8

Experiential Influences on Speech Perception and Speech Production in Infancy

Linda Polka, Susan Rvachew, and Karen Mattock

Mature language users are highly specialized, expert, and efficient perceivers and producers of their native language. This expertise begins to develop in infancy, a time when the infant acquires language-specific perception of native language phonetic categories and learns to produce speech-like syllables in the form of canonical babble. The emergence of these skills is well described by past research but the precise mechanisms by which these foundational abilities develop have not been identified.

This chapter provides an overview of what is currently known about the impact of language experience on the development of speech perception and production during infancy. Throughout we affirm that experiential influences on phonetic development cannot be understood without considering the interaction between the constraints that the child brings to the task and the nature of the environmental input. In the perception and production domains our current understanding of this interaction is incomplete and tends to focus on the child as a passive receiver of input. In our review, we signal a recent shift in research attention to the infant's role in actively selecting and learning from the input.

We begin this chapter by describing what is currently known about the determinants of speech perception and speech production development during infancy while highlighting important gaps to be filled within each domain. We close by emphasizing the need to integrate research across the perception and production domains.

Speech Perception Development

Speech perception development involves a complex interaction between the child and his/her language environment(s). In this section we discuss when and how language input begins to shape speech perception and then consider how speech intake is directed and constrained from the infant side of the interaction. This discussion focuses on the development of speech perception at the segmental or phonetic level, that is, the perception of consonants, vowels, and tones.

Effects of age and language input

Developmental cross-linguistic research has shown that from the first few months of life the ambient language begins to guide the infant toward the goal of language-specific speech perception. Werker and collaborators conducted the first systematic investigations of the effects of age and language experience on phonetic perception. In a series of studies, they found that 6- to 8-month-old English infants and native adult speakers discriminated Hindi and Salish contrasts equally well, but discrimination declined with age for infants tested at 8 to 10 and at 10 to 12 months and was poor for English adults; no decline in discrimination over age was reported for Hindi and Salish infants (Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Tees, 1983, 1984a). These findings showed that infants initially respond to phonetic differences in a language-neutral way but their perception becomes more language-specific by the end of the first year of life.

Werker's findings were initially interpreted as showing that the absence of language experience results in a "loss" of perceptual function. This meshed well with the idea that language experience sets up a perceptual filter for speech, an idea that first emerged to explain poor perception of non-native phonetic contrasts by adults (e.g., Miyawaki et al., 1975). Subsequent developmental cross-linguistic research expanded to include a wider variety of phonetic contrasts from many languages, including consonants, vowels, and tones, as well as more varied infant groups, including infants acquiring languages other than English and infants learning more than one language simultaneously. Now that we have a more detailed picture of the changes in phonetic perception that are shaped by experience with a specific language during infancy, it is time to reassess the filter analogy. For this reason, we will summarize developmental and cross-linguistic research by considering what this work tells us about the functional properties of the native language filter that is emerging in infancy. In the current literature it is clear that the process by which infants arrive at language-specific phonetic perception is more complex than the original interpretation of the early studies by Werker and colleagues. The diverse developmental patterns of speech perception that we describe show that the emerging filter is active and can operate to facilitate as well as to attenuate access to phonetic differences. Furthermore, these patterns reflect an emerging native language filter that is sensitive to the demands of the perceptual task (discrimination vs. labeling), the stimulus domain of the test materials (speech vs. non-speech), and the functional status of the phonetic elements in the native language (phonemic vs. non-phonemic; segmental vs. non-segmental).

Developmental patterns in non-native phonetic perception. The initial view of a "loss" of non-native phonetic perception was re-interpreted when it was found that adult perception of native and non-native contrasts is comparable when task demands are reduced (e.g., Werker & Logan, 1985; Werker & Tees, 1984b), therefore confirming that the perceptual filter established by language experience is not absolute. Similarly, infant speech perception patterns are not so simple. Research clearly indicated that the native language filter emerging in infancy is *domain-specific* and does not impact how infants perceive sounds that are not recognized as speech. The first cross-linguistic evidence to support this claim was provided by Best, McRoberts, and Sithole (1988). They found that English infants failed to show a decline in discrimination of a non-native consonant contrast (lateral–apical clicks) from the Zulu language. The domain-specificity of the native language filter in infants is also confirmed by a study of tone perception in infants. In this study 6- and 9-month-old non-tone (English) and tone language infants (Chinese: Cantonese or Mandarin) were tested on discrimination of lexical tone contrasts cued mostly by fundamental frequency (F0: perceived as speakers' pitch) and non-speech tone analogs with the same F0 differences (Mattock & Burnham, 2006). There was a decline in discrimination over age for lexical tone contrasts in English but not Chinese infants. Discrimination level for non-speech tone was maintained over age for both language groups.

Mattock and Burnham's work (2006) shows that the native language filter operates in a *functionally specific* way – the ability to interpret F0 as a segmental cue was maintained only by infants acquiring a language in which F0 has a segmental (as opposed to a suprasegmental) function. Functional specificity of the native language filter is also supported by findings of Pegg and Werker (1997) showing that infants also fail to maintain perception of an allophonic difference, a phonetic difference that occurs regularly in the ambient language but is not used to signal differences in word meaning. They found that English infants' perception of two allophones of English /p/ – a voiceless aspirated bilabial stop (e.g., the first consonant in "pie") and a voiceless unaspirated bilabial stop (e.g., the second consonant in "spy") – declined between 6–8 and 10–12 months of age despite their exposure to each phone in different phonological contexts.

Cross-linguistic research on the perception of segmental contrasts reveals differences in the timing of the perceptual decline across different non-native contrasts – 6 to 12 months for consonant contrasts (Best, 1995; Best et al., 1990; Best, McRoberts, LaFleur, & Silver-Isenstadt, 1995; Bosch & Sebastián-Gallés, 2003b; Eilers, Gavin, & Oller, 1982; Tsao, Liu, Kuhl, & Tseng, 2000; Tsushima et al., 1994), 6 to 8 months for vowel contrasts (Best et al., 1997; Bosch & Sebastián-Gallés, 2003a; Polka & Werker, 1994), and 6 to 9 months for lexical tone[1] (Mattock & Burnham, 2006).

Although attenuation with age is often observed for perception of non-native contrasts, this is not the only reported developmental pattern. Some non-native contrasts remain discriminable through the native language filter, including certain consonant (see Best, 1991, 1995; Best et al., 1990; Polka, Colantonio, & Sundara, 2001) and vowel contrasts (see Polka & Bohn, 1996). In these studies discrimination levels are consistently quite high across development.[2] It is not yet clear what these findings tell us about the emerging native language filter.

Developmental patterns in native phonetic perception. There is increasing evidence that varied developmental patterns are also observed for *native* language phonetic perception. We observed a facilitative effect of native language experience in our lab when we compared monolingual English and monolingual French listeners in four age groups (6–8 months; 10–12 months; 4 years; adults) on their perception of English /d/ versus /ð/, a contrast that is not phonemic in French. English perceivers showed significant improvement in discrimination of the English /d–ð/ contrast between 10–12 months and 4 years, and further improvement between 4 years and adulthood; French perceivers showed no change with age (Polka et al., 2001; Sundara, Polka, & Genesee, 2006). Tsao et al. (2000) also found evidence for native language facilitation in Chinese-learning infants' perception of a native fricative–affricate contrast. These findings challenge the long established view that language experience serves *only* to prevent a developmental decline in perceptual discrimination of some contrasts. The evidence for facilitation as well as maintenance demonstrates that the language filter is *active*: it can enhance as well as reduce accessibility to phonetic contrasts.

An age-related decline in discrimination in infancy has also been found for a native phonetic contrast, specifically English-learning infants' discrimination of the native /s–z/ fricative voicing contrast (Best & McRoberts, 2003). This finding indicates that some phonetic differences are less salient than others irrespective of language experience, leading to different developmental patterns within the native language. Moreover, findings also show that infant perception of native contrasts also depends on the task at hand. For example, Stager and Werker (1997) observed that 14-month-old English-learning infants were unable to detect a native /bI–dI/ contrast when each syllable was paired with a moving object, but they succeeded when the syllables were paired with a static checkerboard pattern. The authors argue that differences in the task demands and overall processing load in these two tasks are important in understanding these differences in perception within the native language.

Research on infants being raised bilingually also suggests a complex pattern for native phonetic development. Bosch and Sebastián-Gallés (2003a) studied infants at different ages from three language groups (Spanish, Catalan, and bilingual Spanish/Catalan). They examined their perception of the /ε–e/ vowel contrast which is phonemic in Catalan but not in Spanish and is difficult for monolingual speakers of Spanish to discriminate (Pallier, Bosch, & Sebastián, 1997). All three infant groups discriminated this contrast at 4 months of age. As expected, an attenuation of perception with age was observed for the Spanish infants but not for the Catalan infants. The bilingual group showed a U-shaped developmental pattern whereby discrimination declined between 4 and 8 months of age, but then improved to the level of native monolingual Catalan infants at 12 months. A U-shape developmental pattern was reported in two subsequent studies that examined consonant perception in bilingual and monolingual infants (Bosch & Sebastián-Gallés, 2003b; Burns, Werker, & McVie, 2003).

Note that this U-shape pattern was observed when bilingual infants were tested on contrasts that have a conflicting status across their two languages – native in one of the languages and non-native in the other. At present it is not known why perception of such contrasts reveals a temporary weak spot in the bilingual infant's emerging native language filter(s). The conflicting functional status, reduced exposure to such contrasts in the bilingual input, or both factors, may contribute to the U-shape pattern. Alternatively, the bilingual infant may construe the perceptual task differently than the monolingual infant. (See Werker, Hall, & Fais, 2004, for a discussion of U-shape developmental functions.) For now, these data pose a challenge and an opportunity for future research.

Overall, the developmental course that unfolds as infants begin tuning into the phonetic structure of their native language is not uniform. Diverse developmental patterns evident in perception of native and non-native phonetic categories reveal an emerging native language filter that operates in an active, task-specific, domain-specific, and functionally specific way to facilitate as well as to attenuate access to phonetic differences. To retain the filter analogy to conceptualize these assorted effects of language experience, our notion of a filter must be more sophisticated than a passive sieve. The filter evident in the existing data is more akin to a resonator, a filtering device that can selectively enhance as well as selectively attenuate an input signal. Even with a more appropriate analogy, there is no current conceptual framework that can predict when and explain how these diverse developmental patterns arise. (See Best & McRoberts, 2003, for a detailed discussion of hypotheses proposed to explain varied developmental patterns.) A better understanding of how language input is tied to developmental patterns is needed. There is little doubt that phonetic contrasts vary in perceptual salience in ways that are independent of linguistic function and impact perceptual development. One challenge for future research is to provide an acceptable metric to gauge this variation in perceptual salience and determine how it interacts with age and language experience. We must also understand other factors that impact speech intake in infants and explain how these factors interact with language input.

Variables influencing speech intake in infancy

To understand how language input shapes developmental changes in phonetic perception, we must identify factors that constrain and direct the infant's intake of speech information. Typically, researchers present stimuli in very quiet testing environments using test paradigms that tap a perceptual function defined by the experimenter, for example discrimination or categorization. This approach is informative yet it tells us very little about what happens when infants encounter speech in their everyday lives. Learning through exposure to the ambient language requires the infant to selectively attend to speech and detect relevant patterns in the speech stream. Outside the speech lab, infants encounter speech in presence of noise and other competing auditory and visual signals and are not explicitly reinforced for responding to specific phonetic elements. Unlike the speech presented in many laboratory studies, real language input consists of connected speech in which a critical feature or property of the language has not been isolated. Although very little is known about how infants gain access to relevant speech information "in the wild," some studies have explored this question by assessing what infants learn from a brief exposure to a corpus of speech that has been carefully structured by the researcher. Findings obtained using this artificial language approach to study phonetic and phonological development are described in Gerken (this volume). Our discussion will focus on two other lines of research that have explored speech intake in infants.

Effects of noise and hearing loss. Infant speech perception is impacted by conditions that alter the amount or quality of language input available, such as the presence of hearing loss or the presence of background noise. Although hearing loss clearly impacts speech perception development, few systematic studies of its specific effects have been conducted with infants. Houston, Pisoni, Iler Kirk, Ying, and Miyamoto (2003) compared attention to a visual pattern, either paired with an auditory speech signal or with silence, by normal hearing and hearing impaired infants before and after cochlear implantation (CI). Their results demonstrate a strong impact of hearing loss on infant attention to speech. Normal hearing infants showed a strong preference for the stimulus trials with a speech signal. Hearing impaired infants showed no difference in looking time across speech and silent trials before CI; a preference for speech began to emerge in some infants 6 months after CI but was significantly weaker compared with normal hearing infants. Despite their poor selective attention to speech, the hearing impaired infants were able to discriminate an obvious vowel versus syllable difference ("ah" vs. "hop") in a laboratory test situation. However, poor attention to speech is likely to impact their processing of more subtle speech patterns in less optimal listening conditions.

We found that experience with otitis media with effusion (OME) also impacts infant responses to speech in the first year of life (Polka & Rvachew, 2005). Otitis media with effusion causes fluid to accumulate in the middle ear and can create a mild to moderate fluctuating conductive hearing loss. To assess effects of OME on phonetic discrimination, we tested 6- to 8-month-olds on a native phonetic contrast, /bu–gu/, using the conditioned headturn procedure. Tympanometry, performed after discrimination testing, revealed three groups of infants: (a) infants with middle ear effusion on test day, (b) infants who had received medical treatment for OM in the past but showing no middle ear effusion on test day, and (c) infants with no history of OM and no effusion on test day. Discrimination performance was best for infants with no history of OME and worst for infants with middle ear effusion on the day of testing. Infants with a history of OME but normal middle ear function on test day showed an intermediate level of discrimination performance, suggesting that OME negatively impacts phonetic perception even after the middle ear fluid is gone. Furthermore, the poorer performance in the history-only group (b above) cannot be explained as an effect of reduced audibility and suggests that experience with OME affects infant attention to speech.

Background noise present in many natural listening environments can also impede access to relevant speech patterns. When tested in presence of noise, infants can detect phonetic differences (Nozza, Miller, Rossman, & Bond, 1991), segment words from fluent speech (Newman & Jusczyk, 1996), and recognize their name (Newman, 2005) and their mother's voice (Barker & Newman, 2004). However, infants need substantially higher signal-to-noise ratios than adults to succeed at these tasks.

Noise acts to block access to relevant sensory information and pull attention away from the speech signal or critical parts of the speech signal. Although these effects are typically inseparable in the real world, it is informative to explore how each impacts infant perceptual processing. Moreover, as infants mature, their ability to control attention will improve whereas the sensory impact of noise is unlikely to change. In our lab, we tested infant phonetic perception using a distraction masker paradigm to assess the role of selective attention independently of sensory effects of noise (Polka, Rvachew, & Molnar, submitted). We tested infant discrimination of /bu/ versus /gu/ using a habituation procedure. To create a distractor condition a high frequency noise was added to the sound file of each syllable so that it gated on and off with the onset and offset of the syllable. The distractor noise was a recording of bird and cricket songs whose frequencies do not overlap with the test syllables; this added signal does not make it harder to hear the syllables but it can distract infants if they cannot focus their attention well. Infants (6–8 months) tested in quiet (i.e., unmodified syllables) performed significantly better than infants tested in the distractor condition; discrimination scores showed little overlap between the two groups. These findings indicate that perceiving auditory phonetic patterns in the presence of noise poses a substantial cognitive challenge for young infants. Recent findings suggest that the availability of multi-modal speech information – an auditory-visual speech stimulus – may provide the critical support that makes speech perception in noise possible for infants (Hollich, Newman, & Jusczyk, 2005).

Infant perceptual biases. A second way that researchers have explored how infants access speech information is by observing their listening preferences in test paradigms that allow infants to control their access to speech samples. Infant listening preferences are evidence of perceptual biases that guide speech intake. Research has focused on biases related to suprasegmental or indexical properties of speech such as speaker's voice, affect, or stress (see Werker & Curtin, 2005). Recent research in our lab shows that infants display strong biases at the phonetic level as well. Our earlier cross-language studies of infant vowel discrimination revealed very robust and predictable directional asymmetries (see Polka & Bohn, 2003). For example, Polka and Werker (1994) found that infants who were presented with one direction of change, /y/ to /u/, found discrimination easier than infants presented with the same vowel pair in the reverse direction, that is, a change from /u/ to /y/. In this and other examples, we observed that vowel discrimination was consistently easier when infants were tested on a change from a less-peripheral to a more-peripheral vowel within the vowel space, suggesting that infant perception is sensitive to the structure of the vowel space. Subsequent research using a preferential listening task has confirmed that these directional asymmetries reflect a strong perceptual bias for peripheral vowels (i.e., vowels closer to the corners of the F1/F2 vowel space, /i/, /a/, /u/) – the same vowels that are strongly favored in vowel inventories across languages. Vowel biases measured in young infants are similar across infants learning different languages but change as language acquisition unfolds. Experiments with adults also show that the vowel biases observed in infancy have been shaped by language experience, toward optimizing vowel processing in a specific language. We can gain many insights into speech perception from the study of infant perceptual biases because they show us where the infants' perceptual priorities lie and point to the information that they are actively extracting from their input. Perceptual biases can also provide a relative index of perceptual salience and thus potentially help explain the variability in phonetic perception observed in the cross-language research described above.

In an effort to account for different aspects of speech perception development Werker and Curtin (2005) recently outlined a framework that integrates a wide range of speech perception phenomena (e.g., input effects, perceptual biases, word learning) within the broader context of language acquisition. The child's active intake of speech information is a prominent feature of their model, called PRIMIR (Processing Rich Information from Multidimensional Interactive Representations). According to PRIMIR three variables – initial biases, developmental level of the child, and task demands – act as dynamic filters and work together to direct the infant's attention to speech. Further research investigating speech intake processes is needed to test the merits of the PRIMIR model and to advance our overall knowledge of speech perception development.

We have outlined some functional properties of the native language filter that is emerging in early infancy and have discussed a recent shift in research focus toward factors that modulate speech intake in the developing infant. However, the story of infant speech perception is incomplete until we also understand how these perceptual processes interact with and are shaped by emerging speech production skills.

Speech Production Development

Recent developments in the study of infant speech production rest on the pioneering efforts of Oller (1980), Stark (1980), and Koopmans-van Beinum and van der Stelt (1986) to describe the course of infant speech development using metrics specially adapted to the infant context. Oller's (2000) comprehensive overview of this literature can be briefly summarized as follows: Normally developing infants universally progress through a series of stages of vocal development that culminate in the production of canonical babble, typically by 7 months of age; canonical babble is characterized by the production of speech-like consonant–vowel syllables, often produced as rhythmic strings of reduplicated syllables containing stop consonants and front or central vowels; and there is considerable continuity in the phonetic content of babble and early meaningful speech. More recent research has been directed at understanding the biological and experiential influences on the form of infant vocalizations.

Biological influences on early speech production

Vocal tract anatomy. Major developmental changes in vocal tract shape occur shortly after birth, including the descent of the larynx and a sharper angle between the oral and pharyngeal cavities. A significant increase in the length of the vocal tract also occurs during the first year, with growth of the pharyngeal cavity being disproportionately large relative to oral cavity growth throughout development (Fitch & Giedd, 1999; Kent & Vorperian, 1995; Kent, Vorperian, Gentry, & Yandell, 1999).

The constraints that vocal tract morphology might place on infant speech output have been investigated through computer modeling by Ménard, Schwartz, and Boë (2002, 2004). Although infant vocal tract anatomy does partially explain the preference for low and front vowels, they demonstrated that the infant's vocal tract anatomy does not prevent the production of the full range of vowels used in the ambient language. However, while it is possible to produce vowels with an infant vocal tract that are perceptually equivalent to adult vowel categories, in many cases the infant would need to employ different articulatory gestures than the adult to achieve the same perceptual outcome. Ménard et al. (2002) also identified acoustic cues to vowel identity that are valid for the full range of vocal tract sizes. Therefore, infants can potentially perceive equivalency between their own vowel productions and those of adult models using the same normalization algorithms for speech-like vocalizations produced by talkers of all ages.

Speech motor control. Infant speech production is clearly limited by immature speech motor control abilities although the exact nature of these limitations has not been fully determined. Indirect investigations of the infant's ability to control the vocal tract have been conducted within the context of the frame/content theory. MacNeilage (1998) proposed that the "frame" for speech production is the repeated opening and closing of the vocal tract. The ability to modulate the "frame" to produce varied "content" is hypothesized to emerge quite late in development, with "frame dominance" persisting through the early word learning stage. Consequently, very little individual variation in the phonetic content of babble or even early words is expected during infancy, within or across language groups. MacNeilage and Davis (2000) summarized a number of studies that provide support for this hypothesis in the form of apparently universal patterns of consonant and vowel co-occurrence. However, these results are based on phonetic transcriptions that are of questionable reliability and validity for infant speech.

Kinematic studies of infants' articulatory movements and acoustic analyses of infants' speech output have provided a more direct and reliable picture of physiological constraints on infant speech production (Green, Moore, Higashikawa, & Steeve, 2000; Green, Moore, & Reilly, 2002; Sussman, Duder, Dalston, & Caciatore, 1999; Sussman, Minifie, Buder, Stoel-Gammon, & Smith, 1996). Overall these data are consistent with an initial dominance of the "mandibular frame," followed by a progressive differentiation of articulator movements. At the same time, rapid changes in speech motor control appear to be occurring during the first 4 months of the child's babbling experience, and limitations on infant control of the articulatory system do not prevent the infant from producing a wide variety of consonant–vowel combinations even during the first year. This leads to questions about how the infant achieves this ability, given a continuously changing vocal tract anatomy and limited speech motor control.

Guenther (1995) proposed a solution to the infant's problem in the form of a computational model (Directions in Articulatory space to Velocities in Articulator space; DIVA). A fundamental characteristic of this self-organizing model is that the goal of speaking is considered to be the production of certain acoustic products that will be perceived by the listener as the intended target sounds. Auditory perceptual feedback is used to develop the mapping between the acoustic target and the required vocal tract constrictions (Guenther, Hampson, & Johnson, 1998). Callan, Kent, Guenther, and Vorperian (2000) demonstrated that this model adapts well to developmental changes in vocal tract size and shape. This model predicts that access to both adult and self-produced speech is critical to the development of speech motor control. Adult input allows the infant to develop representations for language-specific acoustic–phonetic targets (as described above). Access to the infant's own speech allows the child to use auditory feedback to flexibly adapt articulatory gestures to achieve the production of those targets.

Experiential influences on speech production

Empirical investigations have demonstrated that auditory input is crucial for normal speech development during infancy. The emergence of the canonical babble stage is delayed in infants with sensory–neural hearing loss (Oller & Eilers, 1988) but canonical babble appears in the vocalizations of hearing impaired children shortly after they receive cochlear implants (Ertmer & Mellon, 2001). These studies do not, however, illuminate the exact nature of experiential influences on the course of prelinguistic vocal development. Oller (2000) suggests that a certain amount of auditory experience is required for "triggering the events that lead to well-formed syllable production" (p. 132). It is not clear that the specific phonetic content of the adult input is important to the process and indeed, no direct link between underlying perceptual representations and the characteristics of infant speech has been established. The dominant view posits that babbling is a strongly canalized motor behavior in which feedback of self-produced sounds serves to help the infant coordinate articulation and phonation (Koopmans-van Beinum, Clement, & van den Dikkenberg-Pot, 2001) or regulate the temporal rhythm of babble, in a manner similar to the way in which audible rattles seem to facilitate rhythmic hand-banging (Ejiri & Masataka, 2001).

By contrast, the DIVA model emphasizes the role of speech input in the formation of auditory–perceptual targets for speech output as well as the importance of auditory feedback of the infant's own speech for the achievement of speech motor control. This model predicts that the speech environment will influence the acoustic-phonetic characteristics of the infant's speech output from at least the beginning of the canonical babbling stage. Investigations of this hypothesis sometimes employ laboratory induced variations in speech input to the infant but more frequently examine the impact of natural variations in the auditory environment associated with different language groups.

Laboratory investigations of the impact of speech input on speech production. Two developmental changes in infant vowels occur during the first year: first, more vowels with full oral resonance are produced in comparison to vowels with nasal resonance; subsequently, the infant's vowel repertoire expands to include the point vowels (/i/, /a/, /u/) in addition to the predominant low and central vowels. Two laboratory studies have shown that auditory input plays a role in these developments.

An elegant series of studies was conducted by Bloom (Bloom, 1988; Bloom, Russell, & Wassenberg, 1987) in which both the timing and content of speech input to 3-month-old infants were varied systematically in the laboratory. Infants produced a higher proportion of more speech-like utterances involving full resonance when they received speech input (compared with non-speech vocal input), and this effect was enhanced when the input was provided contingent upon the infants' vocalizations.

In a study focused on the specific phonetic content of the adult input and the infant speech output, Kuhl and Meltzoff (1996) presented one of three point vowels to infants aged 12, 16, and 20 weeks. Perceptual and acoustic analyses showed that the infants shifted their vowel production to better match the target vowel category, even at the youngest age. Kuhl and Meltzoff speculated that this matching-to-target phenomenon, combined with the shift to language-specific perceptual processing, underlies developmental changes in acoustic characteristics of infant vowel production.

Cross-linguistic investigations of the impact of speech input on speech production. Descriptions of the prosodic characteristics of infant speech also suggest that environmental input impacts infant speech output. For example, French-learning infants' babble can be differentiated from the babble of infants learning other languages with respect to intonational contours, syllable structure, and number of syllables per utterance (Hallé, de Boysson-Bardies, & Vihman, 1991; Levitt & Utman, 1992; Whalen, Levitt, & Wang, 1992). Maneva and Genesee (2002) recently reported similar findings for a single child learning both English and French, recorded separately with his English-speaking mother or his French-speaking father. Unfortunately, these studies involved very small samples and are largely based on phonetic transcriptions of the infants' speech; therefore the existence of "babbling drift," in which infant babble "drifts" toward the ambient language in its sound properties, is not universally accepted (Oller, 2000).

In our lab, we have been describing the acoustic characteristics of vowels produced by infants learning Canadian English or Canadian French in a cross-sectional study (Rvachew, Mattock, Polka, & Ménard, 2006). Figure 8.1 shows the mean first (F1) and second (F2) formant frequencies of vowels produced by the 42 infants that we have recorded thus far. The language groups do not differ with respect to F1 or F2 frequency at 8 months of age. At 10 months of age the groups differ in a manner similar to that reported by de Boysson-Bardies, Hallé, Sagart, and Durand (1989), with the mean F2 being significantly lower for French-learning than for English-learning infants. Between 10 and 18 months of age, the mean F2 for the English-learning group falls below that of the French-learning group which shows a rising F2 during this period. The result is a significant interaction between age and language group for F2 [$F\ (3,42) = 3.68$, $p = .02$].

These data do not represent a pattern of initial overlap of the vowel spaces followed by a linear divergence of the two groups' "average vowel" with advancing age, as we had expected. Rather, the patterns of developmental change and cross-linguistic differences suggest a more complex and dynamic situation in which the child's ability to attend to and reproduce specific features of the native language vowel space shifts with age. However, it is difficult to interpret these findings because it is not clear what the language-specific targets are for the infant vowel space. The vowels produced by infants are necessarily described in terms of the mean and dispersion of formant values for the entire vowel space while the adult input is universally presented as formant values for specific vowels produced in an adult-directed fashion. For example, Escudero and Polka (2003) have shown that there are acoustic differences among Canadian English and Canadian French vowels in adult speech, even for shared point vowels. However, these data cannot be used to predict developmental changes in the infant's mean vowels because the adult vowel space is not described using the same method that is used for the description of

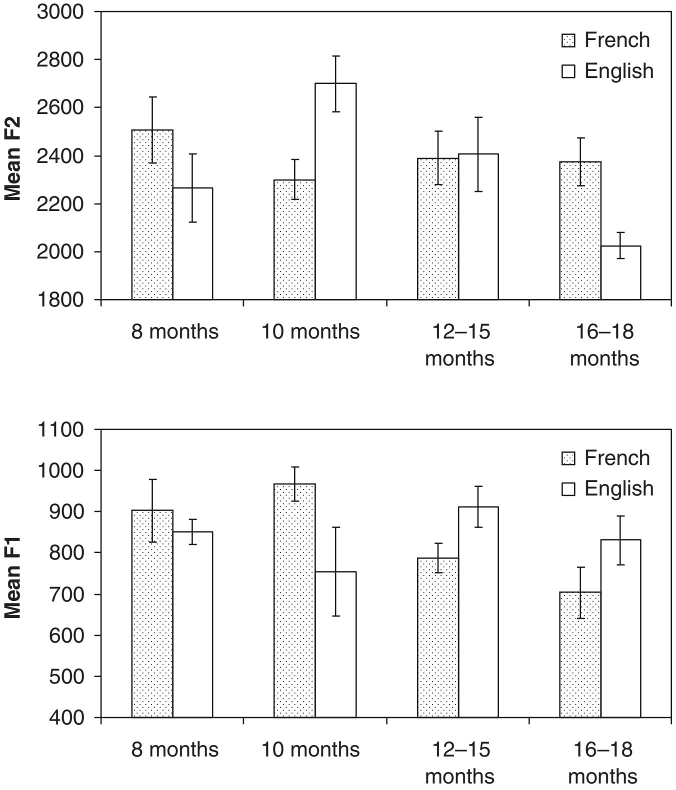

Figure 8.1 Mean second formant frequencies (F2; top panel) and first formant frequencies (F1; bottom panel) for infants learning Canadian English (white bars) or Canadian French (speckled bars), by age group. Error bars represent standard error of the mean.

infant vowels. Furthermore, it is clear that adult-directed speech is significantly different from infant-directed speech and this is likely to have implications for the acoustics of the vowels that infants hear (e.g., see Kuhl et al., 1997). Studies in which the speech addressed to the infant and the speech produced by the infant are recorded in the same context and described using the same procedures are required.

In addition to establishing the characteristics of the input vowel space from the infant's perspective, we need to have a better understanding of how the infant processes that input if we are to predict patterns of developmental change in speech production. For example, the non-linear changes in infant vowel characteristics that are shown in Figure 8.1 might be explained by age-related changes in infant attention to specific properties of the adult input and the infant's own speech. As an example, one characteristic of the French vowel space is the presence of rounded vowels that are cued by a complex integration of the relatively low amplitude second, third, and fourth formant frequencies (Ménard et al., 2004). It is possible that the cue for lip rounding has diminished salience in the child's own speech at a young age. This cue may become available to the child later in the infant period, as a function of increased exposure, changes in vocal tract and laryngeal anatomy, and improved control of lip movements. If the infant then begins to try to manipulate the cue in his or her own speech output, this could cause a non-linear shift in the child's vowel space. However, this hypothesis cannot be tested without more information about the child's ability to attend to and process different acoustic cues in infant and adult speech. This brings us to the issue of infant intake of speech input.

The impact of variations in speech intake. The infant's ability to receive speech input is obviously impacted by the integrity of the child's auditory system, and hearing impairment has a clear impact on infant speech production as described above. To date, almost no research has considered the potential role of the infant's active efforts to listen to the speech input and select the relevant information.

It is now well known that infants listen preferentially to certain types of speech but the developmental implications of variations in selective attention are not clear. Vihman and Nakai (2003) reported a negative correlation between listening preferences for and productive use of a given consonant. The results are interpreted within the context of the articulatory filter hypothesis in which the physical act of producing a given speech sound impacts on the child's perception of speech and helps the child to develop "more robust lexical and phonological representations" (p. 1017). Vihman (2002) further speculates that articulation of adult-like syllables activates neurons that are active when performing or observing an action (i.e., mirror neurons; see Vihman for more detail). This hypothesis presumes that speech perception and speech production development are independent during the first year of life. The onset of canonical babble is attributed to "maturation" and "rhythmic motoric advances" while the child's implicit perceptual knowledge of native language sound categories is attributed to "passive intake"; when the mirror neurons are activated and perception and production become integrated the infants "lay down phonological representations at a new level" (p. 1017; Vihman & Nakai, 2003).

Note that this account of early phonetic development is based upon a correlational study in which it is impossible to determine the direction of causality. The finding that the infants preferred to listen to consonants that they were *less* likely to produce is difficult to interpret. However, the focus on the child's active intake of information from the language environment is extremely important, and there is a clear need for more studies in which the relationship between selective attention and speech production is explored.

Interactions between the integrity of the auditory mechanism and the development of selective attention have been proposed, as discussed above in relation to the reduced attention to speech observed in recipients of cochlear implants (Houston et al., 2003). In particular, selective attending deficits may explain the impact of otitis media on language development. The impact is small and the clinical significance of otitis media-related language delay has been questioned (e.g., Paradise et al., 2000). However, the hearing loss associated with otitis media is so subtle and transitory that any measurable impact on language development is surprising. And yet these impacts can be observed even during the first year of life in the form of slower emergence of the canonical babble stage and a restricted vowel space (Rvachew, Slawinski, Williams, & Green, 1996, 1999). Current hypotheses about the way in which otitis media impacts on speech development recognize that the issue is not simply one of signal audibility. The fluctuating nature of the hearing deficit may make it difficult for the child to discover regularities in the speech input and lead the child to become less attentive to speech (Feagans, 1986; Mody, Schwartz, Gravel, & Ruben, 1999). Furthermore, 12-month-olds with chronic otitis media experience fewer episodes of joint attention with their parents than children with normal hearing (Yont, Snow, & Vernon-Feagans, 2003). Thus it appears that the early vocalizations of children with early-onset otitis media are determined by a complex interaction of biological factors (integrity of the peripheral auditory system), input factors (quality of input provided by parents), and intake factors (child attention to speech and ability to engage in joint attention routines).

Speech Perception and Speech Production Development: Challenges for the Future

The study of infant phonetic development has been marked by tension between competing perspectives about the goal of the enterprise itself, the purported basis for developmental change during infancy, and the significance of these changes for later language acquisition. Specifically, some researchers have focused on the universal characteristics of the infant's phonetic abilities while others have investigated individual variation in phonetic development. Developmental changes in language acquisition are assumed by some to emerge directly from the maturation of certain neurological, physiological, and anatomical structures, while others privilege the infant's active efforts to learn from the input. Although phonetic development during infancy is considered to be an important foundation for later language learning, the nature of its relation to later phonological, lexical, and syntactic development remains a question of debate.

Competing perspectives on the relation between perception and production processes in early development were labeled by de Boysson-Bardies et al. (1989) as the *independence hypothesis* and the *interactional hypothesis*. From the independence perspective, speech perception and production skills develop independently during the first year, with any continuity between infant phonetic skills and later language development viewed as a function of common biological constraints on perceptual and articulatory performance during the prelinguistic and early linguistic stages. For example, Kent and Miolo (1994) suggested that speech perception and speech production "may have somewhat different courses of development, but they are ultimately integrated in spoken language competence" (p. 304). Locke (1994) proposed independent neurological mechanisms and maturational timecourses to account for early phonetic development and phonological development after the first 50 word stage.

From the interactional perspective, the child's developing knowledge of the perceptual and articulatory characteristics of native language phonetic categories is seen as integrated from the beginning. For example, de Boysson-Bardies et al. (1989) stated that "articulatory procedures that are mastered step by step are oriented by auditory configurations" (p. 2). This perspective is associated with strong claims for continuity between phonetic development during this first year and language development during the second year, as exemplified by Tsao, Liu, and Kuhl's (2004) assertion that "phonetic perception plays a critical role in the early phases of language acquisition" (p. 1082).

While the research reviewed here has not yet provided a resolution to this debate, recent changes in theoretical perspectives, methodological tools, and research approaches promise considerable advances in our understanding of how perception and production might interact during the first year to produce the emergence of language in the second year of life. The DIVA model provides a theoretical account of how these domains influence each other throughout the lifespan. The PRIMIR model suggests that apparent discontinuities in infant performance may actually be indicative of a fundamental underlying continuity from the prelinguistic to the linguistic phases of development. New research tools provide us with the opportunity to assess the merits of these theories. Investigation of individual differences in babbling after taking into account universal biomechanical constraints has been facilitated by a number of new technologies (e.g., magnetic resonance imaging of the vocal tract, computational modeling of the impact of vocal tract development on speech output, and kinematic and acoustic descriptions of infant speech production). Infant speech perception research has in the past tapped into natural variation in language input via cross-linguistic comparisons to explore linguistic influences. Researchers are beginning to actively control and manipulate language experience. For example, Kuhl, Tsao, and Liu (2003) manipulated different components of a natural language setting to show the impact of a live social context on phonetic learning. Given the wide acceptance of highly controlled artificial language paradigms, cross-linguistic studies implementing this approach will soon provide insights into the language-specificity (or lack thereof) of infant on-line speech processing. Along with these new methods there is a growing interest in examining individual differences in perceptual responding under conditions that are more akin to the infant's natural environment.

Together these new technologies make it possible for researchers, currently working separately in the domains of speech perception or speech production, to come together to investigate the relationship between individual variations in speech perception and production performance in the same infants. The DIVA and PRIMIR models suggest some specific directions for future research. The PRIMIR model predicts that there will be individual differences and developmental changes in the intake of specific aspects of the rich and multidimensional input available to the infant. Ongoing efforts to understand the infant's role in the selection of input are extremely important. The DIVA model shows that a critical but as yet unexplored aspect of this input is the infant's own speech. Collaborative and interdisciplinary research is needed to focus research on the interplay between perception and production processes within the developing child. The results will have a profound impact on our understanding of early language development.

Notes

1 Note that no studies of lexical tone perception have been conducted with infants younger than 6 or older than 9 months so it is not known whether the perceptual decline actually occurs earlier in development or continues later.

2 The lower discrimination levels in the Polka et al. study are an exception.

References

Barker, B. A., & Newman, R. S. (2004). Listen to your mother! The role of talker familiarity in infant streaming. *Cognition*, *94*, B45–B53.

Best, C. T. (1991, April). *Phonetic influences on the perception of non-native speech contrasts by 6–8 and 10–12-month-olds.* Paper presented at the Meeting of Society for Research in Child Development, Seattle, WA.

Best, C. T. (1995). A direct-realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and early linguistic experience* (pp. 171–206). Baltimore, MD: York Press.

Best, C., & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech*, *46*, 183–216.

Best, C. T., McRoberts, G. W., Goodell, E., Wormer, J. S., Insabella, G., Kim, P., et al. (1990, April). *Infant and adult perception of non-native speech contrasts differing in relation to the listener's native phonology.* Paper presented at the Meeting of the International Conference on Infant Studies, Montreal.

Best, C. T., McRoberts, G., LaFleur, R., & Silver-Isenstadt, J. (1995). Divergent developmental patterns for infants' perception of two nonnative consonant contrasts. *Infant Behavior and Development*, *18*, 339–350.

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

Best, C., Singh, L., Bouchard, J., Connelly, G., Cook, A., & Faber, A. (1997, April). *Developmental changes in infants' discrimination of non-native vowels that assimilate to two native categories.* Paper presented at the Meeting of Society of Research in Child Development.

Bloom, K. (1988). Quality of adult vocalizations affects the quality of infant vocalizations. *Journal of Child Language*, *15*, 469–480.

Bloom, K., Russell, A., & Wassenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language*, *14*, 211–227.

Bosch, L., & Sebastián-Gallés, N. (2003a). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, *46*, 217–243.

Bosch, L., & Sebastián-Gallés, N. (2003b). *Language experience and the perception of a voicing contrast in fricatives: Infant and adult data.* Paper presented at the 15th International Conference of the Phonetic Sciences, Barcelona, Spain.

Burns, T. C., Werker, J. F., & McVie, K. (2003). Development of phonetic categories in infants raised in bilingual and monolingual environments. In B. Beachley, A. Brown, & F. Conlin (Eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Callan, D. E., Kent, R. D., Guenther, F. H., & Vorperian, H. K. (2000). An auditory-feedback-based neural network model of speech production that is robust to developmental changes in the size and shape of the articulatory system. *Journal of Speech, Language, and Hearing Research*, *43*, 721–738.

De Boysson-Bardies, B., Hallé, P., Sagart, L., & Durand, C. (1989). A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language*, *16*, 1–17.

Eilers, R. E., Gavin, W. J., & Oller, D. K. (1982). Cross-linguistic perception in infancy: Early effects of linguistic experience. *Journal of Child Language*, *9*, 289–302.

Ejiri, K., & Masataka, N. (2001). Co-occurrence of preverbal vocal behavior and motor action in early infancy. *Developmental Science*, *4*, 40–48.

Ertmer, D. J., & Mellon, J. A. (2001). Beginning to talk at 20 months: Early vocal development in a young cochlear implant recipient. *Journal of Speech, Language, and Hearing Research*, *44*, 192–206.

Escudero, P., & Polka, L. (2003). *A cross-language study of vowel categorization and vowel acoustics: Canadian English versus Canadian French.* Paper presented at Paper presented at the 15th International Conference of the Phonetic Sciences, Barcelona, Spain.

Feagans, L. (1986). Otitis media: A model for long-term effects with implications for intervention. In J. F. Kavanagh (Ed.), *Otitis media and child development* (pp. 192–208). Parkton, MD: York Press.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America*, *106*, 1511–1522.

Green, J. R., Moore, C. A., Higashikawa, M., & Steeve, R. W. (2000). The physiologic development of speech motor control: Lip and jaw coordination. *Journal of Speech, Language, and Hearing Research*, *43*, 239–255.

Green, J. R., Moore, C. A., & Reilly, K. J. (2002). The sequential development of jaw and lip control for speech. *Journal of Speech, Language, and Hearing Research*, *45*, 66–79.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, *102*, 594–621.

Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review*, *105*, 611–633.

Hallé, P. A., de Boysson-Bardies, B., & Vihman, M. M. (1991). Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants. *Language and Speech*, *34*, 299–318.

Hollich, G., Newman, R. S., & Jusczyk, P. W. (2005). Infant's use of synchronized visual information to separate streams of speech. *Child Development*, *76*, 1–16.

Houston, D. M., Pisoni, D., Iler Kirk, K., Ying, E. A., & Miyamoto, R. T. (2003). Speech perception skills of deaf infants following cochlear implantation: a first report. *International Journal of Pediatric Otorhinolaryngology*, *67*, 479–495.

Kent, R. D., & Miolo, G. (1994). Phonetic abilities in the first year of life. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language* (pp. 303–334). Oxford: Blackwell.

Kent, R. D., & Vorperian, H. K. (1995). Anatomic development of the craniofacial-oral laryngeal systems: A review. *Journal of Medical Speech-Language Pathology*, *3*, 145–190.

Kent, R. D., Vorperian, H. K., Gentry, L. R., & Yandell, B. S. (1999). Magnetic resonance imaging procedures to study the concurrent anatomic development of vocal tract structures: preliminary results. *International Journal of Pediatric Otorhinolaryngology*, *49*, 197–206.

Koopmans-van Beinum, F. J., Clement, C. J., & van den Dikkenberg-Pot, I. (2001). Babbling and the lack of auditory speech perception: a matter of coordination? *Developmental Science*, *4*, 61–70.

Koopmans-van Beinum, F. J., & van der Stelt, J. M. (1986). Early stages in the development of speech movements. In B. Lindblom & R. Zetterstrom (Eds.), *Precursors of early speech* (pp. 37–50). New York: Stockton Press, Inc.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, *277*, 684–686.

Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America*, *100*, 2425–2438.

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*, 9096–9101.

Levitt, A. G., & Utman, J. A. (1992). From babbling towards the sound systems of English and French: A longitudinal two-case study. *Journal of Child Language*, *19*, 19–49.

Locke, J. L. (1994). Development of the capacity for spoken language. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language* (pp. 278–302). Oxford: Blackwell.

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, *21*, 499–546.

MacNeilage, P. F., & Davis, B. L. (2000). On the origin of internal structure of word forms. *Science*, *288*, 527–531.

Maneva, B., & Genesee, F. (2002). Bilingual babbling: Evidence for language differentiation in dual language acquisition. In B. Skarbela et al. (Eds.), *Boston University Conference on Language Development 26 Proceedings* (pp. 383–392). Somerville, MA: Cascadilla Press.

Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy, 10.*

Ménard, L., Schwartz, J., & Boë, L. (2002). Auditory normalization of French vowels synthesized by an articulatory model simulating growth from birth to adulthood. *Journal of the Acoustical Society of America*, *111*, 1892–1905.

Ménard, L., Schwartz, J., & Boë, L. (2004). Role of vocal tract morphology in speech development: Perceptual targets and sensorimotor maps for synthesized vowels from birth to adulthood. *Journal of Speech, Language, and Hearing Research*, *47*, 1059–1080.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J., & Fujimura, O. (1975). An effect of language experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English. *Perception and Psychophysics*, *18*, 331–340.

Mody, M., Schwartz, R. G., Gravel, J. S., & Ruben, R. J. (1999). Speech perception and verbal memory in children with and without histories of otitis media. *Journal of Speech, Language, and Hearing Research*, *42*, 1069–1079.

Newman, R. S. (2005). The cocktail party effect revisited: Listening to one's own name in noise. *Developmental Psychology*, *41*, 352–362.

Newman, R. S., & Jusczyk, P. W. (1996). The cocktail party effect in infants. *Perception and Psychophysics*, *58*, 1145–1156.

Nozza, R. J., Miller, S. L., Rossman, R. N. F., & Bond, L. C. (1991). Reliability and validity of infant speech-sound discrimination-in-noise thresholds. *Journal of Speech and Hearing Research*, *34*, 643–650.

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology: Vol. 1* (pp. 93–112). New York: Academic Press.

Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, NJ: Lawrence Erlbaum Associates.

Oller, D. K., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development*, *59*, 441–449.

Pallier, C., Bosch, L., & Sebastián, N. (1997). A limit on behavioral plasticity in vowel acquisition. *Cognition*, *64*, B9–B17.

Paradise, J. L., Dollaghan, C. A., Campbell, T. F., Feldman, H. M., Bernard, B. S., Colborn, K., et al. (2000). Language, speech sound production, and cognition in three-year-old children in relation to otitis media in their first three years of life. *Pediatrics*, *105*, 1119–1130.

Pegg, J. E., & Werker, J. F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America*, *102*, 3742–3753.

Polka, L., & Bohn, O. S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America*, *100*, 577–592.

Polka, L., & Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication*, *41*, 221–231.

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/–/ð/ perception: Evidence for a new developmental pattern. *Journal of the Acoustical Society of America*, *109*, 2190–2201.

Polka, L., & Rvachew, S. (2005). The impact of otitis media with effusion on infant phonetic perception. *Infancy*, *8*, 101–117.

Polka, L., Rvachew, S., & Molnar, M. (submitted). The role of attention in infant speech perception. *Cognition.*

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 421–435.

Rvachew, S., Mattock, K., Polka, L., & Ménard, L. (2006). Developmental and cross-linguistic variation in the infant vowel space: the case of Canadian English and Canadian French. *Journal of the Acoustical Society of America, 120.*

Rvachew, S., Slawinski, E. B., Williams, M., & Green, C. (1996). Formant frequencies of vowels produced by infants with and without early onset otitis media. *Canadian Acoustics*, *24*, 19–28.

Rvachew, S., Slawinski, E. B., Williams, M., & Green, C. L. (1999). The impact of early onset otitis media on babbling and early language development. *Journal of the Acoustical Society of America*, *105*, 467–475.

Stager, C. L., & Werker, J. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*, 381–382.

Stark, R. (1980). Stages of speech development in the first year of life. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology: Vol. 1* (pp. 73–92). New York: Academic Press.

Sundara, M., Polka, L., & Genesee, G. (2006). Language experience facilitates discrimination of /d–ð/ in monolingual and bilingual acquisition of English. *Cognition*, *100*, 369–388.

Sussman, H. M., Duder, C., Dalston, E., & Caciatore, A. (1999). An acoustic analysis of the development of CV coarticulation: A case study. *Journal of Speech, Language, and Hearing Research*, *42*, 1080–1096.

Sussman, H. M., Minifie, F. D., Buder, E. H., Stoel-Gammon, C., & Smith, J. (1996). Consonant–vowel interdependencies in babbling and early words: Preliminary examination of a locus equation approach. *Journal of Speech, Language, and Hearing Research*, *39*, 424–433.

Tsao, F., Liu, H.-M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, *75*, 1067–1084.

Tsao, F., Liu, H., Kuhl, P. K., & Tseng, C. (2000). *Perceptual discrimination of Mandarin fricative–affricate contrast by English-learning and Mandarin-learning infants.* Paper presented at the International Conference on Infant Studies, Brighton, UK.

Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., et al. (1994). *Discrimination of English /r–l/ and /w–y/ by Japanese infants at 6–12 months: Language-specific developmental changes in speech perception abilities.* Paper presented at the International Conference on Spoken Language Processing.

Vihman, M. M. (2002). The role of mirror neurons in the ontogeny of speech. In M. I. Stamenov & V. Gallese (Eds.), *Mirror neurons and the evolution of brain and language* (pp. 305–314). Advances in Consciousness Research: Vol. 42. Amsterdam: John Benjamins Publishing Company.

Vihman, M. M., & Nakai, S. (2003). Experimental evidence for an effect of vocal experience on infant speech perception. In M. J. Sole, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Science, Barcelona* (pp. 1017–1020). Universitat Autonoma de Barcelona.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Development and Learning*, *1*, 197–234.

Werker, J. F., Gilbert, J. H. V., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, *52*, 349–355.

Werker, J. F., Hall, G., & Fais, L. (2004). Reconstructing U-shaped functions. *Journal of Cognition and Development*, *5*, 147–151.

Werker, J. F., & Logan, J. (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics*, *37*, 35–44.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of nonnative speech sounds. *Canadian Journal of Psychology*, *37*, 278–286.

Werker, J. F., & Tees, R. C. (1984a). Cross-language speech perception: Evidence for perceptual reorganisation in the first year of life. *Infant Behavior and Development*, *7*, 49–63.

Werker, J. F., & Tees, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, *75*, 1866–1878.

Whalen, D. H., Levitt, A. G., & Wang, Q. (1992). Intonational differences between the reduplicative babbling of French- and English-learning infants. *Journal of Child Language*, *18*, 501–516.

Yont, K. M., Snow, C. E., & Vernon-Feagans, L. (2003). Is chronic otitis media associated with differences in parental input at 12 months of age? An analysis of joint attention and directives. *Applied Psycholinguistics*, *24*, 581–602.