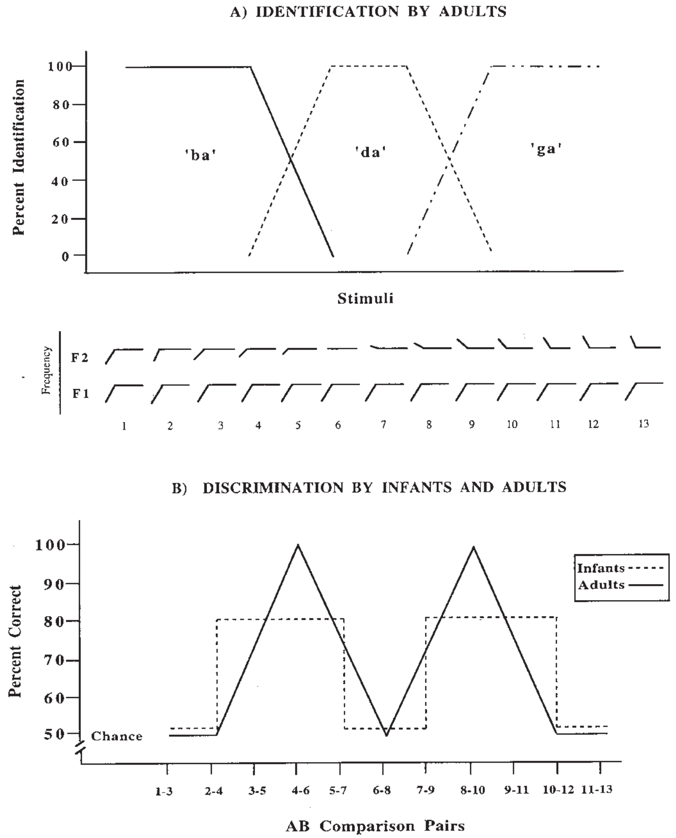

Figure 1 An illustration of adult and infant categorical perception of consonants using 13 syllables taken from the synthetic (and schematically represented) consonant-vowel continuum from /ba/ to /da/ to /ga/, signaled by differences in the starting frequency of the second format transition. (*A*) Idealized identification by adults. (*B*) Idealized discrimination by infants and adults. (Adapted from Eimas 1974, Strange & Jenkins 1978.)

1998). In illustration, McGurk & MacDonald (1976) presented subjects with "talking heads" accompanied by either matching or nonmatching speech sounds. When presented with an auditory /ba/ and a visible /ga/, adults consistently reported perceiving a /da/, a syllable intermediate between the "seen" and the "heard" syllables. Furthermore, when presented with an auditory /ba/ and a visible /va/, adults perceived only what they saw, i.e. /va/. What is remarkable is that although in both cases these percepts are different from the veridical information presented, adults perceive a strong, immediate, and unambiguous syllable.

By 4 months of age, infants can detect the match between heard and seen vowel information, looking preferentially to a talking head producing the matching vowel (Kuhl & Meltzoff 1982, Patterson & Werker 1998). However, the combined percept does not appear to be compelling and unambiguous at this young age . Instead, it appears that infants this young can attend to both the heard and seen sources of information (Rosenblum et al 1997), but that their integration is not yet mandatory or absolute (Desjardins & Werker 1996).

Sensitivity to Syntactic Information in Speech

Infants may be sensitive to the prosodic information that signals the overall grammatical structure of the language. Languages differ in the basic word order allowed. In English, commonly, sentences tend to obey a subject/verb/object structure, wherein complementizing information follows the verb (such languages are called right-branching languages). In left-branching subject/object/verb languages (such as Turkish), the complementizing information precedes the verb (Chomsky 1975). The overall prosodic structure of subject/verb/object languages is different from that of subject/object/verb languages (Nespor & Vogel 1986), yielding perceptible acoustic correlates. Recently, researchers have hypothesized that newborn infants might be sensitive to this prosodic information (Christophe et al 1997, Mazuka 1996). Although there have not yet been studies with newborns, infants 9 months old are able to perceive phonological phrase boundaries (Gerken et al 1994).

Moreover, newborn infants are sensitive to the acoustic information that differentiates broad grammatical classes of words. Acoustic cues to such categories occur in adult speech (Kelly 1992). Across languages, the most fundamental distinction among grammatical classes is the bifurcation into lexical and grammatical words. Lexical words are content words, such as nouns, verb, adjectives, etc, which carry meaning. Grammatical words are function words such as prepositions, articles, etc, which primarily carry structure. Shi et al (1998a) showed that there are acoustic and phonological cues in speech directed to infants that distinguish content and function words. For example, content words tend to be longer and they tend to be spoken more loudly and to have their vowels more fully enunciated than function words. Recently, Shi et al (1998b) tested newborn infants on their ability to distinguish words on the basis of grammatical category. Infants were familiarized first to a list of content (or a list of function) words and then tested on a new list of words either from the same category or from the contrasting category. Infants showed evidence of detecting a switch to the new category but not a switch to a new words from the same category. Thus, one of the perceptual sensitivities infants bring to the task of perceiving speech is a sensitivity to the acoustic and phonological cues that distinguish the two fundamental grammatical categories.

Summary of Initial Infant Sensitivities

Taken together, the findings reviewed above reveal a remarkable level of sophistication in infants' speech processing capabilities. Undoubtedly, these initial biases and capabilities allow the infant to begin to gather the kind of information necessary to move one step closer to the eventual task of language acquisition.

THE EFFECTS OF EXPERIENCE ON INFANT SPEECH PERCEPTION

Languages differ in their inventory of consonants and vowels in both auditory (phonemes) and visible (visemes) speech. They differ as well (*a*) in the frequencies with which their particular consonants and vowels occur, (*b*) in the precise acoustic-phonetic characteristics of the consonants and vowels in different positions in words, i.e. allophonic variations, and (*c*) in the allowable combinations of consonants and vowels (phonotactics) that can occur within words. Also, as noted earlier, the precise set of grammatical categories used (over and above the fundamental lexical versus grammatical word distinction) differs between languages, and the branching parameter is set differently in different languages families. By studying perception in infants raised in different language environments, we can begin to describe how and when infants become attuned to the properties of their native language.

Recognizing One's Native Language

One intriguing finding is that some adaptation to the properties of one's native language occurs either in utero or immediately following birth. English- and Spanish-learning infants as young as 2 days show a preference for listening to their native language (Moon et al 1993). Moreover, within the first few hours and days of life, infants are able to discriminate excerpts from distinct language families. In an early demonstration of this, Mehler and colleagues (Mehler et al 1988, Mehler & Christophe 1994) tested neonates in Paris and infants aged 2 months from Oregon on French versus Russian and English versus Italian. They found that neonates from both environments were able to discriminate between the familiar and the unfamiliar language or between two different unfamiliar languages. More recently, this finding has been qualified. The two unfamiliar languages must be from relatively distinct language families to be discriminable to a newborn. For example, Nazzi et al (1998) reported that although newborn French infants can discriminate English from Japanese, they cannot discriminate English from the rhythmically more similar German. By 4 months of age, however, infants raised in a monolingual environment may be able to distinguish their native language from a very similar unfamiliar language. The evidence for this comes from work of Nazzi et al (1998), showing that English-learning infants aged 4–5 months can distinguish English from the highly similar Dutch, and from the very recent study in which Bosch & Sebastian-Galles (1997) showed that monolingual 4-month-old Spanish- or Catalan-learning infants are able to discriminate Spanish from Catalan. However, 4-month-old infants being raised in bilingual Spanish/Catalan environments failed to provide evidence of discriminating between these two languages. This first study with bilingual infants raises a number of interesting questions regarding the nature of language representations in early bilingualism.

Cross-Language Consonant and Vowel Perception

A research focus over the past 15 years has been on understanding the effects of experience on perception of consonants. Infants younger than 6–8 months of age can not only discriminate categorically native phonetic contrasts (e.g. /ba/ versus /da/, or /ra/ versus /la/), they can also discriminate phonetic contrasts involving syllables that are not used to distinguish meaning in their native language (Aslin et al 1981, Best et al 1988, Polka & Werker 1994, Streeter 1976, Trehub 1976, Werker et al 1981, Werker & Tees 1984a).

In contrast to the language-general sensitivities shown by young infants, adults often have difficulty discriminating between syllables that differ by only a single phoneme, if that particular phonemic contrast is not used in their native language: Japanese adults have difficulty discriminating the difference between /ra/ and /la/ (Strange & Jenkins 1978), and English adults have difficulty discriminating certain Hindi (Werker et al 1981), Nthlakampx (Werker & Tees 1984a), and Czech (Trehub 1976) contrasts. Adults may need short familiarization periods even to discriminate acoustically quite salient non-native distinctions (Pisoni et al 1982, Werker & Tees 1984b). Thus, although the sensitivities of infants allow language-universal phonetic discrimination, their subsequent experience functions to narrow, or "prune," their perceptual sensitivities to enable "mapping on" to the phonology of their native language.

To examine the effects of early experience, we conducted a series of studies comparing infants of different ages, children, and adults on their ability to discriminate non-native phonetic contrasts. The comparison of either Hindi-speaking or English-speaking adults with English-learning infants showed results consistent with the prediction of language-universal infant sensitivities and their subsequent decline. Virtually all subjects in all groups could discriminate the /ba/-/da/ contrast—a distinction common in the world's languages and one that is used in both English and Hindi. The 6- to 8-month-old English-learning infants and the Hindi-speaking adults could easily discriminate both Hindi contrasts. However, the English-speaking adults had trouble discriminating the Hindi contrasts, showing particular trouble with the more difficult retroflex/dental place-of-articulation distinction (Werker et al 1981).

In follow-up studies, we found that this change in ability is evident by 4 years of age (Werker & Tees 1983) and, in fact, occurs within the first year of life. English-learning infants of 6–8, 8–10, and 10–12 months of age were tested on their ability to discriminate the Hindi /Ta/-/ta/ contrast and a Nthlakampx /k'i/-/q'i/ contrast. Although most of the infants 6–8 months old were able to discriminate between both non-English contrasts, few of the infants 10–12 months old were able to discriminate either the Hindi or the Nthlakampx contrast (Werker & Tees 1984a). The pattern of results revealed for infants is shown in Figure 2.

This pattern of change between 6 and 12 months of age has been reported (*a*) for a different retroflex/dental distinction (/Da/-/da/) (Werker & Lalonde 1988); (*b*) for three Zulu contrasts: a bilabial plosive/implosive distinction, a lateral fricative voiced/voiceless contrast, and a velar voiceless/ejective stop distinction (Best 1995); and (*c*) among Japanese infants for the (non-Japanese) English /ra/-lla/ (Kuhl 1998). The change for the Nthlakampx contrast has also been replicated by Best (1995). Importantly, however, the decline in cross-language consonant perception is not always evident at 10–12 months of age. Best and colleagues (see Best 1995) have shown that the decline only occurs for contrasts that involve sounds similar to sounds used in the native language. For example, infants 10–12 months old, and even adults, continue to discriminate the apical/lateral Zulu click contrast <xa>-<ca>, but this contrast sounds to all but the Zulu more like someone clucking to a horse or making a "tsk tsk" sound

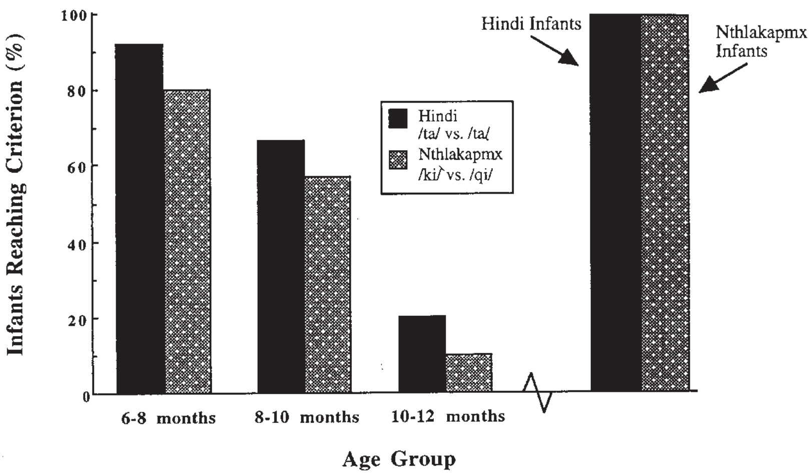

Figure 2 The proportion of infants at each age reaching discrimination criterion on the Hindi and Nthlakampx contrasts. (*Far right*) The performance of infants 11 months old raised in either a Hindi or a Nthlakampx environment. (Adapted from Werker & Tees 1984a.)

than like a linguistic sound. Best suggests that it is experience-dependent assimilability to the phonology of the native language rather than linguistic experience per se that accounts for the maintenance of discriminative abilities.

To further examine what kind of experience is important, Pegg & Werker (1997) conducted a study assessing infants and adults on their ability to recognize a phonetic difference that occurs but is not used to distinguish meaning in the native language. They investigated the ability of English-speaking infants and adults to discriminate the syllable /da/ from the syllable created by removing the [s] in a /sta/ syllable. /Sta/ without the [s] was perceived to be /da/ by the adults, who although labeling all these stimuli as equally good instances of the English phoneme /da/ were nevertheless able to discriminate the two sets of stimuli (the /da/ set from the [s]/ta/ set). Importantly, approximately half of the English-learning infants 6–8 months of age were also able to discriminate the two sets of stimuli, whereas nearly all the English infants 10–12 months of age failed this discrimination task.

One interpretation of the above data is that by 10–12 months of age, infants selectively listen to only that phonetic variation in the native language that conveys meaningful distinctions. A simpler explanation, and the one we prefer, is that by 10–12 months of age, infants are sensitive not only to the phonetic characteristics of the native language, but also to the syllabic context in which that phonetic variation occurs. Thus, when presented with the [s]/ta/ phonetic variate in syllable initial position, infants treated it as an instance of the closest context-appropriate form, which is /da/. This kind of sensitivity would allow infants to represent and attend to acceptable syllable forms in their native language.

The language-specific influences on vowel perception may be evident at an earlier age than those seen for the perception of consonants. For example, Kuhl and colleagues (1992) found that at 6 months of age, Swedish-learning infants showed the magnet effect (described above) when tested on variates of Swedish vowels but not of English vowels, whereas English-learning 6-month-olds showed the magnet effect only for the English vowels.

Polka & Werker (1994) extended the Kuhl work to assess not just possible language-specific influences on the internal structure of vowel categories but also age-related changes in between-category discrimination of non-native vowels. Their work revealed an effect of experience between 4 and 6 months of age, but the full impact of experience was not evident until approximately 10–12 months of age, the same age at which it is found for consonants (but see Polka & Bohn 1996).

Experiential Influences on Bimodal Speech Perception

Although no one has been able to assess experiential influences on bimodal speech perception in infants, postnatal experience does influence bimodal speech perception in children and adults. In an early study, Werker et al (1992) examined bimodal speech perception in five groups of French-speaking Canadians who were studying English in a summer program. Of interest was the perception of the interdental fricative /tha/, which occurs in English but not in French. Canadian French speakers tend to substitute /tha/ with /ta/ or /da/ when speaking English. Perhaps not surprisingly, when tested on their perception of the visible information in /tha/, Canadian French speakers tended to perceive /ta/ or /da/ rather than /tha/. The strength of this effect was inversely proportional to their fluency in English. A similar effect of experience on perception of visible speech was shown by Massaro et al (1993).

Several researchers have presented data indicating that the overall amount of visible influence increases with age (e.g. McGurk & MacDonald 1976). This suggests a role for experience in the perception of visible speech. Recent research is beginning to reveal the kinds of experience that might be important. To illustrate, Desjardin et al (1997) tested children aged 3 and 4 who were either still making or not making typical substitution errors in their production of any one of the consonants /b/, /v/, /th/, or /d/. The children were tested in an engaging bimodal speech perception task, an auditory speech perception task, and a lip-reading task. Although both the substitutors and the nonsubstitutors performed equally well in the auditory speech perception task, the substitutors showed much less influence from the visible speech in their performance in both the bimodal speech perception and the lip-reading tasks. Similarly, Siva et al (1995) showed that the audible and visible information did not appear to be as well integrated in the percepts of adults with cerebral palsy. Taken together, these two studies provide strong evidence that experience producing speech correctly is at least one source of influence on the effective use of visible information.

Prosodic, Phonotactic, and Syntactic Information Processing

In a series of studies, Jusczyk and colleagues have shown that by 6 months of age, infants show a preference for listening to words that correspond to the prosodic (rhythm and intonation) patterns of their native language, but it is not until 9 months of age that they show a preference for listening to words that correspond to the phonetic and phonotactic rules of their native language (Jusczyk 1997). Moreover, by 10 months of age, but not before, infants show a preference for listening to lists of words that conform to the predominant native language strong-weak (SW) stress pattern (Juscyk et al 1993).

Only recently, researchers have begun to assess experiential influences on perception of grammatical information. The question being asked is whether infants can use phonological and acoustic cues to bootstrap into the grammatical structure of the native language. In the first demonstration, Hirsh-Pasek and colleagues (1987) showed that by 7 months of age, infants show a preference for listening to infant-directed speech samples with pauses inserted at phrase boundaries over similar samples wherein the pauses are inserted within a phrase. Jusczyk (1997) extended this work to a cross-language sample. He found that at 4 months, English-learning infants listened equally long to both Polish and English speech samples with pauses inserted at clause boundaries over speech samples with pauses inserted within clauses, but by 6 months of age they only showed this preference for correctly inserted pauses with their native language (English). Thus, between 4 and 6 months of age, infants apparently lose their prior ability to detect the acoustic and phonological cues demarking major grammatical constituents in an unfamiliar language and retain a sensitivity to only those cues that are important in their language.

As noted earlier, Shi et al (1998a) found that newborn infants can discriminate between sets of grammatical and lexical words and that they show detection of a change irrespective of the direction of the change. By 6 months of age, however, infants show a detection of the change in syntactic category only if they are initially familiarized/habituated to grammatical words and then tested on lexical words, i.e. they dishabituate only under those circumstances. Shi and colleagues interpret these findings as showing an emerging preference for lexical words between birth and 6 months of age.

Do These Age-Related Changes Represent Losses in Linguistic/Perceptual Competence?

The language-general perceptual sensitivities in newborns undergo a change and become more language-specific during the first year of life. When first reported, the research community viewed the data as indicating a loss of perceptual capacity due to lack of experience. However, the interpretation has since become more precise. First, as we cited earlier, infants continue discriminating between some non-native phonetic contrasts even though they have never heard them, (e.g. Best et al 1988) and lose the ability to discriminate between others even though they are part of heard speech (Pegg & Werker 1997). Second, even though adults perform poorly on many non-native phoneme contrasts in the testing circumstances we have described, there are other conditions under which continuing adult sensitivity can be demonstrated. Adults can discriminate even difficult non-native contrasts if the critical acoustic information in the speech contrast is presented alone so that the now-truncated syllables no longer sound like speech (Werker & Tees 1984b). Furthermore, adults can be taught to distinguish full syllables if given enough training trials (Tees & Werker 1984), or if tested in sensitive procedures with low memory demands (e.g. Pisoni et al 1982, Werker & Tees 1984b), or if given extensive language instruction (MacKain et al 1981). For these reasons, we have referred to the changes that occur, at least in consonant discriminative ability, as a functional reorganization (Werker 1995) and to the early influences of linguistic experiences as reflective of a sensitive, not critical (i.e. invariant), period (Tees 1986, Werker & Tees 1992). Its nature is further considered at the end of this review.

Summary of Experiential Influences on Infant Speech Processing

In summary, the evidence establishes convincingly that by the end of the first year of life, an infant's phonetic perceptual sensitivities reflect considerable influence from the native language. This influence is reflected both in a preference for highly frequent phonetic patterns and in a narrowing of initial discriminatory abilities to match the contextual distribution of phonetic information from auditory (linguistic) input. As well, by one year of age, infants also show preferential processing for many other aspects of the native language. These changes, which occur during the first year of life, appear to prepare the child for the next functional task—beginning to acquire the ability to understand and speak her native language.

LISTENING FOR WORDS

What Do Infants Know About Words?

Few studies to date have assessed the relationship between the speech perceptual competencies of developing infants and their emerging word-learning. In this section, we briefly review recent work on word-learning in infancy and examine whether or not there is a direct link between changing sensitivity to the native language and the onset of word-learning.

During their first 14–15 months, infants learn to extract words from the speech stream, to recognize word forms they have previously heard, to associate words with objects, to understand the meaning of some words, and even to produce some words. Importantly, although infants are learning a lot about words, none of the kinds of word-learning they show in the first year of life necessarily indicates they have a full referential understanding of words. Indeed, evoked responses show that in 15-month-olds, known words and unknown words are processed differently, but not until 20 months of age (by which time they do have a referential understanding of word meaning) is the recognition of familiar words strongly lateralized to the left hemisphere and to the speech processing areas over the planum temporale (Mills et al 1994). Researchers who study early language acquisition have suggested for some time that children may not attend to fine phonetic detail when they are first learning what words mean, even though they displayed such attention to fine phonetic detail in their perception of speech at a younger age. Indeed, early work suggested that young children either confuse similar-sounding words or tolerate single feature phonetic substitutions in known words (e.g. Eilers & Oller 1976). One proposal to account for this argued that infants initially represent only the vocalic nucleus of the word (Ferguson & Farwell 1975). However, the findings with young children are unconvincing because the methods used to test their perception of words differed greatly in task demands and sensitivity compared with those used to collect evidence about the infant's speech perception. Thus, it remains a matter of conjecture whether children do in fact represent words as meaningful entities with less detail than they do non-meaningful word forms. Recently, investigators have revisited the issue of whether or not infants use fine phonetic detail when they first start learning to understand words. This evidence is discussed below with respect to the various stages of word-learning from recognition and segmentation of word forms, through associative and referential understanding (for a full discussion of various stages of word learning, see Werker & Stager 1998).

Recognizing Word Forms

Even neonates show some sensitivity to the acoustic cues that specify word boundaries. Infants 1–4 days old are able to distinguish two-syllable stimuli in which the two syllables either were excised from a single word (e.g. [mati] from "cinematique") or from two different words (e.g. [mati] from "panorama typique") (Christophe et al 1994). By 4.5 months of age, infants show a preference for listening to their own names, which suggests that even at this early age, infants are beginning to recognize something about word forms (Mandel et al 1995). By 7.5 months of age, when presented with passages containing familiar words, such as "cup," "bike," etc, infants (Jusczyk & Aslin 1995) show a preference for listening to those over unfamiliar words and can remember heard words for up to two weeks (Jusczyk & Hohne 1997). Furthermore, infants of this age can segment and remember bisyllable words that conform to the dominant, SW stress pattern in English (e.g. "doctor"), but fail to pull out the first syllable when familiarized to weak-strong (WS) words. For example, if familiarized to the word "belief," infants of 7.5 months show a preference for listening to "lief" rather than "belief." By 10 months of age, infants successfully pull out words irrespective of whether the SW or WS form is used and successfully recognize even WS forms (Jusczyk 1997).

The question of whether infants detect and represent fine phonetic detail in word-priming tasks has also been addressed. In one study, Jusczyk & Aslin (1995) familiarized infants to a set of words, then tested them on "foils," words that are phonetically similar (e.g. "tup" rather than "cup"). Although infants 7.5 months old once again showed a preference for listening to the words they had been presented with during the familiarization phase, that preference did not extend to the phonetically similar foils. Thus, this suggests that infants are able to detect, access, and use fine phonetic detail to distinguish familiar from unfamiliar items.

Different results were obtained by Halle & de Boysson-Bardies (1996) for slightly older infants. They reported that by 11.5 months of age, infants were able to recognize frequently heard words and showed a preference for listening to such words over the infrequently heard foils, but only if the foils were phonetically quite dissimilar. When tested with phonetically similar foils, infants 11–12 months old treated them as equally familiar. To explain the difference between their findings and those of Jusczyk & Aslin (1995), Hallé & de Boysson-Bardies (1996) proposed that in contrast to 7.5-month-old infants, 11-month-old infants listen to words as potential sources of semantic content, and that such a listening strategy may predispose them to adopt a more holistic, less analytic processing strategy than their younger counterparts.

Word-Object Associative Learning

To test the hypothesis that infants access or store less phonetic detail when they are listening to words as potential sources of meaning than when they are just listening to words as sounds, we decided to test infants on their ability to discriminate among fine phonetic features in more-explicit word-learning tasks. Using a "habituation switch" procedure, infants were familiarized to one or two word-object pairings (Werker et al 1998b). Following the familiarization phase, infants were then tested on their ability to detect a change in either the word, the object, or the pairing of the word with the object. Werker and colleagues (1998a) found that infants 14 months old, but not younger, could learn the association between two words and two moving objects with only minimal exposure when the objects used were physically dissimilar and the words used were phonetically dissimilar (e.g. "lif" and "neem"). In a follow-up series of studies, Stager & Werker (1997) tested infants on their ability to learn the association between phonetically similar words (e.g. "bih" versus "dih") and objects. Surprisingly, when the words were phonetically similar, infants of 14 months failed.

This kind of associative task can be regarded as more demanding than a task that simply requires children to recognize a familiar word form, but less demanding than one that requires referential comprehension. It does not require that the infant adopt or grasp the concept "stands for," but it does require the infant to use the word form in a minimally semantic way. That is, the successful child at least must perceive and remember the "goes with" association between the word and object, whereas no such "goes with" understanding is required in the recognition task used by Jusczyk and colleagues.

The apparent discrepancy between the results obtained when speech perception of infants is assessed and those obtained when word-learning is tested in the same infants is a reflection, we argue, of the real discontinuity in the nature of processing strategies in operation in the two circumstances. It is for this reason that we refer to the inattention to phonetic detail in early word-learning tasks as a second functional reorganization. Further research is required to determine when infants again recover access to fine phonetic detail in word forms. However, we suspect it may not be evident until after the onset of full referential understanding. This referential understanding is thought to allow the rapid spurt in vocabulary growth that occurs at approximately 18–20 months of age. The resulting increased vocabulary could result in sufficient pressure to fill in finer phonetic detail in the lexical representations in order to avoid confusion between similar-sounding, known words. Although we have not yet successfully tested infants 18–20 months old, in the switch procedure, we would predict that at, or shortly after, this age, access to such fine phonetic detail would be evident. The known and proposed relationships between onset of different kinds of word-learning and changes in sensitivity to the amount of phonetic detail is shown in Figure 3.

VOCAL PRODUCTION

Babbling

Although infants begin life with a broad-based, high level of speech perceptual ability, their ability to produce speech is clearly limited by the immaturity of

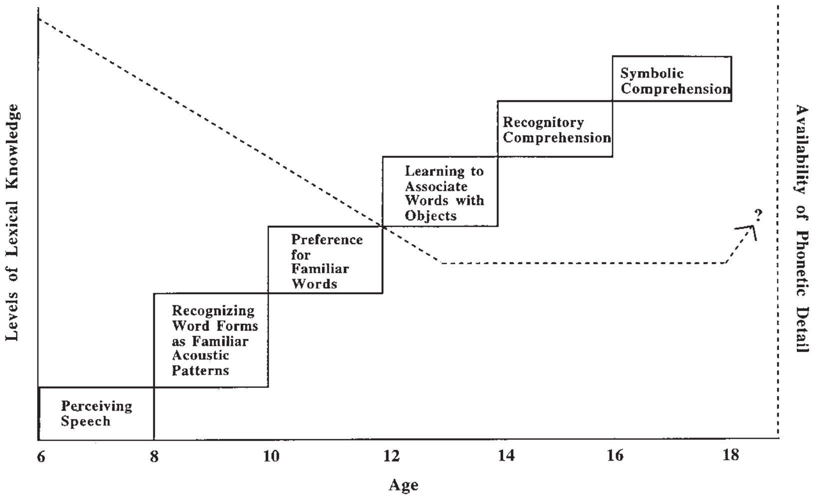

Figure 3 A pictoral representation of the amount of phonetic detail represented or accessed by the infant (*dotted line*) as a function of level of complexity in speech perception and word processing tasks (*boxes*).

both the vocal tract and the related neuromusculature (see Kent & Murray 1982). It is not until approximately 6 months of age that infants first display coordination of the gestures involved in closing the vocal tract for a consonant and the release of a steady state vowel in the form of a consonant-vowel (CV) syllable (Oller 1980). These early canonical forms involve reduplicated productions of the same CV syllable (e.g. /bababa/). By 8–9 months of age, infants begin to be able to combine varying CV syllables in a single production (e.g. /babeedo/). Early appearing syllable forms are those that are the easiest to produce (see Locke 1983). Thus, in the initial stages of development, there do appear to be universals in vocal production.

The question of whether (and when) babbling begins to reflect an influence from the native language has been the subject of great controversy. In the past 15 years, convincing evidence has been collected that infant vocalizations do reflect the properties of the native language and that they begin to do so by 9–10 months of age. By this age, the formant structure of vowels produced by infants is closer to that of their native language than to that of comparison languages (de Boysson-Bardies et al 1989). More convincingly, in a longitudinal study of French, Swedish, English, and Japanese infants, de Boysson-Bardies & Vihman (1991) found that in babbling, the distribution of manner class (e.g. stop consonants such as /b, d, t, k/ versus fricatives such as /sh, th, f, v/ versus nasals such as /m, n/) and of place-of articulation becomes more language-specific by 10 months of age. By that age, both the timing of syllable productions (Levitt et al 1992) and the intonational patterns in infant babbling (Whalen et al 1991) begin to match those of the native language. Thus, we see in babbling the emerging language-specificity we observed in perception.

The Production of Words

What is the relationship between babbling and speech? In his early writings, Jakobson (1941/1968) proposed that there is a discontinuity between babbling and speech. He argued that the sounds produced prior to meaning had no relation at all to those used once an infant attempted to produce words. Indeed, he even suggested that there was a period of silence between the babbling period and the onset of true word production in many children, and that this underlined the fact that one vocal production system was being supplanted by another. At the onset of word production, infants were believed to add sounds in a regular and systematic way, in essence filling out the structure of a formal phonological system.

More recently, it has been shown that infants do not stop babbling prior to beginning to speak. Furthermore, although the phoneme inventories used in babbling and speech may not be identical, careful transcriptions of children's production make clear that those sounds they can articulate well in babbling influence those they attempt to make when first producing words (e.g. Vihman & Miller 1988).

Although infants select sounds they are capable of articulating when attempting to produce words, there are differences between babbling and speech. First, in the initial stages of word learning children do not always produce correctly or consistently even sounds they have mastered in babbling. Indeed, children often eliminate, substitute, or mix the order of segments (e.g. Ferguson & Farwell 1975). Moreover, a child might substitute for one word a sound that she seemed unable to produce in a different setting, for example calling "dog" "gog" but calling "truck" "duck" (Gerken 1994). This variability suggests that when first learning words, infants may not "represent" all the detail found in adult speech. Indeed, it has been suggested that they may only represent sufficient information to contrast the words in their own lexicon (Menyuk et al 1979).

The obvious parallel between production and comprehension is striking, and as reported above, Stager & Werker (1997) showed that infants aged 14 months do confuse similar-sounding words in a word-object association task, which suggests that either not all the phonetic detail of the words is represented by the infant, or that the infant does not use it. Thus, both in the increasing specificity of their babbling toward the end of the first year of life and in the inconsistency of phonetic detail in early word production, we see parallels to the functional reorganizations involved in speech perception and word comprehension.

TOWARD AN EXPLANATION

Several models have been proposed to explain the data detailed above. Any complete explanation must account for the initial sensitivities shown in early infancy, the age-related changes in speech processing which occur during the first year of life, and the inattention to phonetic detail in the initial stages of word-learning. The explanations proposed to date span the spectrum from strong nativist models to extreme empiricist approaches. In this section, we review some of these models and attempt to provide a new synthesis.

Nativist and Ecological Models

According to "strong" nativist models, the initial speech perception capabilities of infants reflect a special-purpose speech-processing module, evolved to detect and analyze the essential properties of human language, available "online" early in ontogeny. The most well-known of these models is the motor theory of speech perception (Liberman & Mattingly 1985), in which specialized computational routines analyze phonetic input in terms of the potential mode of production. Research on the automatic and early coordination of heard-and-seen speech (bimodal speech perception) is seen as support for this explanation, as is evidence that the same areas of brain are activated in both the production and perception of speech (Ojemann 1991, Sams et al 1991).

To account for age-related changes in perceptual performance that reveal increased discriminability of the properties of the native language and decreased sensitivity to non-native phonemes, phonotactics, or rhythmical patterns, proponents of such approaches turn to learnability theories (e.g. Wexler & Culicover 1980). In such models, all possible parameters for language-universal rules (of the proposed device) are "given" at birth, and the contribution of specific experiences is to select some settings for the parameters and effectively "turn off" others. Thus, at birth, infants would be sensitive to both native and non-native phonetic contrasts, and experiential influence would cause those settings that correspond to the consonants and vowels used in the native language to remain activated and allow to be deactivated those that correspond to non-native phones.

Any similarity between production and perception could be easily and logically dealt with. The setting of parameters in such a module could simultaneously affect both perception and production. It is not clear what explanation would be offered to account for the inattention to phonetic detail in the early stages of word-learning, but whatever was proposed would not necessarily need to be tied to that offered to explain age related changes in perception.

Another approach, suggested by advocates of the Gibsonian ecological perspective, explains perception as direct access to the generating properties of the distal object. In the case of speech perception, information in the proximal acoustic wave form is thought to provide direct access to the articulators that produced the sounds. Thus, as in the motor theory, it is believed that perception of at least phonetic information can be explained by reference to production (Fowler & Rosenblum 1991). Furthermore, as in the motor theory, much of the bimodal speech perception research is seen as support for the viability of this approach. However, in the ecological approach, no special-purpose, built-in module need exist: The information is in the world to be "picked up directly." The experiential changes in speech perception that occur during infancy are explained by increasing attention to the specific acoustic information in the wave-form that reflects language-specific properties of vocal production (see Best 1995) for a model in this genre.

No attempt by those adopting this approach has been made to account for the dropping of phonetic detail in the initial stages of word-learning. However, as with the strong nativist approaches, the ecological model need not explain word-learning within the same framework used to explain changes in speech perception.

Evolution and Induction

According to non-nativist models, initial speech perception capabilities reflect the operation of a general auditory perceptual system, a product of mammalian evolution, and age-related changes in speech processing reflect experientially "induced " or "self-organized" systems (e.g. Kuhl 1988, Lindblom 1992). There is no exclusive specialized module or computational device built-in for the perception of speech. Instead, general-purpose auditory sensitivities are seen as contributing to speech perception, and language is said to have evolved either through phylogeny or ontogeny to take advantage of these auditory sensitivities. It is in discussions about the ways experience might induce language-specific linguistic perception that these approaches have made the most exciting advances.

One popular descriptive model of this genre is the "perceptual magnet" approach offered by Kuhl and colleagues (e.g. Kuhl 1993). The basic findings showing the existence of a native language perceptual magnet by 6 months of age were reviewed earlier. The developmental explanation Kuhl offers assumes that, even early in life, some regions in the vowel space are more stable and discriminable than others (see Lindblom 1992), yielding some initial rudimentary neural/perceptual organization to the vowel space. Repeated experience hearing certain acoustic forms (instances of vowels) more often than others remodel the perceptual space by rendering the frequently heard vowels more effective attractors.

The results from the studies of early word-learning, in which infants have been shown to ignore phonetic detail (Halle & de Boysson-Bardies 1996, Stager & Werker 1997), fit into the perceptual magnet model (PK Kuhl, personal communication). Basically, as words are established in the lexicon, the existing words act as a magnet, attracting similar-sounding words to them. Only with the establishment of additional words (following repeated exposure) would new magnets of sufficient strength to establish discrete spheres be created. Although this is an interesting descriptive model, no precise mechanism has been posited.

An approach that is attracting increasing attention is prosodic or phonological bootstrapping. This refers to the possibile existence in input speech of acoustic cues to linguistic structure, and that infants might be able to use this information to help "bootstrap" themselves into knowledge of language (Gleitman et al 1988; for a recent discussion of such phonological bootstrapping, see Morgan & Demuth 1996). A major problem to the prosodic or phonological bootstrapping theory is the lack of consistency in the information. For example, strong syllables provide good cues to word onsets in English, but only approximately 65% of English words start with strong syllables (Cutler & Carter 1987). Nevertheless, adult English speakers have learned to take advantage of this cue in word processing (e.g. Cutler & Norris 1988). Similarly, prosodic cues signaling phonological phrases correspond roughly to syntactic boundaries in English (Demuth 1996), but again, the match is far from perfect. Thus, if infants are to be able to use acoustic and phonological cues to "bootstrap" into a knowledge of language, they need to be able to take advantage of (i.e. learn from) imperfect, probabilistic information.

Recent research shows that even very young infants are indeed able to use frequency, density, and probabilistic information in input speech to establish significant linguistic knowledge (Aslin 1993). For example, in a recent study, Saffran et al (1996) showed that infants 8 months old were able to "learn" acceptable sequences of words in an artificial language when the only information they were given was probabilistic. By 10 months of age, infants learn about phonotactic acceptability on the basis of frequency of occurrence (Jusczyk et al 1993), and by 9 months of age they can use probabilistic cues in input speech to detect phonological phrase boundaries (Gerken et al (1994). In their recent work, Shi and colleagues (1998a) showed that even infants as young as 2–3 days can use probabilistic acoustic and phonological cues to distinguish between the fundamental categories of grammatical and lexical words.

These are but a few of the many demonstrations from the past few years showing an infant's ability to apply general learning techniques to use frequency and probabilistic cues to learn about the structure of his native language (Aslin 1993, Jusczyk 1997). We argue below, however, that although it is undeniable that probabilistic accounts can explain many of the important findings reviewed above, we believe there are limitations. As well, we have difficulty accepting nativist accounts without some mechanism proposed to account for the biases present in the initial state. Thus, we suggest an alternative explanation, neither fully nativist nor fully empiricist, that is more epigenetic in nature.

A Probabilistic Epigenetic Model

We propose a more epigenetic model to account for both the initial state of infant speech perception and the subsequent changes with age. To account for the sensitivities in newborn infants, we suggest there is already a history of interaction between a genetically initiated, overelaborated neural substrate and the normally invariantly occurring critical, species-specific experience (in this case, human speech). Building on Hebb's original idea (1949) about the importance of ubiquitous early experiential factors, Greenough (1986) has characterized this type of influence as involving "experience expectant" brain development. Briefly, decades of research in developmental biology have shown that genes are not deterministic, rather they express themselves in different ways depending on environmental influences (cf Gottlieb 1991). One process that appears common in early neural development is an overproduction of synapses followed by selective retention of a subset. This mechanism appears to have evolved to allow incorporation of expected experiences to remodel, or sculpt, the genetically initiated developing mechanisms that underlie a variety of perceptual and other systems (Cowan et al 1984). The experience-expectant modifications can be relied on because appropriate input is virtually always going to be present for members of the species. Furthermore, although experience-dependent changes occur throughout the life span and thus can often be redone, experience-expectant processes are time-locked, tending to occur in very early development and to result in a stable, and relatively permanent, platform of neural architecture (for further discussion, see Werker & Tees 1992) Psychologists would characterize the result as reflecting "innately guided learning" (e.g. Jusczyk & Bertoncini 1988).

We argue that these experience-expectant, interactive changes account for the speech-specific biases shown by the human neonate. By the sixth month of gestation in humans, the peripheral auditory system is fully functioning. However, connections between areas of the brain, and even between the sense organs and neural structures, have yet to be fully established (Kolb & Fantie 1989), and this creates a situation whereby genetic instructions for the establishment of brain connections can be influenced by information arriving from the (newly functioning) peripheral auditory system. With the human voice being a regular and reliable source of input to the fetus through both air and bone conduction (Moon et al 1993), the neural substrate can become organized to respond preferentially to sounds that could be produced by a human vocal tract, and to process both seen and heard speech from an early age (for a full elaboration, see Werker & Tees 1992). Similarly, preferential sensitivity to linguistic information of other kinds could result from such experience-expectant brain modifications. This resembles nativist theory, but it provides a developmental stimulation history to explain the emergence of language-general sensitivities.

To explain the tuning to the properties of the native language during the first year of life, we also suggest an epigenetic process involving changing brain structures and experiential influences. There is no question that the brain (e.g. frontal cortex) continues to mature during the first year of life (Kolb & Fantie 1989), and the organization of language and cognitive-related neural substrates are remodeled through both additive (i.e. experience-dependent) and subtractive (experience-expectant) processes. Experientially, the child is hearing speech and language and is beginning to vocalize with greater and greater control. Undoubtedly, the input from both heard speech and self-vocalizations has an impact on the emerging neural organization that may lead to a preferential perceptual sensitivity to language-specific phonetic information, and perhaps to the emerging native language specificity in production. And finally, it is apparent that general experience-dependent information storage capabilities of infants are employed to detect, remember, and use both frequency and probabilistic information to learn about the phonetic, rhythmical, and syntactic properties of the native language. We would argue, however, that there are points of discontinuity, leading to the emergence of significant advances in the way infants process their native language. We suggest that epigenetic processes—the continuing transaction between developing brain and experience—are the developmental mechanism that allows the emergence of these discontinuities in perceptual processing.

As noted earlier, there is considerable evidence that infants can use probabilistic information to learn about the characteristics of their native language. However, there are many functionally important tasks to be mastered for the learning of which sensitivity to probabilistic information is insufficient. For example, in order to properly segment words from ongoing speech, infants need not only to be sensitive to language-specific phonetic detail, phonotactics, and stress patterns, they must also be able to coordinate the simultaneous use of these different sources of information. The ability to coordinate two or more sources of information emerges for the first time at approximately 9 or 10 months of age (Diamond et al 1994, Lalonde & Werker 1995). This ability is manifest in speech processing in the ability to coordinate both rhythmical and segmental information simultaneously in the detection and establishment of a "unit" (Morgan & Saffran 1995). Prior to 9 months of age, infants can use a regular SW stress pattern to pull out bisyllabic units, or they can use a regular ordering of the same two syllables to pull out such a unit (Goodsit et al 1993), but it is not until 9 or 10 months of age that they can combine both sources of information and learn to listen for a particular sequence of syllables following a particular stress pattern

With a newly found general coordinative ability, one would expect many language-specific effects to appear simultaneously. And they do. They appear in phonetic perception, in use of phonotactic detail, in preference for native-language stress patterns, and in emerging language-specificity in babbling. The emergence of such an ability can also help explain many other findings in infant speech perception. For example, the full decline in non-native consonant and vowel discrimination seen by the end of the first year of life may rest on the ability to coordinate two sources of information (Lalonde & Werker 1995). In this case, we suggest that the coordination involves information about phonetic detail with position in a word (see Pegg & Werker 1997; see also Jusczyk 1997).

What is left unexplained is which mechanism allows for the appearance of this coordinative ability by 9 to 10 months of age. It is possible, as has been shown by some computational modeling studies (MacWhinney 1998), that self-organizing systems can "jump" to new levels of analysis by the simple accumulation of probabilistic and frequency information. Although we acknowledge this is possible, we find it unlikely. Our skepticism stems primarily from the fact that so many different abilities—both within and outside the domain of speech perception— show the sudden emergence of a coordinative ability at 9 or 10 months of age. For this reason we suggest that the emergence of the ability to coordinate two sources of information rests on some common underlying change in use of information. Specifically, we propose an epigenetic advance in brain development that, once present, allows for coordinative use of information in many different domains. In earlier work, we explored the possibility that the development of prefrontal cortex and its connections might be the advance that allows the functional reorganization in phonetic perception (Diamond et al 1994).

We are not yet able to account for the decline in phonetic detail seen in the initial stages of mapping words to meaning. Yet, in the spirit of the discussion above, we propose another functional reorganization that rests on another discontinuity in development. In the initial stages of constructing a lexicon based on full referential understanding, the child may establish a new level of representation. There may be limits to the amount of information initially encoded in this new level of representation because of additional computational demands of linking sound to objects, or the limits may come from the very act of building a new representation. Although further work is needed to understand this new process, we are confident that there is a discontinuity at this point in development that requires explanation.

SUMMARY

We briefly reviewed a number of existing models that attempt to account for the initial state and age-related changes in speech perception performance during infancy. We have shown that although each explains some aspects of development, none is comprehensive. As well, although different mechanisms are proposed by each, in some ways the differences are potentially complimentary rather than contradictory. We sketched an outline of an integrated, epigenetic model to explain both the language-general sensitivities present in the newborn and the age-related changes that occur during infancy. We, like the others, have been less successful in explaining the relationship between age-related changes in speech processing and the decline in phonetic detail used in the early stages of word-learning. At one level, we can describe each as a functional reorganization in which the child drops detail as she moves on to an increasingly complex task. But at another level, it may be necessary to begin anew when one moves from processing of speech as an acoustic form to processing speech as a unit of meaning. We argue that with the establishment of a new level of representation, a discontinuity is introduced. But we predict that the language-specific processing evident by the end of the first year of life provides the child with a repertoire of sensitivities—perhaps at a different level of representation—to draw on as he fills in that representation required to map sound to communicative intent (Hockett 1954).

Together, the findings reviewed in this chapter show a remarkable preparedness for speech perception and production in the human neonate. This boost from nature enables the infant to attend selectively and efficiently to information in the speech stream and to proceed on to the rapid elaboration of perceptual and productive knowledge of the phonological and syntactic properties of the native language. By, or shortly after, a child's first birthday, the knowledge of acceptable sound and grammatical patterning in the native language is present. The task for the next year of life is to construct a second-order system to effortlessly and efficiently use the medium of speech to map on to meaning.

ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada (OGP0001103 to JF Werker and OGP000179 to RC Tees). We also thank M Bhatnagar, J Pegg, and C Stager for their assistance and comment on the text and figures.

Literature Cited

Aslin RN 1993. Segmentation of fluent speech into words: learning models and the role of maternal input. See de Boysson-Bardies et al 1993, pp. 305–16

Aslin RN, Pisoni DB, Hennessy BL, Perey AJ. 1981. Discrimination of voice onset time by human infants: new findings and implications for the effects of early experience. *Child Dev.* 52:1135–45

Best CT. 1995. Learning to perceive the sound patterns of English. In *Advances in Infancy Research*, ed. C Rovee-Collier, LP Lipsitt, pp. 217–304. Norwood, NJ: Ablex

Best CT, McRoberts GW, Sithole NM. 1988. Examination of the perceptual re-organization for speech contrasts: Zulu click discrimination by English-speaking adults and infants. *J. Exp. Psychol. Hum. Percept. Perform.* 14:345–60

Bosch L, Sebastian-Galles N. 1997. Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition* 65:33–69

Chomsky N. 1975. *Reflections on Language*. New York: Random House

Christophe A, Dupoux E, Bertoncini J, Mehler J. 1994. Do infants perceive word boundaries? An empirical approach to the bootstrapping problem for lexical acquisition. *J. Acoust. Soc. Am.* 95:1570–80

Christophe A, Guasti T, Nespor M, Dupoux E, Van Ooyen B. 1997. Reflections on phonological bootstrapping: its role for lexical and syntactic acquisition. *Lang. Cogn. Proc.* 12:585–612

Cowan WM, Fawcett JW, O'Leary DDM, Stanfield BB. 1984. Regressive events in neurogenesis. *Science* 225:1258–65

Cutler A, Carter DM. 1987. The predominance of strong initial syllables in the English vocabulary. *Comput. Speech Lang.* 2:133–42

Cutler A, Norris D. 1988. The role of strong syllables in segmentation for lexical access. *J. Exp. Psychol. Hum. Percept. Perform.* 14:113–21

de Boysson-Bardies B, Halle P, Sagart L, Durand C. 1989. A cross-linguistic investigation of vowel formants in babbling. *J. Child Lang.* 16:1–17

de Boysson-Bardies B, Jusczyk P, McNeilage

P, Morton J, eds. 1993. *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*. Norwell, MA: Kluwer

de Boysson-Bardies B, Vihman MM. 1991. Adaptation to language: Evidence from babbling and first words in four languages. *Language* 67:297–319

Demuth K. 1996. The prosodic structure of early words. See Morgan & Demuth 1996, pp. 171–84

Desjardins R, Werker JF. 1996. Four month old infants are influenced by visible speech. *Infant Behav. Dev.* 19:762

Desjardins RN, Rogers J, Werker JF. 1997. An exploration of why preschoolers perform differently than do adults in audiovisual speech perception tasks. *J. Exp. Child Psychol.* 66:85–110

Diamond A, Werker JF, Lalonde CE. 1994. Toward understanding commonalties in the development of object search, detour navigation, categorization, and speech perception. See Fischer & Dawson 1994, pp. 380–426

Eilers R, Oller DK. 1976. The role of speech discriminations in developmental sound substitutions. *J. Child Lang.* 3:319–29

Eimas PD. 1974. Auditory and linguistic processing of cues for place of articulation by infants. *Percept. Psychophysiol.* 16:513–21

Eimas PD, Siqueland ER, Jusczyk PW, Vigorito J. 1971. Speech perception in infants. *Science* 171:303–6

Ferguson CA, Farwell CB. 1975. Words and sounds in early language acquisition. *Language* 51:419–39

Fischer KW, Dawson G, eds. 1994. *Human Behavior and the Developing Brain*. New York: Guilford

Fowler CA, Rosenblum LD. 1991. Perception of the phonetic gesture. In *Modularity and the Motor Theory*, ed. IG Mattingly, M Studdert-Kennedy, pp. 33–59. Hillsdale, NJ: Erlbaum

Gerken LA. 1994. Child phonology: past research, present questions, future directions. In *Handbook of Psycholinguistics*, ed. MA Gernsbacher, pp. 781–820. New York: Academic

Gerken LA, Jusczyk PW, Mandel DR. 1994. When prosody fails to cue syntactic structure: nine-month-olds' sensitivity to phonological vs. syntactic phrases. *Cognition* 51:237–65

Gleitman L, Gleitman H, Landau B, Wanner E. 1988. Where the learning begins: initial representations for language learning. In *The Cambridge Linguistic Survey*, ed. F Newmeyer, 3:150–93. Cambridge, MA: Harvard Univ. Press

Goodsit JV, Morgan JL, Kuhl PK. 1993. Perceptual strategies in prelingual speech segmentation. *J. Child Lang.* 20:229–52

Gottlieb G. 1991. Experiential canalization of behavioral development: theory. *Dev. Psychol.* 27:4–13

Green KP. 1998. The use of auditory and visual information in phonetic perception. In *Speech Reading by Humans and Machines*, ed. D Stork, ME Henneck. Berlin: Springer-Verlag. In press

Greenough WT. 1986. What's special about development? Thoughts on the bases of experience-sensitive plasticity. See Greenough & Juraska 1986, pp. 387–407

Greenough WT, Juraska JM. 1986. *Developmental Neuropsychobiology*. Orlando, FL: Academic

Grieser D, Kuhl PK. 1989. Categorization of speech by infants: support for speech-sound prototypes. *Dev. Psychol.* 25: 577–88

Halle PA, de Boysson-Bardies B. 1996. The format of representation of recognized words in infants' early receptive lexicon. *Infant Behav. Dev.* 19:463–81

Hebb DO. 1949. *The Organization of Behavior*. New York: Wiley

Hirsh-Pasek K, Kemler Nelson DG, Jusczyk PW, Cassidy KW, Druss B, Kennedy L. 1987. Clauses are perceptual for young infants. *Cognition* 26:269–86

Hockett, CA. 1954. Two models of grammatical description. *Word* 10:210–31

Jakobson R. 1941. *Child Language, Aphasia and Phonological Universals*. Transl. A Keller, 1968. The Hague, Netherlands: Mouton

Jusczyk PW. 1997. *The Discovery of Spoken Language*. Cambridge MA: MIT Press

Jusczyk PW, Aslin RN. 1995. Infants' detection of the sound patterns of words in fluent speech. *Cogn. Psychol.* 29:1–23

Jusczyk PW, Bertoncini J. 1988. Viewing the development of speech perception as an innately guided learning process. *Lang. Speech* 31:217–35

Jusczyk PW, Cutler A, Redanz N. 1993. Preference for the predominant stress patterns of English words. *Child Dev.* 64:675–87

Jusczyk PW, Hohne EA. 1997. Infants' memory for spoken words. *Science* 277: 1984–86

Kelly M. 1992. Using sound to solve syntactic problems: the role of phonology in grammatical category assignments. *Psychol. Rev.* 9:349

Kent RD, Murray AD. 1982. Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *J. Acoust. Soc. Am.* 72:353–65

Kolb B, Fantie B. 1989. Development of

child's brain and behavior. In *Handbook of Clinical Child Neuropsychology,* ed. CR Reynolds, E Fletcher-Janzen, pp. 17–39. New York: Plenum

Kuhl PK. 1979. Speech perception in early infancy: perceptual constancy for spectrally dissimilar vowel categories. *J. Acoust. Soc. Am.* 66:1168–79

Kuhl PK. 1988. Auditory perception and the evolution of speech. *Hum. Evol.* 3(1–2): 19–43

Kuhl PK. 1993. Innate predispositions and the effects of experience in speech perception: the native language magnet theory. See de Boysson-Bardies et al 1993, pp. 259–74

Kuhl PK. 1998. Effects of language experience on speech perception. *J. Acoust. Soc. Am.* 103(5):2931

Kuhl PK, Meltzoff AN. 1982. The bimodal perception of speech in infancy. *Science* 218:1138–41

Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B. 1992. Linguistic experiences alters phonetic perception in infants by 6 months of age. *Science* 255:606–8

Lalonde C, Werker JF. 1995. Cognitive influences on cross-language speech perception in infancy. *Infant Behav. Dev.* 18:459–76

Levitt AG, Utman J, Aydelott J. 1992. From babbling towards the systems of English and French: a longitudinal two-case study. *J. Child Lang.* 19:19–49

Liberman AM, Mattingly IG. 1985. The motor theory of speech perception revised. *Cognition* 21:1–36

Lindblom B. 1992. Phonological units as adaptive emergents of lexical development. In *Phonological Development: Models, Research, Implications*, ed. CA Ferguson, L Menn, C Stoel Gammon, pp. 131–63. Timonium, MD: York

Locke JL. 1983. *Phonological Acquisition and Change*. New York: Academic

MacKain KS, Best CT, Strange W. 1981. Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Appl. Psycholinguist.* 2:369–90

MacWhinney B. 1998. Models of the emergence of language. *Annu. Rev. Psychol.* 49: 199–227

Mandel DR, Jusczyk PW, Pisoni DB. 1995. Infants' recognition of the sound patterns of their own names. *Psychol. Sci.* 6:315–18

Massaro DW, Cohen MM, Gesi A, Heredia R, Tsuzaki M. 1993. Bimodal speech perception: an examination across languages. *J. Phon.* 21:445–78

Mazuka R. 1996. Can a grammatical parameter be set before the first word? Prosodic contributions to early setting of a grammatical parameter. See Morgan & Demuth 1996, pp. 313–30

McGurk H, MacDonald J. 1976. Hearing lips and seeing voices. *Nature* 264:746–48

Mehler J, Christophe A. 1994. Language in the infant's mind. *Philos. Trans. R. Soc. London Ser. B* 346:13–20

Mehler J, Juczyk PW, Lambertz G, Halsted N, Bertoncini J, et al. 1988. A precursor of language acquisition in young infants. *Cognition* 29:144–78

Menyuk P, Menn L, Silber R. 1979. Early strategies for the perception and production of words and sounds. In *Studies in Language "Acquisition,"* ed. P Fletcher, M Garman, pp. 198–222. Cambridge, MA: Cambridge Univ. Press

Mills DL, Coffey-Corina SA, Neville HJ. 1994. Variability in cerebral organization during primary language acquisition. See Fischer & Dawson 1994, pp. 427–55

Moon C, Cooper RP, Fifer WP. 1993. Two-day old infants prefer native language. *Infant Behav. Dev.* 16:495–500

Morgan JL, Demuth K, eds. 1996. *Signal to Syntax: Bootstrapping from Speech to Grammer in Early Acquisition*. Mahwah, NJ: Erlbaum

Morgan JL, Saffran JR. 1995. Emerging integration of sequential and suprasegmental information in prelingual speech segmentation. *Child Dev.* 66:911–36

Nazzi T, Bertoncini J, Mehler J. 1998. Language discrimination by newborns: towards an understanding of the role of rhythm. *J. Exp. Psychol. Hum. Percept. Perform.* In press

Nespor M, Vogel I. 1986. *Prosodic Phonology*. Dordrecht, Netherlands: Foris

Ojemann GA. 1991. Cortical organization of language. *J. Neurosci.* 11:2281–87

Oller DK. 1980. The emergence of speech sounds in infancy. In *Child Phonology*, ed. GH Yeni-Komshian, JF Kavanagh, CA Ferguson, pp. 92–112. New York: Academic

Patterson ML, Werker JF. 1998. Matching phonetic information in lips and voice by 40-month-old infants. Poster presented at 11th Bienn. Int. Conf. Infant Stud., Atlanta, GA

Pegg JE, Werker JF. 1997. Adult and infant perception of two English phones. *J. Acoust. Soc. Am.* 102:3742–53

Pisoni DB, Aslin RN, Perey AJ, Hennessy BL. 1982. Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *J. Exp. Psychol. Hum. Percept. Perform.* 8: 297–314

Polka L, Bohn O-S. 1996. A cross-language comparison of vowel perception in English-learning and German-learning infants. *J. Acoust. Soc. Am.* 100:577–92

Polka L, Werker JF. 1994. Developmental changes in perception of non-native vowel contrasts. *J. Exp. Psychol. Hum. Percept. Perform.* 20:421–35

Rosenblum LD, Schmuckler MA, Johnson JA. 1997. The McGurk effect in infants. *Percept. Psychophysiol.* 59:347–57

Saffran JR, Aslin RN, Newport EL. 1996. Statistical learning by 8-month old infants. *Science* 274:1926–28

Sams M, Aulanko R, Hamalainen M, Mari R, Lounsamaa OV. 1991. Seeing speech: visual information from lip movements modifies activity in the human auditory cortex. *Neurosci. Lett.* 127:141–45

Shi R, Morgan JL, Allopenna P. 1998a. Phonological and acoustic bases for earliest grammatical category assignment. A cross-linguistic persepctive. *J. Child Lang.* 25:169–201

Shi R, Werker JF, Morgan JL. 1998b. Young infants' perception of lexical and functional categories. Poster presented at 11th Bienn. Int. Conf. Infant Stud., Atlanta, GA

Siva N, Stevens EB, Kuhl PK, Meltzoff AN. 1995. A comparison between cerebral-palsied and normal adults in the perceptions of auditory-visual illusions. *J. Acoust. Soc. Am.* 98:2983

Stager CL, Werker JF. 1997. Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature* 388:381–82

Strange W, Jenkins JJ. 1978. Role of linguistic experience in the perception of speech. In *Perception and Experience*, ed. RD Walk, HL Pick, pp. 125–69. New York: Plenum

Streeter LA. 1976. Language perception of 2-month old infants shows effects of both innate mechanisms and experience. *Nature* 259:39–41

Tees RC. 1986. Experience and visual development: behavioral evidence. See Greenough & Juraska 1986, pp. 317–61

Tees RC, Werker JF. 1984. Perceptual flexibility: maintenance or recovery of the ability to discriminate non-native speech sounds. *Can. J. Psychol.* 38:579–90

Trehub SE. 1973. Infants' sensitivity to vowel and tonal contrasts. *Dev. Psychol.* 9:91–96

Trehub SE. 1976. The discrimination of foreign speech contrasts by infants and adults. *Child Dev.* 47:466–72

Vihman MM, Miller R. 1988. *Word*s and babble at the threshold of language acquisition. In *The Emergent Lexicon*, ed. MD Smith, JL Locke, pp. 151–83. New York: Academic

Werker JF. 1995. Exploring developmental changes in cross-language speech perception. In *An Invitation to Cognitive Science.* Part I: *Language,* ed. L Gleitman, M Liberman, pp. 87–106. Cambridge, MA: MIT Press

Werker JF, Cohen LB, Lloyd V, Stager C, Cassosola M. 1998a. Acquisition of word-object associations by 14-month-old infants. *Dev. Psychol.* In press

Werker JF, Gilbert JHV, Humphrey K, Tees RC. 1981. Developmental aspects of cross-language speech perception. *Child Dev.* 52:349–55

Werker JF, Lalonde CE. 1988. Cross-language speech perception: initial capabilities and developmental change. *Dev. Psychol.* 24:672–83

Werker JF, McGurk HE, Frost P. 1992. La langue et les lévres: cross-language influences on bimodal speech perception. *Can. J. Psychol.* 46:551–68

Werker JF, Shi R, Desjardins R, Pegg JE, Polka L, et al. 1998b. Three methods for testing infant speech perception. In *Perceptual Development: Visual, Auditory, and Speech Perception in Infancy*, ed. AM Slater. London: Univ. Coll. London Press. In press

Werker JF, Stager C. 1998. Developmental changes in infant speech perception and early word learning. Is there a link? In *Papers in Laboratory Phonology,* ed. M Broe, J Peirre-Humbert. Cambridge, UK: Cambridge Univ. Press. In press

Werker JF, Tees RC. 1983. Developmental changes across childhood in the perception of non-native speech sounds. *Can. J. Psychol.* 37(2):278–86

Werker JF, Tees RC. 1984a. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* 7:49–63

Werker JF, Tees RC. 1984b. Phonemic and phonetic factors in adult cross-language speech perception. *J. Acoust. Soc. Am.* 75:1866–78

Werker JF, Tees RC. 1992. The organization and reorganization of human speech perception. *Annu. Rev. Neurosci.* 15:377–402

Wexler K, Culicover P. 1980. *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press

Whalen DH, Levitt A, Wang Q. 1991. International differences the reduplicative babbling of French- and English-learning infants. *J. Child Lang.* 18:501–6