Learning and representation in speech and language

Patricia K Kuhl

University of Washington, Seattle, USA

Infants learn language with remarkable speed. By the end of their second year they speak in sentences with an 'accent' typical of a native speaker. How does an individual acquire a specific language? While acknowledging the biological preparation for language, this review focuses on the effects of early language experience on infants' perceptual and perceptual–motor systems. The data show that by the time infants begin to master the higher levels of language — sound–meaning correspondences, contrastive phonology, and grammatical rules — their perceptual and perceptual–motor systems are already tuned to a specific language. The consequences of this are described in a developmental theory at the phonetic level that holds promise for higher levels of language.

Current Opinion in Neurobiology 1994, **4**:812–822

Introduction

Speech exhibits a 'language-general' pattern that becomes 'language-specific' by the end of the first year of life [1••–3••]. This transition is one of the most intriguing problems in language acquisition. What accounts for the transition?

Early in life, infants discern differences between all the phonetic units used in the world's languages, and demonstrate exquisite sensitivity to acoustic change in the region of the boundaries between phonetic categories [4]. Because non-human animals perceive the same discontinuities in speech, infants' initial abilities have been attributed to more basic auditory processing mechanisms rather than ones that evolved specifically for language [5]. By 12 months of age, infants fail to discriminate the foreign-language contrasts they once distinguished [3••]. As adults, our abilities are greatly reduced; we often find it difficult to perceive differences between sounds not used to distinguish words in our native language [3••]. Adult native speakers of Japanese, for example, have great difficulty discriminating American English /r/ and /l/ [6••], although Japanese infants do make this distinction [7].

Speech production follows a similar pattern. Regardless of culture, all infants progress through a set of universal stages during the first year [8]. By the end of the first year, however, the utterances of infants reared in different countries begin to diverge, reflecting the ambient language [9••,10]. In adulthood, the speech motor patterns that contribute to one's 'accent' are very difficult to alter [11••].

This review focuses on the role of early linguistic experience in bringing about this language-general to language-specific change in speech perception and production. The thesis is that linguistic experience results in an interesting kind of learning. Given linguistic input, the perceptual and perceptual–motor systems underlying speech show self-organization accompanied by a loss in flexibility. The vehicle for change is argued to be representations stored in memory that capture the regularities of a specific language. These representations alter the perceptual and perceptual–motor skills of infants. The consequence of this is that in the absence of formal language understanding or use, infants' perceptual and perceptual–motor systems are strongly biased towards the characteristics of the ambient language. A model is described at the phonetic level that shows how this structure aids the acquisition of phonology and accounts for the transition. Implications for theories of language learning and several problems in cognitive neuroscience are discussed.

Language experience alters perception

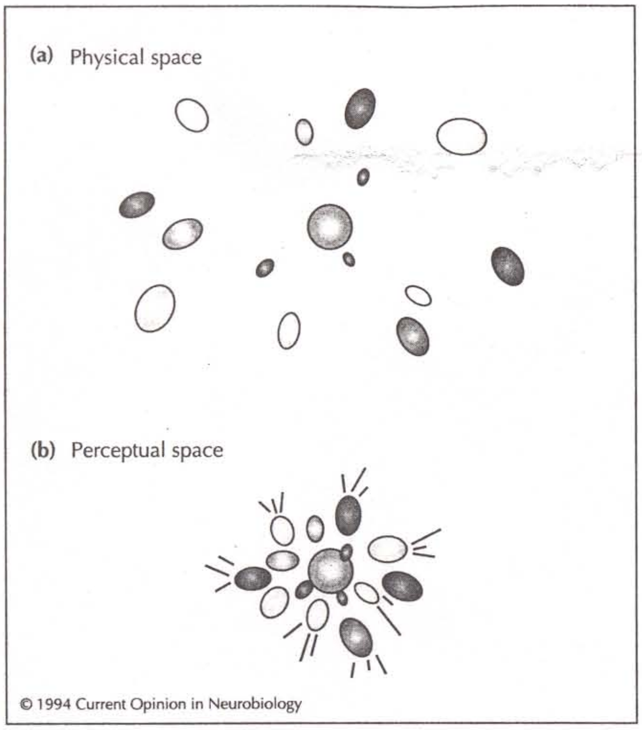
Recent work in my laboratory has produced an effect that helps explain how language experience affects speech perception and production. The effect shows that linguistic experience alters the perceived distances between speech stimuli; in effect, experience 'warps' the perceptual space underlying speech. The end result is that perceptual categories are formed, ones that mir-

---

Abbreviations

ERP—event-related potential; **F1**—first formant; **F2**—second formant; **F3**—third formant; **fMRI**—functional MRI; **MDS**—multidimensional scaling; **MEG**—magnetoencephalography; **MMN**—mismatched negativity; **MRI**—magnetic resonance imaging; **NLM**—native language magnet; **PET**—positron emission tomography.

© Current Biology Ltd ISSN 0959-4388



(a) Physical space

(b) Perceptual space

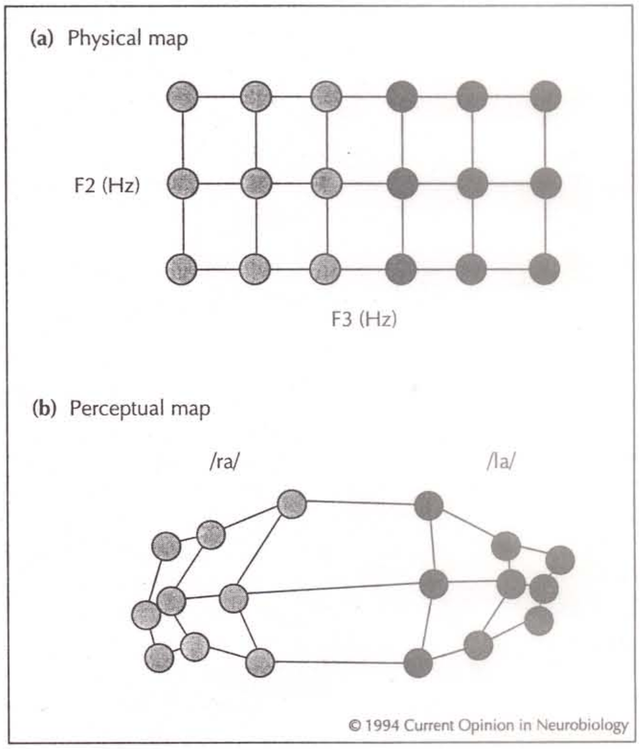
© 1994 Current Opinion in Neurobiology

Fig. 1. The perceptual magnet effect. **(a)** When a variety of sounds in a category surround the category prototype, **(b)** they are perceptually drawn toward the prototype. The prototype appears to function like a magnet or attractor for other stimuli in the category.

ror the phonological categories of the ambient language. The phenomenon — called the 'perceptual magnet effect' — is demonstrated in experiments using phonetic 'prototypes', the best or most representative instances of phonetic categories. The experiments show that the best instances of phonetic categories function like 'perceptual magnets' for other sounds in the category (Fig. 1) [12]. When listeners hear a phonetic prototype and attempt to discriminate it from sounds that surround it in acoustic space (Fig. 1a), the prototype displays an attractor effect on the surrounding sounds. It perceptually pulls other members of the category toward it, making it difficult to hear differences between the prototype and surrounding stimuli (Fig. 1b). Poor instances from the category (non-prototypes) do not function in this way. Several experimental tasks produce this result [13••,14,15]. Other studies show that phonetic prototypes are language specific [16] and context specific [17••,18•].

Developmental studies revealed that the perceptual magnet effect is exhibited by 6 month old infants for the sounds of their native language [12]. Moreover, cross-language experiments demonstrated that the magnet effect is the product of linguistic experience [19]. Kuhl and colleagues [19] conducted a cross-language experiment with adults and infants in America and Sweden. Subjects from both countries were tested with two vowel prototypes, an American English vowel prototype, /i/ (as in 'peep'), and a Swedish vowel prototype, /y/ (as in 'fye'). Adults from both cultures perceived the foreign vowel as a non-prototype. The results demonstrated that the perceptual magnet effect in 6 month old infants is affected by exposure to a particular language. American infants demonstrated the magnet effect only for the American English /i/; they treated the Swedish /y/ like a non-prototype. Swedish infants showed the opposite pattern, demonstrating the magnet effect for the Swedish /y/, and treating the American English /i/ as a non-prototype.

Work by Polka and Werker [20••] both supports and extends these findings. They tested Canadian English infants in a discrimination task involving two German phonetic contrasts. Their results confirm the presence of the magnet effect in 6 month old infants and show a decline in the discrimination of foreign-language vowel contrasts between 4 months and 6 months, earlier for vowels than for consonants but following the same pattern.



(a) Physical map

(b) Perceptual map

© 1994 Current Opinion in Neurobiology

Fig. 2. Physical (acoustic) versus perceptual maps. **(a)** Consonant tokens of /r/ (light gray dots) and /l/ (dark gray dots) were generated to be equally distant from one another in acoustic space. **(b)** However, when listeners perceive them, distance is distorted. Perceptual space is shrunk near the best instances of /r/ and /l/ and stretched at the boundary between the two. The resulting perceptual map differs for speakers of different languages. Abbreviations: F2, second formant; F3, third formant.

What can we say about the mechanics and basis of the magnet effect? Recent studies using multidimensional scaling (MDS) techniques provide clues to the way in which the magnet effect distorts perception (Fig. 2) ([21••,22]; K Davis, PK Kuhl, J Acoust Soc Am 1994, 95:2976). The studies show that the best instances of phonetic categories yield increased perceptual clustering while the category's worst instances yield reduced perceptual clustering (see [23] for related work on cognitive categories outside the domain of speech). Iverson and Kuhl [21••] computer synthesized a set of syllables beginning with /r/ and /l/. The syllables were created by varying critical acoustic components of the signals, the second and third formant (F2 and F3, respectively) frequencies. (Formants are bands of highly concentrated acoustic energy; the formant frequency specifies the center of the band.) The syllables were spaced at equal intervals in a two-dimensional grid (Fig. 2a). Listeners identified each syllable as beginning with either /r/ or /l/, rated its category goodness, and estimated the perceived similarity for all possible pairs of stimuli using a scale from '1' (very dissimilar) to '7' (very similar). Similarity ratings were scaled using MDS techniques.

The results revealed that perceived distances differed from real physical distances; in other words, perception distorted physical space. The physical (acoustic) differences between pairs of stimuli were equal (Fig. 2a); however, perceived distance was 'warped' (Fig. 2b).

The perceptual space around the best /r/ and the best /l/ was greatly reduced — shrunk — while the space near the boundary between the two categories was expanded — stretched. The results suggest that linguistic experience results in the formation of 'perceptual maps' specifying the perceived distances between stimuli. These maps increase internal category cohesion while maximizing the distinction between categories. The critical point for the theory is the hypothesis that the map is defined differently for speakers of different languages [22]. Japanese adults tested with the same /r/ and /l/ stimuli will very likely show a different perceptual map, one without perceptual clusters around the American /r/ and /l/ prototypes.

A theory of speech development

These findings have been incorporated in a three-step theory of speech development, called the native language magnet (NLM) theory [1••,2••,24]. NLM describes infants' initial state as well as changes brought about by experience with language. It explains how infants' developing native-language speech representations alter both speech perception and production. The example developed here is for vowels, although the same principles apply to consonant perception and higher order units such as words.

Phase 1 describes infants' initial abilities (Fig. 3). At birth, infants partition the sound stream into gross categories separated by natural auditory boundaries. In Fig. 3, these perceptual boundaries divide the vowel space, separating the vowels of all languages. According to NLM, these boundaries are inherent in auditory processing; infants' abilities at this stage do not depend on specific language experience. The boundaries initially structure perception in a phonetically relevant way. However, they are not due to a 'language module' but to more basic auditory perceptual processing mechanisms. This notion is buttressed by the fact that these same perceptual boundary phenomena are exhibited in the same places in acoustic space by non-human animals, suggesting some interesting hypotheses about the evolution of speech [5].

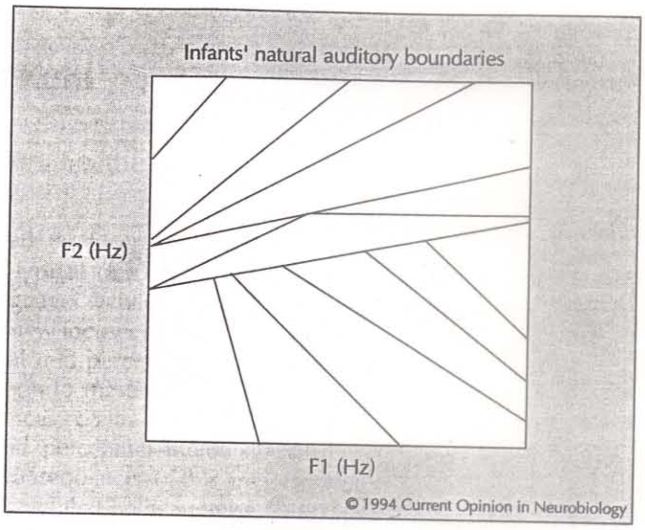


Fig. 3. At birth, infants perceptually partition the acoustic space underlying phonetic distinctions in a universal way. They are capable of discriminating all phonetically relevant differences in the world's languages. Abbreviations: F1, first formant; F2, second formant.

Phase 2 describes perception at 6 months of age for infants reared in three very different language environments, Swedish, English and Japanese (Fig. 4). By this age, infants show more than the innate boundaries shown in Phase 1. By six months, infants have heard hundreds of thousands of instances of particular vowels (K Gustafson, unpublished data). According to NLM, infants represent this information in memory in some form. Moreover, the distributional properties of vowels heard by infants raised in Sweden, America, and Japan differ. As shown in Fig. 4, their stored representations also differ, reflecting these distributional differences. In each case, linguistic experience has produced stored representations that mirror the vowel system of the ambient language. Language-specific magnet effects, produced by the stored representations, are now exhibited by infants.

Phase 3 shows how magnet effects recursively alter the initial state of speech perception. Magnet effects cause certain acoustic differences to be minimized (those near the magnet attractors) while others are maximized (those near the boundaries between two magnets). The consequence is that some of the boundaries that initially divided the space 'disappear' as the perceptual space is reconfigured to incorporate a language's particular magnet placement. This is schematically illustrated in Fig. 5. In Phase 3, sensory perception has not changed, but higher order memory and representational systems have altered listeners' responses. Magnet effects functionally erase certain boundaries — those relevant to foreign but not native languages. At this stage, a perceptual space once characterized by simple boundaries has been replaced by a dynamically warped space dominated by magnets.

The important point for theory is that infants at six months of age have no awareness of phonemes or the fact that sound units are used contrastively in language to name things. Yet the infant's perceptual system has organized itself to reflect language-specific phonetic categories. At the next stage in linguistic development, when infants acquire word meanings by relating sounds to objects and events in the world, the language-specific mapping that has already occurred in their perceptual systems will greatly assist this process.

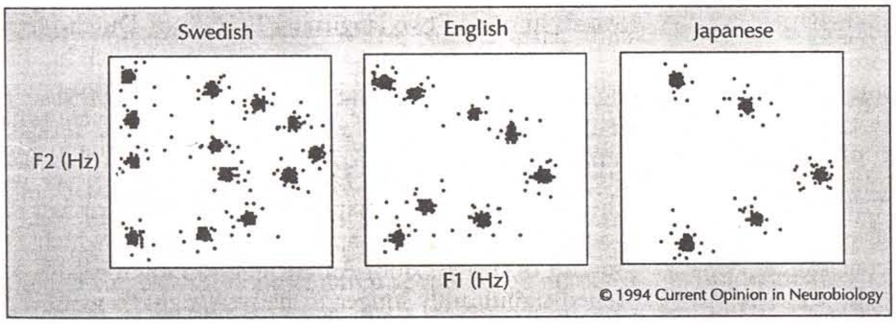
NLM theory offers an explanation for the transition in speech perception observed by Werker and Polka [3••]. A developing magnet pulls sounds that were once discernible towards it, making them less discernible. Magnet effects should therefore developmentally precede changes in infants' perception of foreign-language contrasts. Preliminary data indicate that they do [20••]. The magnet effect also helps account for the results of studies on the perception of sounds from a foreign language by adults [3••,6••,11••,25••]. For example, NLM theory may explain Japanese listeners' difficulty with American /r/ and /l/ by predicting that the magnet effect for their category prototype (which is neither American /r/ nor /l/) will attract both /r/ and /l/, making the two sounds difficult for native-speaking Japanese people to discriminate. Best [6••] has also presented ideas about the relative discriminability of foreign-language contrasts by examining the relationship of specific foreign sounds to native-language categories; these predictions are consistent with the view described here.

The perceptual–motor link

Infants learn to produce the sound patterns of language by listening to native speakers and imitating the sounds they hear. Speech motor control is a complex process, but by adulthood, we possess detailed information about the consequences of speech movements on sound [26••,27]. When do infants forge the perceptual–motor link? By 1 year of age language-specific patterns of speech production have begun to appear in infants' spontaneous utterances [9••,10]; by 30 months, subtle differences that differentiate sounds in two different languages are observed [28].

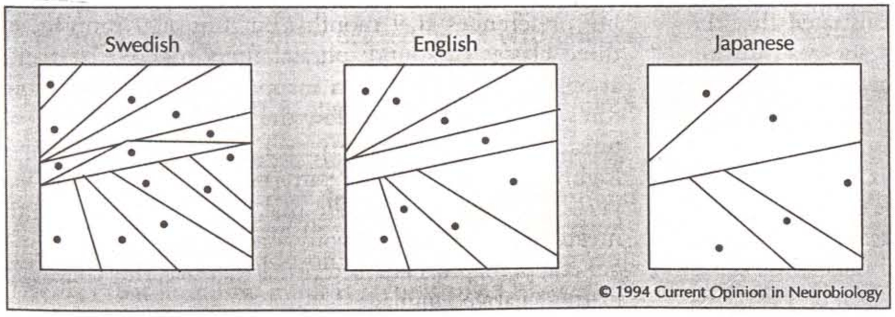
Recent studies, however, suggest that the link is in place much earlier if it is tapped by an appropriately sensitive measure. Kuhl and Meltzoff [29••] recorded infant utterances at 12, 16, and 20 weeks of age while the infants watched and listened to a video recording of a woman producing a vowel, either /a/, /i/, or /u/. Infants watched the video for 5 min on each of three consecutive days. Infants' utterances were analyzed both perceptually (phonetic transcription) and instrumentally (computerized spectrographic analysis). There was developmental change in infants' vowel productions between 12 and 20 weeks of age. The areas of vowel space occupied by infants' /a/, /i/, and /u/ vowels become progressively more tightly clustered at each age (Fig. 6).

Infants also imitated the vowels they heard. The total amount of exposure was only 15 min, yet this was sufficient to alter their vowel sounds. If 15 min of laboratory exposure to a vowel is sufficient to influence infants' vocalizations, then listening to ambient language for weeks would be expected to provide a powerful influence on infants' speech production. These data suggest that infants' stored representations of speech alter not only infant perception, but production as well, serving as targets



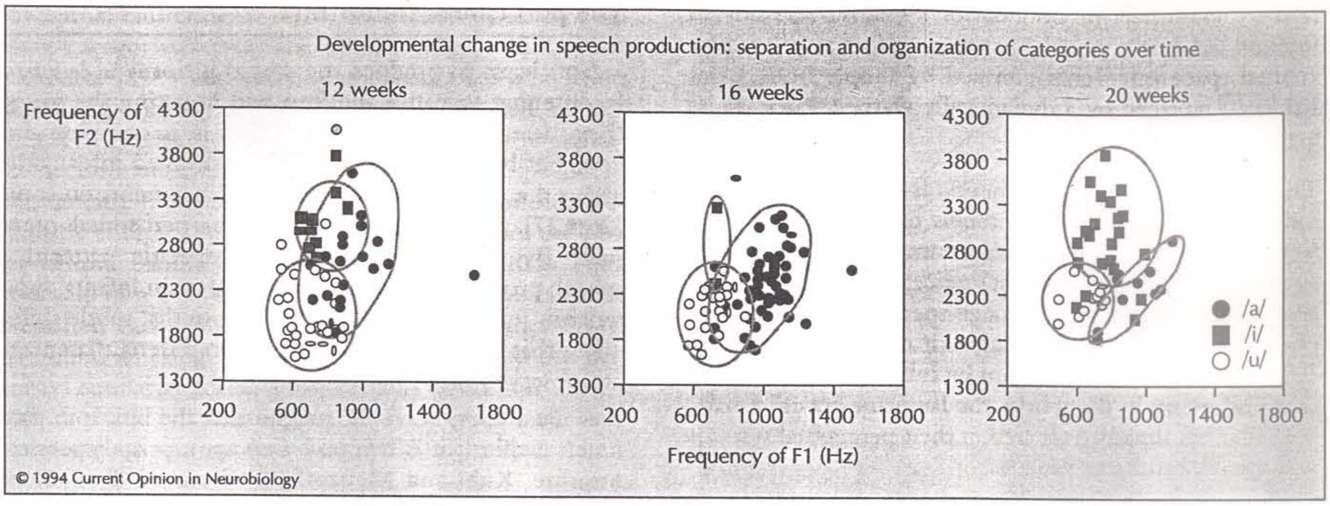
© 1994 Current Opinion in Neurobiology

Fig. 4. By 6 months of age, infants raised in different linguistic environments show an effect of language experience. Infants store incoming speech information in memory in some form. The resulting representations are language specific, and reflect the distributional properties of vowels in the three different languages. Abbreviations: F1, first formant; F2, second formant.



© 1994 Current Opinion in Neurobiology

Fig. 5. After language-specific magnet effects appear (shown by the dots), some of the natural boundaries that existed at birth 'disappear'. Infants now fail to discriminate foreign-language contrasts they once discriminated.



© 1994 Current Opinion in Neurobiology

Fig. 6. The location of /a/, /i/, and /u/ vowels produced by 12, 16, and 20 week old American infants. Infants' vowel productions show progressively tighter clustering in vowel space over the 8 week period and reflect differences between the three vowel categories seen in adults' productions. The curves were drawn by visual inspection to enclose 90% or more of infants' utterances. Abbreviations: F1, first formant; F2, second formant.

that guide motor production. Stored representations are thus viewed as the common cause for both the tighter clustering observed in infant vowel production and infant vowel perception (Fig. 7) [2••,29••].

This pattern of learning and self-organization [30], in which perceptual patterns stored in memory serve as guides for production, is strikingly similar to that seen in other domains involving auditory–perceptual learning, such as birdsong [31], and in visual–motor learning, such as gestural imitation [32]. In each of these cases, perceptual experience establishes a representation that guides sensory-motor learning. In the case of infants and speech, perception affects production in the earliest stages of language learning, reinforcing the idea that the perceptual–motor link is in place early in life [2••,29••,33,34••,35,36].

Avid learning at all levels

Learning commences prenatally with the more global, prosodic aspects of language [37]. Recent studies confirm earlier reports suggesting that by the time infants are born, exposure to sound *in utero* has resulted in a preference for native-language over foreign-language utterances [38••]. Previous work demonstrated that the mother's voice [39] and simple stories she read during the last trimester [40] are also recognized by infants at birth. Studies on the acoustics of speech and the intrauterine environment suggest that intense (>80 dB), low-frequency sounds (particularly <300 Hz, but as high as 1000 Hz with some attenuation) penetrate the womb [41]. This means that the prosodic patterns of speech, including voice pitch and the stress and intonation characteristics of a particular language and speaker, are transmitted to the fetus, while the sound patterns that allow phonetic units and words to be identified are greatly attenuated. (This can be compared to listening to speech through the wall of a room — a human voice can be identified, but words cannot be made out.)

Postnatally, infants' processing of the prosodic aspects of speech provides additional information about language-specific sound patterns. Jusczyk and colleagues (see [42]) have focused on infant learning of the sound patterns typical of native-language words, phrases and sentences. This work shows that between 6 and 9 months of age, infants develop listening preferences for sound patterns typical of the native language. Infants indicate preference by turning their heads to the right or left. In one study, infants could listen to lists of 15 two- and three-syllable words in one of two languages, English or Dutch [43••]. The two languages have similar prosodic patterns but distinct phonetic units and rules for combining those units (phonotactic rules). The words on these lists were unfamiliar (e.g. 'trustworthy' or 'jostle' in English and 'uitsteeksel' or 'oprecht' in Dutch). The basis of choice was therefore not previous experience with the words but the sounds they contained and the sequences in which the sounds occurred. Both American and Dutch infants listened significantly longer to native-language word lists. At 6 months of age, infants showed no listening preferences (see also [44]). Additional work shows listening preferences at 9 months, but not at 6 months, for three classes of sound: phrasal units marked by pauses at appropriate rather than inappropriate linguistic junctures [45]; words that follow the predominant stress pattern of the language [46••]; and non-words that contain highly frequent phonetic patterns in the native language [47]. These studies indicate that before infants learn the meanings of individual words or phrases, they recognize general perceptual characteristics that describe such units in their native language.

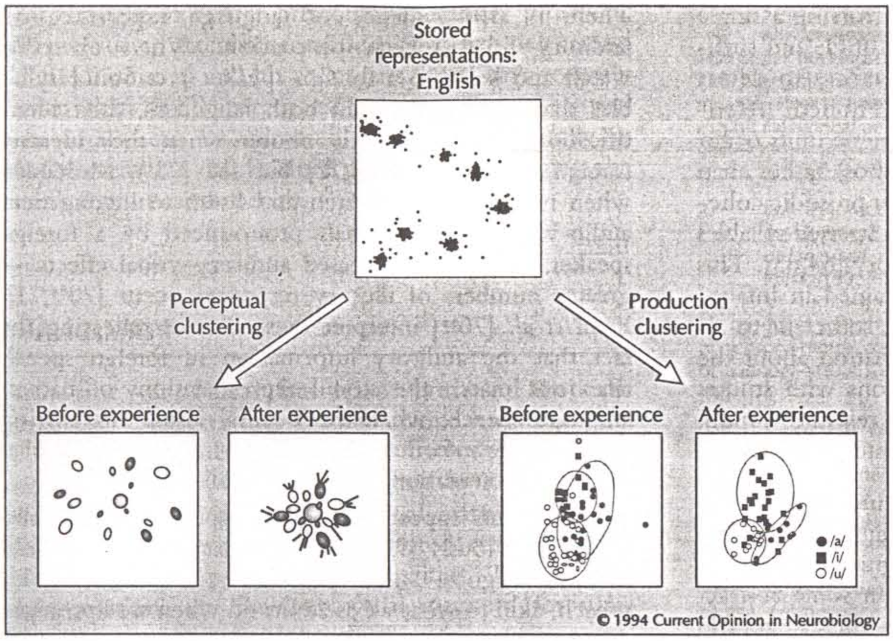


Fig. 7. Stored representations of native-language speech affect both speech perception, producing the perceptual clustering evidenced by the magnet effect, as well as speech production, producing the increased clustering seen in infants' vocalizations over time.

Language input to the child

Language input is essential for learning. How much and what kind of speech do infants hear?

Estimates indicate that a typical listening day for a 2 year old includes 20 000 – 40 000 words [48]. Speech addressed to infants (often called 'motherese' or 'parentese') is unique: It has a characteristic prosodic structure which includes a higher pitch, a slower tempo and exaggerated intonation contours; it is also syntactically and semantically simplified. Research supports the idea that this speaking style is near universal in the speech of caretakers around the world and that infants prefer it [49].

Significant units in speech are prosodically marked in infant-directed speech and these features aid infants' speech processing [50–52]. When parents teach their infants new words, object names are accompanied by a peak in prosodic cues [53,54]. New words are highlighted in other ways. They are placed in word-final position which helps encode the item in memory [50]. Language input affects the order in which certain classes of words are acquired by children. In a cross-language study, Gopnik and Choi [55••] show that although English-speaking children acquire large numbers of nouns, reflecting language input by American caretakers, Korean children at the same age have acquired larger numbers of verbs, reflecting the fact that in Korean, noun phrases are often deleted while verb phrases are emphasized.

The motherese pattern of speech, with its higher pitch and expanded intonation contours, is probably not necessary for learning. However, the context in which language is presented to the child, both its auditory and visual characteristics (greatly exaggerated facial expressions), fix infant attention on the talking caretaker [56]. The fact that linguistic input is accompanied by acoustic features that not only attract infant attention but mark its significant features probably helps infants learn.

Memory and representation

In speech learning, what kind of memory system is involved, what kind of information is retained, and in what form is it stored?

Kuhl and Meltzoff [2••] argue that the learning and memory involved in infants' perceptual and perceptual–motor speech learning is not conscious learning of specific facts or events. In terms of the kinds of distinctions discussed by modern cognitive- and neuroscientists [57,58], it could not be described as explicit, 'declarative' memory. Infants listening to ambient language learn unconsciously, automatically and without extrinsic reinforcement. The learning that results is difficult to undo: it involves relatively permanent changes in speech production (an accent) and in speech perception (perceiving phonetic distinctions). It is probably best thought of as non-declarative memory of some (as yet undefined) type. Recent data and theorizing by Pisoni and his colleagues (see [59••]) led to the suggestion that implicit memory systems are involved in processing speech information. Considerably more research will need to be done to uncover the types of memory systems involved in speech processing, especially those involved in the initial phases of development when the impact of linguistic experience is long-lasting and profound.

Exploration of infant memory and representation for speech information has only just begun. Newborns can retain a word in memory over a 24-hour delay [60]. Two to three month old infants can retain sufficient speech information over a 2 min delay interval to discriminate a change in the sequence of syllables involving a single phonetic feature of one of the syllables [61], and sufficient information about bisyllabic utterances to detect the presence of a new syllable [62]. Phonetic retention appears to be enhanced when phonetic units occur within the same prosodic unit (a sentence) rather than across a sentence juncture, indicating that prosodic cohesion may help linguistic processing [63]. Stressed syllables may be particularly prominent in memory [64,65]. This work suggests that the information retained in infants' speech representations is quite detailed, sufficient to allow infants to encode and store information about the structure of language input. Comparisons with studies on early memory for other biologically relevant stimuli, such as faces [32], are of considerable interest.

When information about speech is retained, what form does the representation take? Two possibilities have been discussed in the adult literature on cognitive categories [66]. The first assumes that people form some abstract version (a 'prototype') that characterizes the category as a whole. The second, 'exemplar-based' model of categorization, assumes that individual exemplars are stored and retrieved. Research by Miller and her colleagues (see [17••]) on adults shows that the location of best exemplars of a phonetic category shifts with changes in context such as the rate of speech, suggesting that speech representations are context-specific. Moreover, data collected by Pisoni and colleagues (see [59••]) on the effects of talker variability, suggest that adult listeners encode fine details about the voice of the talker who produced the utterance, and that listeners' subsequent recognition of speech information spoken by the same talker is improved. These data suggest either exemplar-based representations that contain context- and talker-specific instances or a number of prototype-based representations that are themselves context- or talker-specific.

Polymodal speech

Speech perception has classically been considered an auditory process. This belief has been modified by data showing that speech perception is strongly affected by the sight of a talker producing speech. One of the most compelling examples of the polymodal nature of speech is auditory–visual illusions that result when discrepant information is sent to two separate modalities. One such illusion can be demonstrated when auditory information for /b/ is combined with visual information for /g/ [67,68]. Perceivers report the phenomenal impression of an intermediate articulation (/da/, /tha/, or /za/) despite the fact that this information was not delivered to either sense modality. Recent data suggest the robustness of the effect by demonstrating it is maintained even when the cross-modal information cannot have derived from the same biological source, as when a male face is combined with a female voice [69].

There is also evidence of language experience on auditory–visual speech perception. When observers watch and listen to a foreign speaker pronounce syllables that are contained in both languages, their identifications are significantly poorer when they identify foreign speech ([70••,71,72] but see [73]). Moreover, when native speakers watch and listen to incongruent audio-visual speech signals pronounced by a foreign speaker, they show increased auditory–visual effects — greater numbers of illusory responses occur [70••,71]. Kuhl *et al.* [70••] interpret these data as reflecting the fact that the auditory information in foreign speech does not match the stored representations of native-language speech; when this occurs, visual information may be more informative. The data support the idea that speech representations are polymodally mapped.

Young infants appear to represent speech polymodally. It was previously demonstrated that 18–20 week-old infants recognize auditory–visual correspondences for speech, akin to what we as adults do when we lipread; in these studies, infants looked longer at a face pronouncing a vowel that matched the vowel sound they heard rather than a mismatched face [74]. These results have been extended to show that infants do not match speech sounds to faces when the auditory stimulus contains an isolated feature of speech that does not allow the vowel to be identified [75]. Infants respond to novel face–voice combinations that are articulatorally possible, but not to those that are articulatorally impossible [76]. Young infants demonstrate knowledge about both the auditory and visual information contained in speech, supporting the notion that their stored speech representations contain information of both kinds.

Neural correlates of speech processing

Various techniques of neuroscience (PET, MRI, fMRI and MEG) have not yet been applied to phonetic processing in infants, although adult studies are beginning to appear [77]. However, high-density event-related potentials (ERPs) have recently been used to study word processing in young infants [78]. These techniques have been applied to speech processing in two month olds [79••] and newborns [80].

In the study of two month olds, Dehaene–Lambertz and Dehaene [79••] presented infants with strings of syllables (/ba/ or /ga/), that were either identical or contained one deviant syllable. They observed two distinct peaks in electrical activity: peak 1, which occurred within 290 ms of the onset of the syllable, was insensitive to phonetic changes; peak 2, which reached its maximum about 390 ms after syllable onset, showed significant change when the deviant syllable was presented. Thus, a single instance of a deviant syllable was recognized in less than 400 ms in the infant brain.

Additional work in adults suggests that the mismatched negativity (MMN) response, an ERP component thought to reflect preattentive auditory processes, will provide an interesting measure of speech perception [15,81]. Aaltonen and his colleagues [82••] have recently shown that Kuhl's perceptual magnet effect measured behaviorally can be mirrored electrophysiologically in MMN measures. ERPs hold promise for mapping the brain's responses to speech.

Conclusions

In the first year of life, infants learn a great deal about the perceptual characteristics of their native language. Perceptual learning subsequently alters both the perception and production of speech. According to Kuhl's 'native language magnet' theory, perceptual learning early in life results in the formation of stored representations that capture native-language regularities. These stored representations act like 'perceptual magnets' for similar patterns of sound. Magnet effects distort perceptual space creating perceptual maps that specify the perceived distances between sounds. The map shrinks perceptual distances near a category's most typical instances and stretches perceptual distances between categories. Perceptual maps differ in adults who speak different languages. The magnet effects and the perceptual maps they yield also affect speech production. This sheds light on the course of development in speech production and helps explain why, as adults, we do not hear or produce foreign-language sounds very well. During the language-learning period, our perceptual maps are tuned to our native language. The data implicate non-declarative learning and memory of some type. Future work will be aimed at mapping the brain changes that accompany language learning using the techniques of modern neuroscience.

Acknowledgements

Financial support from the National Institutes of Health (DC00520, HD18286 and HD22514) and the Virginia Merrill Bloedel Hearing Research Center is gratefully acknowledged. I thank A Meltzoff for helpful comments and discussion on the manuscript, and E Stevens for assistance on all aspects of the research and preparation of the manuscript.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. •• Kuhl PK: **Innate predispositions and the effects of experience in speech perception: the native language magnet theory.** In *Developmental neurocognition: speech and face processing in the first year of life.* Edited by de Boysson-Bardies B, de Schonen S, Jusczyk P, McNeilage P, Morton J. Dordrecht, Netherlands: Kluwer Academic Publishers; 1993:259–274.

This paper develops the native language magnet theory in detail. It reviews comparative studies on categorical perception and the perceptual magnet effect in infants and non-human animals, showing that animals exhibit categorical perception but not the perceptual magnet effect. It discusses results in terms of biological predispositions for language and the effects of language experience.

2. •• Kuhl PK, Meltzoff AN: **Evolution, nativism, and learning in the development of language and speech.** In *The biological basis of language.* Edited by Gopnik M. New York: Oxford University Press; 1995:in press.

Results of studies on infants' perception of speech are related to studies on infant speech production. The authors discuss the development of sensory–motor connections for speech, and relate them to studies of auditory–visual speech perception, vocal imitation and gestural imitation.

3. •• Werker JF, Polka L: **The ontogeny and developmental significance of language-specific phonetic perception.** In *Developmental neurocognition: speech and face processing in the first year of life.* Edited by de Boysson-Bardies B, de Schonen S, Jusczyk P, McNeilage P, Morton J. Dordrecht, Netherlands: Kluwer Academic Publishers; 1993:275–288.

The authors review the substantial literature on infants' perception of foreign-language distinctions, showing that between 10 and 12 months of age, performance on foreign-language consonant contrasts declines. They discuss various linguistic and cognitive explanations for the decline.

4. Eimas PD, Miller JL, Jusczyk PW: **On infant speech perception and the acquisition of language.** In *Categorical perception: the groundwork of cognition.* Edited by Harnad S. New York: Cambridge University Press; 1987:161–195.

5. Kuhl PK: **Perception, cognition, and the ontogenetic and phylogenetic emergence of human speech.** In *Plasticity of development.* Edited by Brauth SE, Hall WS, Dooling RJ. Cambridge, Massachusetts: MIT Press; 1991:73–106.

6. •• Best CT: **Emergence of language-specific constraints in perception of non-native speech: a window on early phonological development.** In *Developmental neurocognition: speech and face processing in the first year of life.* Edited by de Boysson-Bardies B, de Schonen S, Jusczyk P, McNeilage P, Morton J. Dordrecht, Netherlands: Kluwer Academic Publishers; 1993:289–304.

The author shows that there is a great deal of variability in adults' abilities to discriminate foreign-language contrasts. Some contrasts are relatively easy to discriminate, while others are difficult, even after extensive training. Relative difficulty is predicted by the relationship between the foreign sounds and native-language speech categories, in terms of their gestural similarity, with the notion that certain foreign-language sounds are assimilated by native-language categories. A description of the various ways in which foreign-language contrasts can relate to native-language categories is developed.

7. Tsushima T, Takizawa O, Sasaki M, Shiraki S, Nishi K, Kohno M, Menyuk P, Best C: **Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6–12 months: language-specific developmental changes in speech perception abilities.** *Proceed Int Conf Spoken Lang Process* 1994, **4**:1695–1698.

8. Ferguson CA, Menn L, Stoel-Gammon C: *Phonological development: models, research, implications.* Timonium, Maryland: York Press; 1992.

9. •• De Boysson-Bardies B: **Ontogeny of language-specific syllabic productions.** In *Developmental neurocognition: speech and face processing in the first year of life.* Edited by de Boysson-Bardies B, de Schonen S, Jusczyk P, McNeilage P, Morton J. Dordrecht, Netherlands: Kluwer Academic Publishers; 1993:353–363.

The author reviews studies on the early influences of ambient language on speech production at the babbling stage in infants reared in different cultures. Just before the age of one year, both vowel and consonant productions show differences.

10. Vihman M M, de Boysson-Bardies B: **The nature and origins of ambient language influence on infant vocal production and early words.** *Phonetica* 1994, **51**:159–169.

11. •• Flege JE: **Production and perception of a novel, second-language phonetic contrast.** *J Acoust Soc Am* 1993, **93**:1589–1608.

Chinese subjects' production and perception of the vowel lengthening cue in words ending in /t/ and /d/ were examined. The results correlated with the degree of foreign accent in English and show a close relationship between subjects' perceptual representations of speech and their abilities to produce distinctions in a foreign language they are trying to learn.

12. Kuhl PK: **Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not.** *Percept Psychophys* 1991, **50**:93–107.

13. •• Iverson P, Kuhl P: **Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling.** *J Acoust Soc Amer* 1995, in press.
The perceptual magnet effect is examined using the techniques of signal detection theory. The results show that measures of absolute sensitivity (indexed by d') are minimal near category prototypes and large near boundaries between categories. Multidimensional scaling analyses show increased perceptual clustering near prototypes and decreased perceptual clustering near category boundaries.

14. Sussman JE, Lauckner-Morano VJ: **Identification and change/no-change discrimination of [i] stimuli: further tests of the 'magnet' effect.** *J Acoust Soc Am* 1994, in press.

15. Aaltonen O, Eerola O, Lang AH, Uusipaikka E, Tuomainen J: **Automatic discrimination of phonetically relevant and irrelevant vowel parameters as reflected by mismatch negativity.** *J Acoust Soc Amer* 1994, in press.

16. Kuhl PK: **Psychoacoustics and speech perception: internal standards, perceptual anchors, and prototypes.** In *Developmental psychoacoustics.* Edited by Werner LA, Rubel EW. Washington, DC: American Psychological Association; 1992:293–332.

17. •• Miller JL: **On the internal structure of phonetic categories: a progress report.** *Cognition* 1994, **50**:271–285.
Overt ratings of the perceptual adequacy or goodness of tokens of phonetic categories were made by adult listeners. The results show that phonetic categories have internal structure; members vary systematically in perceived goodness and goodness ratings shift with changes in context such as the rate of speech.

18. • Wayland SC, Miller JL, Volaitis LE: **The influence of sentential speaking rate on the internal structure of phonetic categories.** *J Acoust Soc Amer* 1994, **95**:2694–2701.
The authors report studies on the effect of sentence- and syllable-level speaking rate on the perceived goodness ratings for members of a phonetic category. The results show differential effects for the two rate manipulations. A change in syllable-level rate altered the location of the best instances and widened the area of best exemplars; a change in sentence-level rate altered only the location of the best exemplars, not the entire range.

19. Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B: **Linguistic experience alters phonetic perception in infants by 6 months of age.** *Science* 1992, **255**:606–608.

20. •• Polka L, Werker JF: **Developmental changes in perception of nonnative vowel contrasts.** *J Exp Psych: Hum Percept Perform* 1994, **20**:421–435.
The authors report on infants' perception of foreign-language vowel contrasts, demonstrating that vowel perception changes from a pattern of language-general to language-specific perception at an earlier age than that indicated for consonants. For consonants, the change occurs between 10 and 12 months; for vowels, the change occurs between 4 and 6 months. The authors also verify the existence of the perceptual magnet effect at 6 months of age for vowels.

21. •• Iverson P, Kuhl PK: **Tests of the perceptual magnet effect for American English /r/ and /l/.** *J Acoust Soc Am* 1994, **95**:2976.
The authors use multidimensional scaling techniques to show how the magnet effect alters the perceptual space underlying phonetic categories. Listeners tested with syllables beginning with /r/ and /l/ were asked to identify and rate the perceived goodness of each syllable. Then they rated the perceived similarity for all pairs of stimuli. Perceived similarity measures were scaled using MDS techniques. The results show that perceptual space is reduced in the region of the category's best instances and expanded in the region of the category boundary.

22. Kuhl P, Iverson P: **Linguistic experience and the 'perceptual magnet effect'.** In *Speech perception and linguistic experience: theoretical and methodological issues in cross-language speech results.* Edited by Strange W. Timonium, Maryland: York Press; 1995:in press.

23. Nosofsky RM: **Exemplar-based accounts of relations between classification, recognition, and typicality.** *J Exp Psych: Learn Mem Cog* 1988, **14**:700–708.

24. Kuhl PK: **Infants' perception and representation of speech: development of a new theory.** In *Proceedings of the International Conference on Spoken Language Processing.* Edited by Ohala JJ, Nearey TM, Derwing BL, Hodge MM, Wiebe GE. Edmonton: University of Alberta; 1992:449–456.

25. •• Strange W: *Speech perception and linguistic experience: theoretical and methodological issues in cross-language speech research.* Timonium, Maryland: York Press; 1995:in press.
Provides a comprehensive survey of recent literature and approaches in cross-language studies of speech perception, speech production, and second-language learning.

26. •• Gracco VL, Löfqvist A: **Speech motor co-ordination and control: evidence from lip, jaw, and laryngeal movements.** *J Neurosci* 1994, in press.
Kinematic patterns of different articulators (lower lip, jaw, and larynx) were examined during speech to identify their coordination. The authors conclude that vocal tract actions for sounds are stored in memory as motor programs and sequenced together during speaking. Speech articulatory motions were influenced by the identity of the surrounding units, suggesting that speech movements are modified to take into account larger numbers of segments.

27. Perkell JS, Matthies ML, Svirsky MA, Jordan MI: **Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: a pilot motor equivalence study.** *J Acoust Soc Amer* 1993, **93**:2948–2961.

28. Stoel-Gammon C, Williams K, Buder E: **Cross-language differences in phonological acquisition: Swedish and American /t/.** *Phonetica* 1994, **51**:146–158.

29. •• Kuhl PK, Meltzoff AN: **Infant vocalizations in response to speech: vocal imitation and developmental change.** *J Acoust Soc Am* 1995, in press.
Vocalizations of infants watching a video of a female talker were recorded at 12, 16, and 20 weeks of age. Infant utterances (n = 144) were measured acoustically and phonetically transcibed. Twenty-six acoustic measures were taken from each infant utterance. The results show developmental change between 12 and 20 weeks of age and also provide evidence of vocal imitation in infants at all three ages.

30. Edelman GM: *Neural Darwinism: a theory of neuronal group selection.* New York: Basic Books; 1987.

31. Konishi M: **Birdsong for neurobiologists.** *Neuron* 1989, **3**:541–549.

32. Meltzoff AN, Moore MK: **Imitation, memory, and the representation of persons.** *Inf Behav Develop* 1994, **17**:83–99.

33. Suomi K: **An outline of a developmental model of adult phonological organization and behaviour.** *J Phonetics* 1993, **21**:29–60.

34. •• Vihman MM: **Variable paths to early word production.** *J Phonetics* 1993, **21**:61–82.
Reviews the sources of variability in infants' productions of words. Infants show developmental articulatory preferences that may provide an 'articulatory filter' on infants' productions, causing them to favor the production of words that contain units that they frequently produce.

35. Studdert-Kennedy M: **Discovering phonetic function.** *J Phonetics* 1993, **21**:147–155.

36. Liberman AM: **In speech perception time is not what it seems.** *Annals NY Acad Sci* 1993, **682**:264–271.

37. Cutler A, Mehler J: **The periodicity bias.** *J Phonetics* 1993, **21**:103–108.

38. •• Moon C, Cooper RP, Fifer WP: **Two-day-olds prefer their native language.** *Inf Behav Develop* 1993, **16**:495–500.
The authors measure sucking behavior in infants using samples of native-language as opposed to foreign-language speech as reinforcers. For half of the infants, Spanish was the maternal language while for the other half, English was the maternal language. Both groups of infants modified their sucking patterns in such a way as to activate the native-language sounds, indicating their preference for the pattern they experienced *in utero.*

39. DeCasper AJ, Fifer WP: **Of human bonding: newborns prefer their mothers' voices.** *Science* 1980, **208**:1174–1176.

40. DeCasper AJ, Spence MJ: **Prenatal maternal speech influences newborns' perception of speech sounds.** *Inf Behav Develop* 1986, **9**:133–150.

41. Lecanuet JP, Granier-Deferre C: **Speech stimuli in the fetal environment.** In *Developmental neurocognition: speech and face processing in the first year of life.* Edited by de Boysson-Bardies B, de Schonen S, Jusczyk P, McNeilage P, Morton J. Dordrecht, Netherlands: Kluwer Academic Publishers; 1993:237–248.

42. Jusczyk PW: **From general to language-specific capacities: the WRAPSA model of how speech perception develops.** *J Phonetics* 1993, **21**:3–28.

43. Jusczyk PW, Friederici AD, Wessels JMI, Svenkerud VY, Jusczyk
•• AM: **Infants' sensitivity to the sound patterns of native language words.** *J Mem Lang* 1993, **32**:402–420.

The authors examined infants' listening preferences for unfamiliar words that either observed or violated native-language phonetic and phonotactic patterns. American nine month olds, but not six month olds, listened significantly longer to words with English, rather than Dutch, sound patterns. Dutch nine month olds showed the opposite pattern of preference. No preferences occurred for low-pass-filtered versions of the words, suggesting that infants were responding to phonetic and phonotactic properties rather than to prosodic ones. When words differed more fundamentally in their prosodic organization (when English was tested against Norwegian), American six-month-olds listened significantly longer to English words.

44. Friederici AD, Wessels JMI: **Phonotactic knowledge of word boundaries and its use in infant speech perception.** *Percept Psychophys* 1993, **54**:287–295.

45. Jusczyk PW, Hirsh-Pasek K, Kemler Nelson DG, Kennedy LJ, Woodward A, Piwoz J: **Perception of acoustic correlates of major phrasal units by young infants.** *Cog Psych* 1992, **24**:252–293.

46. Jusczyk PW, Cutler A, Redanz NJ: **Infants' preference for the
•• predominant stress patterns of English words.** *Child Develop* 1993, **64**:675–687.

English words typically have a strong initial stress. Nine month old infants' preferences were tested for words with a strong–weak as opposed to a weak–strong stress pattern. The results demonstrated a listening preference for the strong–weak pattern. Six month olds showed no preferences.

47. Jusczyk PW, Luce PA, Charles-Luce J: **Infant's sensitivity to phonotactic patterns in the native language.** *J Mem Lang* 1994, **33**:630–645.

48. Chapman RS, Streim NW, Crais ER, Salmon D, Strand EA, Negri NA: **Child talk: assumptions of a developmental process model for early language learning.** In *Processes in language acquisition and disorders.* Edited by Chapman RA. St. Louis: Mosby Year Book; 1992:3–19.

49. Fernald A: **Human maternal vocalizations to infants as biologically relevant signals: an evolutionary perspective.** In *The adapted mind: evolutionary psychology and the generation of culture.* Edited by Barkow JH, Cosmides L, Tooby J. New York: Oxford University Press; 1992:391–428.

50. Aslin RN: **Segmentation of fluent speech into words: learning models and the role of maternal input.** In *Developmental neurocognition: speech and face processing in the first year of life.* Edited by de Boysson-Bardies B, de Schonen S, Jusczyk P, McNeilage P, Morton J. Dordrecht, Netherlands: Kluwer Academic Publishers; 1993:305–316.

51. Werker JF, Lloyd VL, Pegg JE, Polka L: **Putting the baby in the bootstraps: toward a more complete understanding of the role of the input in infant speech processing.** In *Signal to syntax.* Edited by Morgan JL, Demuth K. Hillsdale, New Jersey: Erlbaum; in press.

52. Morgan JL: *From simple input to complex grammar.* Cambridge, Massachusetts: MIT Press; 1986.

53. Fernald A, Mazzie C: **Prosody and focus in speech to infants and adults.** *Develop Psych* 1991, **27**:209–221.

54. Fernald A, McRoberts GW, Herrera C: **Effects of prosody and word position on infants' ability to recognize words in fluent speech.** *J Exp Psych: Learn Mem Cog* 1994, in press.

55. Gopnik A, Choi S: **Names, relational words and cognitive de-
•• velopment in English and Korean speakers: nouns are not always learned before verbs.** In *Beyond names for things: young children's acquisition of verbs.* Edited by Tomasello M, Merriman W. Hillsdale, New Jersey: Erlbaum; in press.

The authors show that language input varies in different languages, with American mothers using a greater number of nouns and Korean mothers using a greater number of verbs in speech to children. Measures of nouns and verbs in the utterances of the children indicated that American and Korean children differed in the relative amount of the two word classes, with Americans producing more nouns and Koreans producing more verbs. The authors argue that linguistic input alters both linguistic and cognitive achievements of young children.

56. Burnham D: **Visual recognition of mother by young infants: facilitation by speech.** *Perception* 1993, **22**:1133–1153.

57. Sherry DF, Schacter DL: **The evolution of multiple memory systems.** *Psych Rev* 1987, **94**:439–454.

58. Squire LR: *Memory and brain.* New York: Oxford University Press; 1987.

59. Pisoni DB: **Long-term memory in speech perception: some
•• new findings on talker variability, speaking rate and perceptual learning.** *Speech Comm* 1993, **13**:109–125.

Studies of the effects of talker variability on speech perception suggest that listeners retain information about the talker's voice in memory and that these details later improve the recognition of novel items spoken by that talker. The results suggest that speech representations include talker-specific information.

60. Swain IU, Zelazo PR, Clifton RK: **Newborn infants' memory for speech sounds retained over 24 hours.** *Develop Psych* 1993, **29**:312–323.

61. Jusczyk PW, Kennedy LJ, Jusczyk AM: **Young infants' retention of information about syllables.** *Inf Behav Develop* 1994, in press.

62. Jusczyk PW, Jusczyk AM, Kennedy LJ, Schomberg T, Koenig N: **Young infants' retention of information about syllables.** *J Exp Psych: Hum Percept Perform* 1994, in press.

63. Mandel DR, Jusczyk PW, Kemler Nelson DG: **Does sentential prosody help infants organize and remember speech information?** *Cognition* 1994, in press.

64. Gerken LA: **Young children's representation of prosodic phonology: evidence from English-speakers' weak syllable omissions.** *J Mem Lang* 1994, **33**:19–38.

65. Bijeljac-Babic R, Bertoncini J, Mehler J: **How do 4-day-old infants categorize multisyllabic utterances?** *Devel Psych* 1993, **29**:711–721.

66. Estes WK: **Concepts, categories, and psychological science.** *Psych Science* 1993, **4**:143–153.

67. McGurk H, MacDonald J: **Hearing lips and seeing voices.** 1976, *Nature* **264**:746–748.

68. Massaro DW: *Speech perception by ear and eye: a paradigm for psychological inquiry.* Hillsdale, New Jersey: Erlbaum; 1987.

69. Green KP, Kuhl PK, Meltzoff AN, Stevens EB: **Integrating speech information across talkers, gender, and sensory modality: female faces and male voices in the McGurk effect.** *Percept Psychophys* 1991, **50**:524–536.

70. Kuhl PK, Tsuzaki M, Tohkura Y, Meltzoff AN: **Human process-
•• ing of auditory-visual information in speech perception: potential for multimodal human-machine interfaces.** In *Proceedings of the international conference on spoken language processing.* Kyoto: The Acoustical Society of Japan; 1994:539–542.

Auditory–visual perception of speech is compared for individuals who speak different languages. In particular, American and Japanese observers watched and listened to both American and Japanese speakers on videotape. The extent to which discrepant auditory and visual information led to 'illusory' combinations of auditory and visual information was examined. The results demonstrated that both groups identified auditorially presented foreign-language syllables less well. In addition, when observers identified auditory-visual speech syllables, there was a strong tendency, especially for Japanese observers, to show greater numbers of illusory responses when watching a foreign speaker.

71. Sekiyama K, Tohkura Y: **Inter-language differences in the influence of visual cues in speech perception.** *J Phonetics* 1993, **21**:427–444.

72. Werker JF, Frost P, McGurk H: **La langue et les lévres: cross-language influences on bimodal speech perception.** *Canadian J Psych* 1992, **46**:551–568.

73. Massaro DM, Tsuzaki M, Cohen MM, Gesi A, Heredia R: **Bimodal speech perception: an examination across languages.** *J Phonetics* 1993, **21**:445–478.

74. Kuhl PK, Meltzoff AN: **The bimodal perception of speech in infancy.** *Science* 1982, **218**:1138–1141.

75. Kuhl PK, Williams KA, Meltzoff AN: **Cross-modal speech perception in adults and infants using nonspeech auditory stimuli.** *J Exp Psych: Hum Percept Perform* 1991, **17**:829–840.

76. Walton GE, Bower TGR: **Amodal representation of speech in infants.** *Inf Behav Develop* 1993, **16**:233–243.

77. Mehler J, Dupoux E, Pallier C, Dehaene-Lambertz G: **Cross-linguistic approaches to speech processing.** *Curr Opin Neurobiol* 1994, **4**:171–176.

78. Mills DL, Coffey-Corina SA, Neville HJ: **Language acquisition and cerebral specialization in 20-month-old infants.** *J Cog Neurosci* 1993, **5**:317–334.

79. Dehaene-Lambertz G, Dehaene S: **Speed and cerebral correlates of syllable discrimination in infants.** *Nature* 1994, **370**:292–295.
••

Authors use event-related potentials (ERPs) to examine speech perception in young infants, demonstrating that infants can detect a change in a syllable within 400 ms.

80. Cheour-Luhtanen M, Alho K, Kujala T, Sainio K, Reinikainen K, Renlund M, Aaltonen O, Eerola O, Näätänen R: **Mismatch negativity shows discrimination of speech stimuli in newborns.** *Hearing Res* 1995:in press.

81. Sharma A, Kraus N, McGee T, Carrell T, Nicol T: **Acoustic versus phonetic representations of speech as reflected by the mismatch negativity event-related potential.** *EEG Clin Neurophys* 1993, **88**:64–71.

82. Aaltonen O, Eerola O, Lang AH, Uusipaikka E, Tuomainen J: **Automatic discrimination of phonetically relevant and irrelevant vowel parameters as reflected by mismatch negativity.** *J Acoust Soc Am* 1994, **96**:1489–1493.
••

The mismatched negativity (MMN) index was used to measure speech discrimination. The authors demonstrate in studies of vowel perception that subjects exhibiting a large perceptual magnet effect show parallel effects in their MMN responses. This suggests that the MMN measure is responsive to the perceived goodness of a phonetic unit.

PK Kuhl, Department of Speech and Hearing Science, JG-15, University of Washington, Seattle, Washington 98195, USA.