

# Chapter 7

## Development of Speech Perception

Robin Panneton and Rochelle Newman

### 1 Introduction

During infancy, interactions with caretakers provide young humans with opportunities to discover important aspects of their world, including ways in which they and others communicate. In the current view, “communication” involves dynamic, real-time processing of multiple sources of information arising both within and between participants, directing attention, enabling co-action, and enriching understanding of the surrounding world (Hollich et al. 2000; Locke 2001). Early human communication can take many forms, including the perception and production of music (e.g., maternal singing; see Trainor and Unrau, Chap. 8) and gesture, and clearly (for most children) involves their mastery of a native language system. The pathway to language in human infancy is not a quick one, but it is remarkably robust in that most children find their way to being fully communicative by the age of 3 years (Jusczyk 1997). That being said, there are important exceptions to this statement in that not all children reach the same level of language proficiency (Klee et al. 2000; see also Eisenberg et al., Chap. 9), with some showing marked deficiencies in communication by early toddlerhood, as in severe autism. In the end, developmental research must account for the full spectrum of functioning in this important domain. This is even more pertinent for practitioners who must bring resources to bear on children who are challenged in their abilities to communicate fully and effectively with others.

---

R. Panneton (✉)  
Department of Psychology, Virginia Tech, Blacksburg, VA, USA  
e-mail: panneton@vt.edu

R. Newman  
Department of Hearing & Speech Sciences, University of Maryland, College Park, MD, USA  
e-mail: newman1@umd.edu

This chapter provides an overview of the general time course for emerging skills related to the perception and production of speech across infancy and early childhood. The presentation begins with a brief description of methodologies used to assess the skills and deficits seen over the course of early language learning. Next, language learning is set within a motivational framework based on infants' early attentional focus; success in language learning critically depends on the ability of infants and young children to direct and maintain attention to sources providing information about their language systems. The ability to regulate infants' attention in language-related tasks is an essential aspect of research in this field. But beyond the laboratory, it is important to appreciate that language learning arises out of dynamic partnerships involving real-time adjustments in attention flow as relevant properties unfold and partners "communicate." Language learning occurs in such relationships. Subsequent sections summarize research that creates a portrait of the typical language-learning infant as one moving from a diffuse, superficial perceiver to a more highly focused and selective consumer of language. Last, the importance of attention regulation is emphasized in the context of challenges that infants face in natural language-learning conditions.

## **2 Measuring Speech Perception and Language Skills in Infancy**

Infants are not always cooperative or accessible research participants. Nonetheless, great strides have been achieved in understanding early auditory psychophysics, perception, and language (Saffran et al. 2006), and in the ability of developmental scientists to probe infants for various abilities or limitations. Infants display an array of perceptual and cognitive skills that are themselves undergoing rapid development (e.g., statistical learning) and emerging in real time (i.e., participating in experimental protocols engenders learning in the moment). Developmental research on infants' perception of speech continues to advance in important ways as both methods and techniques evolve. Several excellent discussions are available to educate new workers in this field (McCardle et al. 2009; Johnson and Zamuner 2010), so the following is a selective summary of common practices and marked advances.

Research with young infants capitalizes on rudimentary motor responses (e.g., sucking on pacifiers, turning supported heads toward a sound) and responsive physiology (e.g., heart rate changes; cortical activation patterns). With age, infants become proficient at controlling their own body movements, allowing protocols to include volitional action (e.g., turning one's own head for specific consequences; Werker et al. 1997). Two common protocols for measuring speech perception involve either selective visual fixation of, or head-turn toward targets associated with speech (Johnson and Zamuner 2010). Visual fixation studies are designed to examine infants' speech preferences or speech discrimination. For example, some are designed such that fixation of a repeated visual target produces two kinds of speech streams, allowing investigators to gauge speech preferences by comparing

differences in looking times (Panneton-Cooper and Aslin 1990; Pegg et al. 1992). Other studies involve periods of familiarization (or habituation) during which fixation of a repeated visual target produces the same speech event, followed by discrete trials during which fixation produces either familiar or novel speech. Often, infants increase attention (i.e., look longer at the visual target) during novel trials, indicating speech discrimination (Fais et al. 2009; Sato et al. 2010).

Although the visual fixation technique works well across many ages, it is particularly well suited for younger infants given that it does not require head turns. Speech preference and discrimination studies with older infants are more likely to involve infants' head-turning to peripherally located lights, and involve some period of familiarization during which infants hear discrete ("cup") or fluent ("where is the cup?") speech. Next, infants hear both familiar and novel speech events and looking time is compared. This technique is typically referred to as the "head-turn preference procedure."

In spite of the utility in these behavioral techniques, three major methodological issues have surfaced (Aslin 2007). First, findings are typically group-level, and not always reflective of individual infants' performances. In the end, this compromises the predictive validity of research findings with respect to emerging language function, a problem for basic and clinical interests alike. Houston et al. (2007) recently addressed this concern by empirically validating a hybrid methodology for testing infants' speech discrimination by integrating robust elements from multiple protocols into one. After infants habituated to a nonsense word (e.g., "boodup"), they received a series of test trials including familiar ("boodup") and novel ("seepug") words, with two innovative design features. First, the ratio of novel to familiar trials was low (e.g., 3/14), increasing the saliency of novel presentations in the stream of test trials, as in classic "oddball" paradigms. Second, the two word types alternated during novel trials ("boodup/seepug/boodup/seepug...") decreasing the cognitive load for discrimination. Analyses of individual infants' data showed significantly higher rates of discrimination compared to other conditions including one feature but not the other. Moreover, these authors also found significant positive correlations between individual infants' performance across 2–3 days. Thus, the hybrid discrimination protocol promotes better internal as well as predictive validity.

A second methodological concern is the inconsistency of expected outcomes in terms of infants showing more attention to familiar versus novel information during test. Some infant speech perception studies find more attention to familiar than novel presentations, whereas other studies find more attention to novel than familiar presentations. Either outcome demonstrates discrimination, although whether attention is enhanced by the familiar or the novel affects conceptual interpretations (Houston-Price and Nakai 2004; Aslin and Fiser 2005). A more serious concern is when both patterns appear in a given sample (i.e., some infants prefer familiar speech whereas others prefer novel speech) leading to a null effect (and erroneous interpretations) at the group level, when discrimination was evident at the individual level (see Houston et al. 2007 for a discussion on this point). Multiple factors determine whether any given infant attends more to familiar or novel test events, such as stimulus complexity and individual processing strategies. Moreover, individual infants' attention to familiar versus novel information is related to concurrent and long-term measures of sustained attention and recognition memory (Colombo 2001).

Third, most protocols used to investigate infants' speech perception do not provide graded responses. For example, in a segmentation task, the primary measure of interest is whether infants attend longer on familiar or novel trials (an either/or measure). Typically, there is no quantification of response strength; even if all infants look more at the familiar event, do some infants look longer (or faster) than others, or show a stronger effect in one task than another? Estimating strength of preference or discrimination could clarify whether some cues to segmentation make the task easier or harder than others. Improved measurement may emerge from advances in using eye-tracking systems to assess the speed, direction, and duration of infants' fixations on visual targets that have been associated with verbal labels (e.g., McMurray and Aslin 2004; Fernald et al. 2008).

A complementary approach to the use of behavioral protocols for understanding speech perception in infants is provided by physiological techniques (e.g., autonomic responses such as heart rate or central responses such as brain activity), some of which can stand alone as measures of processing or work in conjunction with behavioral tasks. Researchers also use different scalp-level recording methods to specify cortical and subcortical involvement in early language processing (Friederici and Oberecker 2008). Brain-relevant recording procedures continue to be refined for infants and young children, such as scalp electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and magnetoencephalography (MEG). One emerging technology involves the use of near-infrared spectroscopy (NIRS), a technique measuring changes in hemodynamic flow in cortex as a function of event exposure (Mehler et al. 2008; Lloyd-Fox et al. 2010). The appeal of NIRS stems in part from its greater cortical source precision compared to other scalp recording systems, and from the fact that it can be used in awake, alert infants (Aslin and Mehler 2005). Recent studies employing fMRI or NIRS with infants across a wide age span have shown increased responding in temporal cortex to voices (e.g., superior temporal sulcus), either presented alone or accompanied by faces, along with patterns of lateralization that reveal the early specialization of information processing in development (Dehaene-Lambertz et al. 2010; Grossmann et al. 2010). Perhaps the need for the aforementioned graded measurements will be achieved by coupling the use of behavioral tasks with psychophysiology (e.g., heart-rate change, amplitudes of deflections in ERP waveforms; see Junge et al. 2010).

In the meantime, developmental studies on speech perception depend importantly on cross-laboratory confirmation of both positive and negative findings related to emerging skills and how they correspond to later levels of language functioning. Although this kind of collaborative validity is ideal, some important questions remain unverified when participant groups are harder to recruit or techniques are more difficult. A key element in the success of developmental studies with infants is eliciting and maintaining task-related attention. Laboratory studies (c.f. Kuhl et al. 2003; Goldstein and Schwade 2008) demonstrate that infant learning is dependent on gaining infants' attention. Even more so, language learning outside of the laboratory demands recruiting and sustaining attention in situations that involve multiple sources of information. For that reason, the next section of the chapter offers a discussion of the context within which language learning occurs during infancy, both in and outside the laboratory.

### 3 Attention to Language Function in Infancy

One robust finding in the early language-learning literature is that caretakers around the world modify their communicative style when addressing infants. These modifications are dynamic, in that they are tuned both to the sociolinguistic development of the child and to the context in which the interaction occurs (Panneton-Cooper 1993; Kitamura and Burnham 2003). Importantly, some aspects of these changes recruit and maintain attention to language, occurring not only in caretakers' vocal acoustics, but in their facial expressions and gestures as well.

Infant-directed speech (IDS) is distinguished acoustically by higher pitch (fundamental frequency or F0), more exaggerated pitch contours, larger pitch range (the difference between F0 maximum and minimum), slower tempo, longer pauses, and higher rhythmicity than adult-directed speech (ADS; Fernald and Mazzie 1991; Katz et al. 1996). Such prosodic exaggeration most likely scaffolds acquisition of linguistic structure in infants as it appears to do in adults (Golinkoff and Alioto 1995). Developmentally, even newborn infants prefer IDS when the alternative is ADS (Panneton-Cooper and Aslin 1990; Pegg et al. 1992). In addition, newborns prefer recordings of their mothers' voices compared to those of unfamiliar females (DeCasper and Fifer 1980), with evidence that such learning is influenced by prenatal experience (DeCasper and Spence 1986; Mehler et al. 1988). Integrating across these early biases leads to the prediction that newborns' attention would be maximally heightened by recordings of maternal IDS. However, young infants (1-month-olds) do not prefer IDS to ADS when both are spoken by their own mothers, although this preference is evident if *nonmaternal* recordings are used (Panneton-Cooper et al. 1997). This finding supports the view that the recruitment of infants' attention to language is heightened by whatever context is most familiar and meaningful to infants at that developmental time. For newborns, the maternal voice is a primary attractor; with more experience, preference for the maternal voice becomes refined to include the kinds of acoustic exaggerations that are shown in caretaking episodes around the world (Fernald et al. 1989). In fact, 4-month-olds do prefer maternal IDS over maternal ADS (Panneton-Cooper et al. 1997) as well as IDS over ADS when spoken by unfamiliar females (Fernald 1985), suggesting that early in the first year, infant–mother exchanges promote the extension of infants' selective attention to speakers and speaking style.

Differential attention to IDS persists at least until the last quarter of the first postnatal year with the acoustic and lexical characteristics of IDS shifting as infants' abilities improve and parental communicative intentions become more complex (Kitamura and Burnham 2003). However, the ability of vocal prosody alone to promote infants' attention appears to diminish with age (Hayashi et al. 2001; Newman and Hussain 2006), suggesting that the contribution of IDS to language learning is more evident in early rather than later infancy.

Early on, IDS most likely garners its perceptual pull in moderating infants' attention from its association with vocal emotion (Kitamura and Burnham 1998; Singh et al. 2002) rather than from the fact that it is speech to infants per se. Several researchers (e.g., Trainor et al. 2000; Santesso et al. 2007) argue that infants are drawn toward

IDS primarily because of its exaggerated emotional tone, which then supports language acquisition. As such, research on IDS has shifted from an analysis of acoustic properties to the affective components that underlie infant preferences. Exaggerated IDS prosody typically reflects vocal expression of emotion, such that the ability of IDS to regulate infant attention may stem from emotional engagement (Kitamura and Burnham 1998). Infants prefer positive over negative vocal affect at least through the first half of the first year (Panneton et al. 2006). Infants also discriminate specific categories of IDS at different points during development, with those signaling emotional intent more preferred by younger, but not by older infants (Moore et al. 1997; Kitamura and Lam 2009). In the adult literature, emotional tone of voice enhances processing of lexically ambiguous words (Nygaard and Lunders 2002) and improves word recognition memory (Dietrich et al. 2000), even though the necessary and sufficient acoustic correlates of vocal emotion that aid perception remain unclear (Scherer 1995).

Importantly, the ability of positive-emotion speech to increase infants' attention is contingent on the nature and quality of exchanges between infant and caretaker. As is clear from studies looking at infant–caretaker dynamics (Stern 1985), early communication is bidirectional in nature, and the ability to increase and maintain infant attention is compromised when contingency is diminished, even in the face of positive emotion in voice and other gestures. An interesting study in this regard involved recording mothers' speech to their infants while they both interacted over a monitor (i.e., the mother and infant were in separate rooms). Mothers were asked to use their voices to increase positive emotion in their infants, only unbeknownst to the mothers the infants could not hear them. A confederate female either engaged the infants positively or negatively, whenever their mothers spoke. Acoustic analyses of the mothers' voices indicated significant elevation of pitch and pitch variance in the group whose infants received surreptitious positive regard, supporting the notion that vocal adjustments in caretakers' speech are highly influenced by contingent adjustments in infant behaviors during interaction (Smith and Trainor 2008). Thus, infants' primary motivation for attending to information in the speech stream arises out of the dynamic exchange known as the infant-directed context. So what is learned in these ongoing interactions?

#### **4 Attention to Language Structure in Infancy and Early Childhood**

Within the early context of heightened attention to IDS, perceptual shaping of attention to language takes root. One clear benefit of infants' heightened attention to IDS is access to important aspects of native language structure (Fisher and Tokura 1996; Christophe et al. 2003). Considerable research effort has focused on how IDS bootstraps infants' early lexical and syntactic awareness. Prosodic information plays a vital role even in adult speech processing, especially in English where intonational emphasis is used to accent novel information (Gerken 1996). Boundaries of both

sentences and restrictive relative clauses in English are marked by falling pitch contours and final vowel lengthening, and words (even syllables) conveying new information are higher pitched and longer. Infants take advantage of prosody in learning how to segment and categorize speech input (Gerken et al. 2005; Thiessen et al. 2005), and in syntactic acquisition (Soderstrom et al. 2003).

#### ***4.1 In the Beginning: Perceptual Shaping of Attention to Speech in the First Six Months***

In general, good empirical support exists for the ability of young infants to attend to and recognize various aspects of human speech (e.g., preferences for the native language in newborns, Moon et al. 1993), prompting focus on aspects of languages that lend to their uniqueness. Languages differ from one another in a number of ways, but many of those differences are in suprasegmental properties, particularly in terms of prosody and rhythm, and these seem particularly critical for infant language preferences. Languages can be classified into three main categories on the basis of their general rhythmic structure (Pike 1945; Abercrombie 1967), providing one potential cue to infants as to native versus nonnative designation. Languages are viewed as “syllable-timed” in which each syllable has an equivalent duration in production, “stress-timed” in which time between stressed syllables is more-or-less constant and unstressed syllables are shortened to fit between the stressed-syllable beats, or “morae-timed” in which syllables consist of either one or two subsyllabic durational units, called morae. Classifying languages into rhythmic groups is now viewed as overly simplistic given that languages exist along a continuum of timing patterns, and languages within a “category” may still differ from one another rhythmically. Nonetheless, this basic rhythmic distinction is readily perceived by adults, and even by nonhuman primates when most of the segmental information has been removed (Ramus et al. 2000). Infants’ familiarity with a particular language rhythm may enhance discrimination and provide an early basis for native language preferences.

As mentioned previously, newborn humans prefer their native language by as young as 2 days of age (Mehler et al. 1988; Moon et al. 1993), demonstrating the ability to distinguish between languages, and appear particularly sensitive to language rhythm. Both 5-month-old infants and newborns discriminate two unknown languages when they fall into different rhythmic classes (Nazzi et al. 1998), but fail to do so when the languages come from the same rhythmic class (e.g., English vs. Dutch), unless one is native (Nazzi et al. 2000). However, Christophe and Morton (1998) found that 2-month-olds did not discriminate languages from different rhythm classes, either indicating a U-shaped developmental trend, or resulting from methodological differences across studies.

Thus, discriminating languages within a rhythm class depends on whether infants are highly familiar with one. Similarly, infants discriminate two dialectal versions of their native language, but only if one of the dialects is familiar (Nazzi et al. 2000; Butler et al. 2010). It is not clear how this process works. One possibility is that

infants gradually refine discrimination within a rhythmic class, moving from reliance on gross prosodic differences to more fine-grained differences, based on experience with their native language (Nazzi et al. 2000). However, interpreting these patterns of results is complicated by the methodological differences across studies and the lack of longitudinal approaches in this area of work. Another way to address this issue developmentally is to examine bilingual infants, given the demand on these infants to discriminate two native systems. Although scarce, available results suggest that bilingual infants discriminate their two languages by 5 months, even when the language systems fall into the same rhythmic class (Bosch and Sebastián-Gallés 2001). How consistent this discrimination ability is across infants or across languages has not been studied, despite frequent concerns from parents regarding the benefits versus risks of raising children bilingually. Interestingly, one recent study found differences between mono- and bilingual infants' discrimination of native from nonnative *visual* speech, an area of interest given that prosody and rhythm in language can not only be heard, but can also be seen (Munhall et al. 2004). As speakers engage in conversation, they move their facial muscles, heads, and bodies in ways that spatiotemporally correlate with their speech. Weikum et al. (2007) found that 4- and 6-month-old English-learning infants discriminated silent videotapes of women speaking English versus French, suggesting that infants attune to how speech is *visually* represented in human facial movement. In contrast, 8-month-olds did not make this discrimination across visual languages unless they were English–French bilinguals.

Although newborns' preference for natural recordings of their native language may be primarily based on rhythm (Moon et al. 1993), newborns also show preferences for canonical consonant-vowel compounds (compared to synthetic analogs; Vouloumanos and Werker 2007). Extending to slightly older infants, Panneton-Cooper and Aslin (1994) also found preferences for normal IDS speech over sine-wave analogs of IDS in 4-month-olds. These preferences may reflect early bias toward natural speech, and are supported by neurophysiological studies on early sensitivity to language-specific information (Peña et al. 2003). This early perceptual shaping of preferences for suprasegmental aspects of native languages is complemented by attention to more fine-grained segments, such as consonants and vowels. Perception of microstructure in speech led to early studies examining infants' categorical perception of phonemes. Categorical perception is seen when listeners are poor at distinguishing sounds from the same phonetic category (e.g., different tokens of the sound “b”) but successful at distinguishing sounds from different phonetic categories (e.g., “b” from “d”). Categorical perception poses a puzzle for development because phonetic categories are language specific; the same acoustic difference could occur within a category in one language (and be ignored by native adult listeners) but could signal an important phonetic distinction in another language. Thus, categorical perception cannot be fully “set” from birth, implying that during development infant listeners must either learn to distinguish phonetic categories or learn to ignore the differences within categories.

Remarkably, categorical perception of phonemes appears early in infancy. Eimas et al. (1971) presented 1- and 4-month-old infants with a synthetic CV syllable,



either /ba/ or /pa/. After the infants' attention habituated to the item, infants either heard further examples of the same item (control group), a switch to a new token within the same category, or a switch to a token from the opposite category. Critically, the two "switch" tokens were equivalently dissimilar acoustically from the original item. Infants' attention showed significantly greater recovery for the item in the opposite category, indicating greater discrimination of items across a category boundary than within a category. Follow-up studies reported similar results for a range of phoneme distinctions, in a range of syllabic contexts (e.g., Eimas and Miller 1980; Cohen et al. 1992). Importantly, these studies also showed that infants discriminated contrasts that did not occur in their native language, contrasts adult speakers of the same language failed to discriminate (e.g., Trehub 1976). Infants also failed to distinguish some contrasts that were within their native language (Lasky et al. 1975). Thus infants' phoneme discrimination appears to be language-universal at first, becoming increasingly attuned to the native system as infants gain experience with hearing speech.

#### ***4.2 Second Half of the First Year: Perceptual Attunement and Attentional Pruning***

In the first 6 months of life, infants focus their attention on rhythmic and prosodic aspects of language. In contrast, their initial perception of segmental properties appears to be less colored by their native language. Comparatively, attention pruning dominates speech perception during the next 6 months, with infants attending more selectively to finer-grained aspects of their native language. Progressive attunement to experientially dependent aspects of information across infancy may extend to other domains as well (e.g., face processing; Scott et al. 2007).

Werker and Tees (1984) explored the time course of increasing attunement to native phonemes across infancy. They tested English-learning infants ages 6–8 months, 8–10 months, and 10–12 months on two nonnative contrasts, from Hindi and Thompson/Nlaka'pamux, that adult English-speakers failed to distinguish. The youngest infants discriminated both contrasts, but older infants failed to do so, suggesting that there was a shift in sensitivity to nonnative phonemes between 8 and 10 months of age. However, not all nonnative contrasts present this difficulty for older infants. Best et al. (1988) found no attenuation of discrimination in English infants for Zulu clicks (see also McMurray and Aslin 2005). Nonetheless, changes in discrimination have been found at roughly the same age for a variety of segmental distinctions (Polka and Werker 1994; Tsao et al. 2006), including both declines in the discrimination of nonnative contrasts, and improvements in, or in some cases development of, discrimination of native contrasts (Hoonhorst et al. 2009; Sato et al. 2010).

Languages differ not only in their phonemic patterns, but also in their use of vocal tone. According to some estimates, half the people in the world (and perhaps as many as 70%) speak a tone language (Yip 2002), in which differences in pitch shape (i.e., fundamental frequency) serve to distinguish meanings, much the same

way phonetic differences do. Currently, little is known about how tonal languages are perceived by infants (Yip 2002). Studies of infants learning both Chinese (Mattock and Burnham 2006) and Yorùbá, a tonal language of west Africa (Harrison 2000) found that infants respond to changes in tone across syllables that are phonemic in function, and appear to do so in a roughly categorical manner. In contrast, infants learning English show early discrimination of tonemes, but reduced discrimination by 9 months of age (Mattock et al. 2008), corroborating similar findings that infants' attention to nonnative segmental distinctions diminishes as experience with the native language increases.

Thus, there is strong evidence for some form of attunement to one's native language during the second half of the first postnatal year (often referred to as "perceptual reorganization"). As infants gain more experience with their native language, their perception (and production; Boysson-Bardies et al. 1984) both change, molding themselves to the typical input. This learning takes place within the attention framework of IDS, in that IDS often consists of elongated and more clearly enunciated vowels (e.g., Kuhl et al. 1997), and the degree of hyperarticulation (i.e., vocal clarity) present in mothers' speech positively correlates with infants' performance on a phoneme discrimination task as well as their vocabulary growth (Liu et al. 2003). Exaggerated pitch contours alone appear to facilitate vowel discrimination in 6–7-month-olds (Trainor and Desjardins 2002). Thus infants' inclination to attend selectively to IDS may enhance their learning of the segmental properties of fluent speech, which then leads to better attention to the input, and better subsequent learning of other important components of their native language.

What of phonemic attunement in infants raised in multilingual environments? Bosch and Sebastián-Gallés (2003) compared monolingual Spanish, monolingual Catalan, and bilingual Spanish/Catalan infants on a phonetic contrast (/e-eh/) that occurs in Catalan, but not Spanish. Four-month-olds discriminated the contrast regardless of which language(s) they learned in the home, but by 8 months only monolingual Catalan infants made the distinction. Diminished discrimination was expected in the monolingual Spanish infants, but surprising in the bilingual infants, for whom the contrast was still relevant. In a follow-up study, 12-month-old bilingual infants regained discrimination of the contrast, suggesting that the time course for attunement might be different in infants learning more than one language (see also Sebastián-Gallés and Bosch 2009). In contrast, other studies with bilingual infants show a time frame for language-specific phonetic reorganization similar to that for monolinguals (Burns et al. 2007; Sundara et al. 2008). It is unclear whether these inconsistencies are the result of the methods used, the specific languages or phonetic contrasts tested, or some other difference. Sundara and Scutellaro (2010) suggest that infants learning two rhythmically different languages may perform more similarly to monolinguals than do infants learning two rhythmically similar languages, perhaps because infants can use rhythm as a means of sorting the input, thus avoiding confusion. However, additional work comparing monolingual and bilingual populations is needed, particularly given the high rate of bilingualism across the world.

The mechanism for phonemic attunement derives from infants' perception of whether a particular sound distinction makes a meaningful difference in the

language, rather than being based on whether the sound distinction is heard in the ambient surroundings. For example, many English speakers produce prevoicing in their stop consonants, but this does not affect the interpretation of the words, and thus prevoiced stop consonants are not treated as a separate sound category. So how do infants learn which sounds are linguistically important? This is perplexing given that native attunement occurs when infants do not yet know many words, and thus could not know which phonetic distinctions make lexically-important distinctions in their language. Maye et al. (2002) suggest that the statistical distribution of sounds provides information as to their importance. More specifically, if variation along some acoustic measure was important in the language, the distribution of sounds would form a bimodal pattern: speakers would avoid producing tokens that fell at the category boundary, and productions would instead cluster around two (or more) distinct category centers or prototypes. In contrast, if variation along a given acoustic measure was irrelevant, speakers' productions would form a unimodal distribution. Maye et al. (2002) first familiarized two separate groups of infants with speech sounds from either a unimodal or bimodal distribution, then tested infants' discrimination of tokens along the entire continuum. Only the infants who had been familiarized with a bimodal distribution of sounds discriminated changes in speech tokens, suggesting that infants track the distribution of sounds they hear. This kind of sensitivity to distributional properties of speech provides a potential mechanism whereby attunement could occur.

Interestingly, distributional learning of phoneme categories in infants appears to also be influenced by visual speech information (Teinonen et al. 2008). Given that distributional cues for phonetic categories occur in the input (Werker et al. 2007) and that these are learnable by computer models (Vallabha et al. 2007), they could presumably be learned by infants as well. That said, it remains unclear whether speech input to infants is consistent in this respect for all phoneme categories and for suprasegmental distinctions, such as tone categories. Determining the number of tone categories in a given language is a critical prerequisite for learning the language, and determining whether the distributional properties would signal this to infants is of particular importance. It is also unclear how well infants' ability to track such distributions in a laboratory setting will "scale up" to more real-world settings, or whether there might be individual differences in this ability.

In addition to learning the phonetic structure of their native language, infants also learn *patterns* of phoneme distributions, referred to as phonotactics, toward the end of their first postnatal year. Jusczyk et al. (1994) showed that 9-month-olds, but not 6-month-olds, listened longer to lists of items that had more common phonetic patterns than to lists with less common patterns. However, the high-probability sequences contained both high positional phoneme frequency (that is, the segments were common in that word position), and contained more common phoneme combinations (or biphones) and higher-probability phonemes. Thus, there were multiple forms of phonotactic information that infants could have perceived from the input, and it is not clear which drove infants' preferences, or whether infants are equally sensitive to them at this age. In addition, the items in this study were entirely CVC tokens; thus aspects of phonotactics having to do with consonant clusters were not

investigated. Languages differ substantially in the number and types of consonant clusters allowed, from languages that forbid all consonant clusters within a syllable (such as Japanese) to languages that allow large strings (such as Slovak in which there is an entire tongue twister without vowels, *strc prst skrz krk*; Hanulíková et al. 2010), and infants might be expected to pick up on this aspect of phonotactics quite early. Although infants become aware of phonotactic probabilities during their first year of life, the specifics of how this process occurs, or how gradual this acquisition might be is unknown; it seems likely that the acquisition of phonotactic biases takes place along a continuum, with some aspects being recognized earlier than others, and there may be limits on which statistics infants will track. Little research has attempted to compare different types of statistical patterns, or to examine how infants determine which statistical computations might be relevant (Soderstrom et al. 2009).

Finally, it is not clear *how* infants track such patterns. One possibility is that infants store patterns of input, comparing new input to the combined set of prior exemplars. Preferences for more common statistical patterns would arise out of the process of recognition, with infants storing large amounts of relatively unprocessed data. An alternative approach is that infants track patterns during their original perception, and store these outcomes rather than the raw data. Certain statistical patterns would be more likely to be observed than others, since untracked properties could not emerge at a later date. This distinction is akin to that of prototype versus exemplars in categorization, but has received less attention in the infant language literature (c.f. Polka and Bohn 2011 for an excellent conceptualization of this prototype model for speech perception).

Although segmental changes may be evidence of a focusing of infant attention, the underlying goal of communication is not to discriminate phonemes but instead to understand the speaker's intention, which requires recognizing meaningful units (words). Most of the speech that infants hear consists of multiword utterances without obvious pauses or breaks demarcating boundaries; learning from this requires that infants first separate the fluent speech into individual words, a task referred to as word segmentation. Jusczyk and Aslin (1995) developed the first paradigm for evaluating the development of this segmentation ability. Using the head-turn preference procedure discussed earlier, these authors familiarized infants with two target words (either *cup* and *dog*, or *bike* and *feet*), spoken in isolation. Next, infants were presented with four fluent speech passages, each of which contained one of the potential target words. Infants age 7.5 months, but not those age 6 months, listened longer to the passages containing familiarized words, showing evidence of segmentation.

More recent work has shown that the precise age at which segmentation first occurs depends on both the method of testing (e.g., ERP results demonstrate slightly earlier segmentation ability than that shown behaviorally; Kooijman et al. 2005), and the type of lexical unit (words beginning with vowels, words with atypical stress patterns, and verbs show slightly later segmentation abilities; Mattys and Jusczyk 2001; Nazzi et al. 2005). However, across these studies, the ability to segment is consistently shown to develop as infancy advances. Thus, although the exact age varies, the general pattern is consistent—segmentation precedes most types of

word learning. Moreover, infants who show stronger segmentation skills also show enhanced language-learning skills at later ages (Newman et al. 2006; Junge et al. 2010), supporting the notion that this ability to segment may be a critical skill underpinning language acquisition.

As older infants' skills at segmentation increase, they begin to rely on a greater variety of cues to segment successfully. One knowledge-based cue promoting segmentation is word familiarity (sometimes referred to as *segmentation by lexical subtraction*, as in White et al. 2010). Words that are highly familiar to infants promote segmentation of adjacent words, even if those are less familiar. For example, infants can segment words adjacent to the word "mommy" (*Mommy's feet* vs. *Lola's dog* produces better segmentation of *feet*; Bortfeld et al. 2005). Another knowledge-based cue to segmentation is infants' ability to use statistical patterns in the spoken input. Certain phonemes are more likely to occur in particular word positions (e.g., "ng" in English only occurs syllable-finally), and certain pairs of syllables are more likely to occur together (as part of the same word) than others (e.g., /fnt/is more likely to occur subsequent to /In/ than as the start of a new word). Computer models tracking such probabilistic patterns consistently identify word boundaries (Brent and Cartwright 1996; Christiansen et al. 1998). Importantly, Saffran et al. (1996) demonstrated that infants use statistical cues in segmentation; after hearing a 2-min stream of continuous speech in which some syllables were adjacent to one another more regularly than were others, infants listened longer to atypical syllable combinations (those that had occurred adjacently less frequently) than to more common syllable combinations. This suggests infants treat high-probability sequences as potential new words. Interestingly, infants are more successful at linking these high-probability sequences with objects than they are at linking low-probability sequences with objects (Graf Estes et al. 2007).

Infants also use a wide variety of acoustic cues in segmentation, such as prosodic or stress cues, coarticulatory cues, phonotactic cues, allophonic cues, and cues to phonological phrase boundaries (see Saffran et al. 2006, for a review), although none of these cues provides definitive information in all settings. Another aid to segmentation is word position within an utterance, as 8-month-old infants are better able to segment utterance-initial or utterance-final words than utterance-medial words (Seidl and Johnson 2006). Overall, infants perform better at segmentation tasks when words have been "partially segmented" for them, and the process of segmentation appears to be one of integrating a wide array of probabilistic sources of information.

Given that segmentation cues appear to become available to infants at different stages in development, infants' weighting of potential cues also changes as they gain more experience with their native language (e.g., Johnson and Jusczyk 2001; Thiessen and Saffran 2003). In general, infants appear to move from depending on syllabic properties, such as statistical regularities of syllables and lexical stress, to being able to use more detailed segmental information, in the form of phonotactic and allophonic cues. Mattys et al. (2005) proposed one detailed cue hierarchy, but arguments over the relative weightings of different cues remain and these could vary across languages. As infants become more efficient in their processing, they are able

to integrate more types of information simultaneously (Morgan and Saffran 1995). As a result of these changes, the ability to segment the speech stream has a drawn-out developmental time course.

There are several important issues about the development of segmentation skills that remain virtually unaddressed. First, the vast majority of research on segmentation has focused on either English or other Germanic languages. Cross-linguistic differences in syllabic structure or in the use of affixes is likely to influence segmentation strategies, but only a few studies have explored segmentation in other European languages, let alone languages that may be more dissimilar either acoustically or structurally. Second, the literature on early segmentation has arisen from the study of infants raised in a monolingual environment. Many of the cues for segmentation, such as stress patterns, are language specific, and segmentation processes may be quite different for infants learning multiple languages simultaneously. Third, most of the research on segmentation has focused on finding the youngest age at which infants could reliably succeed on a task. Presumably, the ability to segment speech may not be a skill that infants either have or do not have, but may instead be a skill that they continue to develop over an extended time frame, and far less work has explored this developmental progression. Finally, there are a few studies that have suggested that infants who demonstrate more advanced segmentation skills are likely to show enhanced language-learning skills at later ages, perhaps because their early segmentation ability has provided them with more opportunities to learn words and morphemes. Although segmentation has been shown to be delayed in children with Williams syndrome (Nazzi et al. 2003), most research has focused on variation within a typically developing cohort; testing segmentation abilities in at-risk populations will be particularly fruitful for the early identification of language delay or disorder.

### ***4.3 Is Speech Perception in Infancy Related to Emerging Language Proficiency after the First Year?***

As is evident in Sects. 4.1 and 4.2, a great deal of research has focused on infants' perception of various cues to language structure between 6 and 12 months of age (i.e., parsing/segmenting words sensitivity to conditional probability of adjacent units). Clearly, infants are progressively attuning perception to speech information such that they are primed for processing within native-language contexts (Kuhl 2007). As a result, one would predict older infants and toddlers would show impressive competencies in their initial grammatical constructions (e.g., rich and flexible surface structure) as they begin to communicate. Although pattern recognition is a first step in language learning, toddlers who attempt to communicate with others need to invoke such patterns in context, and construct utterances with them in ways that accurately reflect intention and meaning. In the process of communicating, one can afford to minimize and even omit a fair amount of structure as long as meaning and intention are preserved. The important question at stake for those interested in

language trajectories, then, is whether there is any demonstrable link between the perceptual acumen of infants and their subsequent communicative skills.

One interesting paradox between the literatures on speech perception in infancy and lexical/grammatical development in toddlerhood is that toddlers show lower levels of grammatical/syntactic sophistication than expected. That is, the productive strategies that characterize toddlers' first attempts at communicating do not seem to include many of the lexical units or relations between units that they perceived during infancy. Children's vocabulary certainly makes rapid advances after the first postnatal year, but it is not clear how infants' perception of form and function translate into their own productive strategies later on. At least one study has found promising predictive validity with the head-turn preference procedure. Newman et al. (2006) analyzed the relationship between infants' segmentation of words from fluent speech and subsequent measures of language proficiency (e.g., vocabulary size). Aggregate data from several studies showed that after familiarizing 7.5- to 8-month-olds to single words, infants (as a group) attended more to sentences that contained those same words, than to those composed of novel words. Importantly, a longitudinal follow-up of individual infants from these studies revealed that infants who performed *poorly* in this word segmentation task also showed significantly lower vocabulary sizes at 24 months.

Although a complete treatment of this paradox is beyond the scope of the current chapter, readers are directed to a thoughtful discussion by Naigles (2002), in which she argues that speech perception in infancy has much less to do with *meaning* than does the active utterance construction by young children. That is, infants operate on continuous speech from the perspective of naïve pattern recognition (e.g., adjacent phonemes with higher conditional probabilities are recognized, processed, and retrieved as clusters), rather than from the perspective of what a given cluster actually denotes about an object or object relations. An excellent example of this comes from studies using the "switch paradigm" to study object–word association learning. In one case, infants at age 14 months readily learn the associations between ObjectA-LabelA and ObjectB-LabelB when the objects and labels are novel, and when the labels are maximally contrastive (e.g., "neem" vs. "lif"; Werker et al. 1998). However, same-age infants do not show evidence of object–word association learning when the labels involve minimal-pair distinctions (i.e., "bin" vs. "din"; Stager and Werker 1997; c.f. Fennell and Werker 2003), even though younger infants have no difficulty discriminating "bin" from "din." Thus, although infants at this age can distinguish minimal pair contrasts, the increased cognitive demand of object–word pairing makes the minimal pair contrast harder to process. In this example, knowing something about younger infants' perceptual acuity does not lead to accurate predictions about their later ability to apply this skill in a learning context.

As speech perception and discrimination abilities continue to evolve, toddlers make significant gains in information processing during their 2nd/3rd years. Bernhardt et al. (2007) found significant, positive correlations between 14-month-olds' word–object association learning and measures of their expressive and productive language skills up to 2 years later. Two-year-olds, compared to 15-month-olds, more accurately and rapidly locate named objects without even hearing entire verbal

labels, or when the labels themselves are acoustically degraded (Fernald et al. 2001; Zangl et al. 2005). Such processing efficiency correlates positively with better lexical and grammatical skill over age, not only in absolute performance, but in growth trajectories as well (Fernald et al. 2006). Moreover, speed of spoken word recognition and vocabulary size during toddlerhood correlate positively with multiple indices of linguistic skill (e.g., expressive vocabulary, formulating sentences, word structure) and working memory performance at 8 years of age (Marchman and Fernald 2008). Thus, the speed with which older infants and toddlers process continuous speech is one factor that plays an important role in emerging linguistic competency across childhood.

Nonetheless, an important tension exists between being fast and being accurate in the continuous perception and production of speech. Other studies have concentrated on identifying mediating factors for toddlers' comprehension of spoken words. For example, 11-month-olds recognize familiar words, but only if their onset and offset phonemes remain intact [e.g., infants recognize the familiar word "dirty" but do not recognize the similar nonwords "nirty" (onset violation) or "dirny" (offset violation); Vihman et al. 2004; Swingley 2005]. Such sensitivity to perceptually salient portions of words indicates that even as toddlers increase their speed of processing, violations of expected phonotactic sequences can attenuate comprehension.

Graf Estes et al. (2007) found that 17-month-olds were more likely to learn object-label associations from words that they had previously segmented from running speech compared to nonwords. Similarly, 2-year-olds use a combination of familiarity and prosody (e.g., word stress) to disambiguate function in sentences, such that their ability to accurately locate a picture of a familiar noun ("doggy") was unaffected by preceding *unstressed* adjectives, whether familiar ("good doggy") or unfamiliar ("glib doggy"). However, if the unfamiliar adjective was stressed ("GLIB doggy"), accurate localization of the named noun decreased (Thorpe and Fernald 2006), suggesting that prosody continues to play an important role in language perception. Swingley et al. (1999) showed that when 2-year-olds heard the word "dog," they more quickly looked at a picture of a dog when the alternate picture was of a tree, compared to a picture of a doll. In this situation, the picture of a doll (a known object) interfered with processing the object-label relation of "dog," presumably because the children required more lexical specificity before being able to make the correct choice.

As is evidenced in the aforementioned studies, much of the research in children older than 1 year of age has focused on lexical and grammatical development, with far less attention paid to continuing developments in speech perception *per se*, except with regard to atypical populations. Yet speech perception and discrimination abilities continue to evolve beyond the first year, as do children's strategies for integrating different sources of information (see, e.g., work by Nittrouer 1996 among others). Children's strategies for integrating different sources of information appear to shift as they place more weight on dynamically changing aspects of signals (and less weight on more static aspects) than do adult listeners (Robinson and Sloutsky 2004; Nittrouer and Lowenstein 2010). Many of these changes appear to be driven by differences in perceptual attention to components of the speech signal. Additional research with these older children, and particularly research examining the role of linguistic experience on perceptual abilities, is clearly warranted.



## 5 Perception of Speech in Challenging (Nonoptimal) Listening Conditions

The bulk of this chapter has summarized evidence on infants' abilities to attend to and extract information from the speech signal, and suggested that this ability is related to emerging knowledge of language structure. Moreover, growing interest in how early perceptual finesse predicts emerging language production importantly links these often separate research literatures. One final area of concern, however, is whether the kinds of experimental situations commonly employed in studies on infants' perception of speech adequately reflect the challenges faced by novice language learners. As discussed by Leibold (Chap. 5), infants and toddlers are often faced with multiple sound sources occurring simultaneously in modality-rich contexts, requiring then that speech be segregated from the background to make sense of the auditory signal. This final section of the chapter addresses different aspects of infants' speech perception under challenge: situations in which infants must focus and maintain attention to one of many sources of language-relevant inputs.

Several studies have reported that infants experience multitalker environments quite frequently (van de Weijer 1998), and thus their listening behavior in such settings may be a more realistic indication of their typical language exposure than is their listening behavior in a quiet laboratory setting. In fact, Trehub et al. (1981) found that infants' thresholds for detection of speech in noise were approximately 12 dB higher than those of adults, and Nozza et al. (1990, 1991) found that infants could distinguish phonemes in the presence of band-passed noise, but were more negatively affected than were adult listeners. These two studies, along with similar results for detecting tones in noise, suggest that infants' speech perception in noise is substantially poorer than that found in adults. Moreover, infants are not able to utilize top-down knowledge to "fill in" information masked by noise. When the source of noise is loud but brief (e.g., a car honking outside), such that it entirely masks a portion of the incoming signal, adult listeners "restore" the missing information, based on prior knowledge of the language (Warren 1970). Infants and toddlers do not show this same pattern (Newman 2006), and may be negatively affected by transient noises to a greater extent than adults.

More recent research has explored infants' speech perception in the presence of other types of noise, particularly natural sounds such as other people talking. In either word recognition tasks (e.g., Newman 2005), or segmentation tasks (Newman and Jusczyk 1996), the general conclusion is that infants' performance in the presence of background sounds is more compromised than that seen in adult listeners. Infants require substantially higher signal-to-noise ratios than adults in order to identify known words, especially during the first year of life (Newman 2005). Other studies suggest that infants' performance is also qualitatively different from adults' performance. For example, infants show better recognition when the distracter stream consists of multi-talker babble than of a single voice speaking, a pattern opposite that of adults (Newman 2009).

Another qualitative difference in perception between infants and adults is seen in the presence of auditory distracters that do not share acoustic space with speech (i.e., no acoustic masking). In one study, 6-month-old English-learning infants

showed no group-level discrimination of simple, native phonemes (e.g., *boo* vs. *goo*) in the presence of a high-frequency, natural distracter (Polka et al. 2008). However, the task itself had excluded factors (discussed previously) that most likely enhance infants' processing skills under challenge: the discrimination task involved a male, monotone voice, and a black-and-white checkerboard display. To extend this work, additional experiments by the first author (Panneton; in collaboration with Polka) have recently tested 6-month-olds' discrimination of the same easy phonemes, in the presence of the same distracters, but with various combinations of dynamic face and voice presentations. Preliminary results show that infants' discrimination in noise is improved when IDS is the style of vocal presentation. Similar studies have found that infants are better able to recognize familiar words in the presence of distracters when they can either see the face of the speaker (Hollich et al. 2005) or when the voice speaking is familiar to them (Barker and Newman 2004).

Collectively, these results reinforce the primary framework articulated earlier (Sect. 3) that the development of infants' perception of language is embedded in a typically rich, multimodal context, involving caretakers and others who adjust their style of communication to maximize infants' attention regulation, even if these adjustments flow naturally from emotional intentions. Presumably, infants who are better able to perceive speech in noise may be expected to have better language skills later on (another way in which perception in infancy can be related to emerging productive skills at later ages). According to the current view, individual differences in this domain most likely emerge in a variety of ways. One, the degree to which caretakers engage in sensitive, contingent, infant-directed styles of communication may enhance infants' resilience to the negative effects of noise and distracters on language processing. Second, variance in home-related ecologies (e.g., the overall amount of noise and distraction) may be related to infants' emerging adeptness at stream segregation. That is, living with more perceptual challenge may actually engender better performance at attending to a signal in noise. Clearly, future research extending these ideas to various populations of infants and young children is needed.

Lastly, infants and young children also face the ongoing challenge of processing speech emanating from speakers with whom they are not familiar (i.e., different vocal registers from their caretakers), who come from different natural groupings with distinct vocal signatures (e.g., males vs. females; adults vs. children), and who may be presenting their native language in unfamiliar accents or dialects. Initially, young infants fail to generalize familiar speech across genders, although they will generalize across talkers within a gender (Houston and Jusczyk 2000). Likewise, infants fail to generalize to speakers with novel accents (Schmale and Seidl 2009) and toddlers have difficulty recognizing words spoken in novel accents (Best et al. 2009). Thus, variability across talkers in gender or accent can pose difficulty for young infants, although with experience, such variation may actually afford infants the ability to normalize their representations of lexical forms such that they learn to ignore surface variability (Rost and McMurray 2010). Infants raised in bi-dialectal families may have greater opportunities to learn which acoustic variation is irrelevant than do infants raised in a single-dialect home. On the other hand, if accommodating dialectal differences requires additional cognitive resources from the child, it may

continue to pose problems, particularly in situations that are already difficult for the child (e.g., listening in noise) or which are near the limit of what the child is capable of doing at that stage of development.

## 6 Summary

This chapter began by placing the development of speech perception in a motivational framework, wherein infants' attention drives their language experience. Properties of caretakers' interactive styles encourage infants' attention, creating opportunities for learning about the native language. Language learning emerges from this dyadic context, with infants developing important skills from which they are able to bootstrap more advanced processing abilities. As a result, infants' ability to discriminate different languages, to segment the fluent speech stream, and to track statistical patterns in the input unfold, as attunement to native language properties increases. With greater experience, infants bring more advanced processing skills to the task of perceiving spoken language around them, progressively distributing their attention to essential aspects of the communicative environment.

Future work will contribute substantially to our understanding of language development if guided by awareness of important issues. First, extant research on children's speech perception has focused primarily on early skills, such as perceptual attunement and segmentation, and phonemic awareness in preschool children just before learning to read. New studies need to bridge infants' perceptual skills with children's developing production and comprehension skills. Second, future developmental studies need to address individual differences, rather than only group-level performance, in all domains of speech perception. The literature is dominated by studies on English or European language acquisition in monolingual families. More work is needed across disparate languages, and multilingual learning environments. Research is also needed on how speech perception skills differ in children at risk for clinical conditions. Moreover, not all learning takes place in the types of settings that mimic those in laboratories. Outside of the laboratory, infants face challenges such as the presence of noise, distracters, and signal variability. The degree to which infants accommodate perceptual challenge is not well understood, and may be an important source of individual differences in task performance. As perceptual challenge taxes cognitive resources, infants may fail to recognize or discriminate speech signals. As a result, infants' speech perception is situationally dependent as well as jointly influenced by individual processing abilities, motivation, and effective caretaking strategies (Werker and Curtin 2005).

This expanded focus opens up understanding of diverse pathways by which infants acquire their native language. Although nearly all children become successful language users, they do so via more than one trajectory. Examining individual differences will lead to more nuanced theories of how language acquisition builds on early perceptual skills and experiences. Addressing these questions will compel shifts in methodological approaches, including more graded responses

involving different functional systems, and multiple-point assessments. Developing methodologies that better enable collaborations, across laboratories as well as disciplines, will yield great benefits given inherent difficulties in securing infant samples with the size, diversity, and power to address new and important issues in this field.

## References

- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: University of Edinburgh Press.
- Aslin, R. N. (2007). What's in a look? *Developmental Science*, 10, 48–53.
- Aslin, R. N., & Fiser, J. (2005). Methodological challenges for understanding cognitive development in infants. *Trends in Cognitive Sciences*, 9, 92–98.
- Aslin, R. N., & Mehler, J. (2005). Near-infrared spectroscopy for functional studies of brain activity in human infants: Promise, prospects, and challenges. *Journal of Biomedical Optics*, 10, 11009.
- Barker, B. A., & Newman, R. S. (2004). Listen to your mother! The role of talker familiarity in infant streaming. *Cognition*, 94, B45–B53.
- Bernhardt, B. M., Kemp, N., & Werker, J. F. (2007). Early word-object associations and later language development. *First Language*, 27, 315–328.
- 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. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science*, 20, 539–542.
- Bortfeld, H., Morgan, J. L., Golinkoff, R., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science*, 16, 298–304.
- Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, 2, 29–49.
- Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, 46, 217–243.
- Boysson-Bardies, B. de, Sagart, L., & Durand, C. (1984). Discernable differences in the babbling of infants according to target language. *Journal of Child Language*, 11, 1–15.
- Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactics are useful for segmentation. *Cognition*, 61, 93–125.
- Burns, T. C., Yoshida, K. A., Hill, K., & Werker, J. F. (2007). Bilingual and monolingual infant phonetic development. *Applied Psycholinguistics*, 28, 455–474.
- Butler, J., Floccia, C., Goslin, J., & Panneton, R. (2010). Infants' discrimination of familiar and unfamiliar accents in speech. *Infancy*, 16, 392–417.
- Christiansen, M. H., Allen, J., & Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes*, 13, 221–268.
- Christophe, A., & Morton, J. (1998). Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science*, 1, 215–219.
- Christophe, A., Gout, A., Peperkamp, S., & Morgan, J. (2003). Discovering words in the continuous speech stream: The role of prosody. *Journal of Phonetics*, 31, 585–598.
- Cohen, L. B., Diehl, R. L., Oakes, L. M., & Loehlin, J. C. (1992). Infant perception of /aba/ versus /apa/: Building a quantitative model of infant categorical discrimination. *Developmental Psychology*, 28, 261–272.
- Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, 52, 337–367.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, 208, 1174–1176.

- DeCasper A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior & Development*, 9, 133–150.
- Dehaene-Lambertz, G., Montavont, A., Jobert, A., Alliol, L., Dubois, J., Hertz-Pannier, L., & Dehaene, S. (2010). Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain and Language*, 114, 53–65.
- Dietrich, D. E., Emrich, H. M., Waller, C., Wieringac, B. M., Johannes, S., & Münted, T. F. (2000). Emotion/cognition-coupling in word recognition memory of depressive patients: An event-related potential study. *Psychiatry Research*, 96, 15–29.
- Eimas, P. D., & Miller, J. L. (1980). Discrimination of information for manner of articulation. *Infant Behavior & Development*, 3, 367–375.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorrito, J. (1971). Speech perception in infants. *Science*, 171, 303–306.
- Fais, L., Kajikawa, S., Amano, S., & Werker, J. F. (2009). Infant discrimination of a morphologically relevant word-final contrast. *Infancy*, 14, 488–499.
- Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, 46, 245–264.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behaviour and Development*, 8, 181–195.
- Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, 27, 209–221.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mother's and father's speech to preverbal infants. *Journal of Child Language*, 16, 477–501.
- Fernald, A., Swingle, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize words using partial phonetic information. *Child Development*, 72, 1003–1015.
- Fernald, A., Perfors, A., & Marchman, V.A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, 42, 98–116.
- Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. Sekerina, E. M. Fernandez, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). Amsterdam: John Benjamins.
- Fisher, C., & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development*, 67, 3192–3218.
- Friederici, A. D., & Oberecker, R. (2008). The development of syntactic brain correlates during the first years of life. In A. D. Friederici & G. Thierry (Eds.), *Trends in language acquisition research* (pp. 215–231). Amsterdam: John Benjamins.
- Gerken, L. A. (1996). Phonological and distributional information in syntax acquisition. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 411–425). Mahwah, NJ: Lawrence Erlbaum.
- Gerken, L. A., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, 32, 249–268.
- Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, 19, 515–522.
- Golinkoff, R. M., & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: Implications for language acquisition. *Journal of Child Language*, 22, 703–726.
- Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? *Psychological Science*, 18, 254–260.
- Grossmann, T., Oberecker, R., Koch, S. P., & Friederici, A. D. (2010). Developmental origins of voice processing in the human brain. *Neuron*, 65, 852–858.
- Hanulíková, A., McQueen, J. M., & Mitterer, H. (2010). Possible words and fixed stress in the segmentation of Slovak speech. *The Quarterly Journal of Experimental Psychology*, 63, 555–579.

- Harrison, P. (2000). Acquiring the phonology of lexical tone in infancy. *Lingua*, 110, 581–616.
- Hayashi, A., Tamekawa, Y., & Kiritani, S. (2001). Developmental change in auditory preference for speech stimuli in Japanese infants. *Journal of Speech & Hearing Disorders*, 44, 1189–1200.
- Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. (2000). Breaking the language barrier: An emergentist coalition model of word learning. *Monographs of the Society for Research in Child Development*, 65, 3 (Serial No. 262).
- Hollich, G., Newman, R. S., & Jusczyk, P. W. (2005). Infants' use of synchronized visual information to separate streams of speech. *Child Development*, 76, 598–613.
- Hoonhorst, I., Colin, C., Markessis, E., Radeau, M., Deltenre, P., & Serniclaes, W. (2009). French native speakers in the making: From language-general to language-specific voicing boundaries. *Journal of Experimental Child Psychology*, 104, 353–366.
- Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1570–1582.
- Houston, D. M., Horn, D. L., Qi, R., Ting, J., & Gao, S. (2007). Assessing speech discrimination in individual infants. *Infancy*, 12, 119–145.
- Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, 13, 341–348.
- Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44, 548–568.
- Johnson, E. K., & Zamuner, T. S. (2010). Using infant and toddler testing methods in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 73–94). Amsterdam: John Benjamins.
- Junge, C., Hagoort, P., Kooijman, V., & Cutler, A. (2010). Brain potentials for word segmentation at seven months predict later language development. In K. Franich, K. M. Iserman & L. L. Keil (Eds.), *Proceedings of the 34th annual Boston University Conference on Language Development* (Vol. 1, pp. 209–220). Somerville, MA: Cascadilla Press.
- Jusczyk, P. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29, 1–23.
- Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, 33, 630–645.
- Katz, G. S., Cohn, J. F., & Moore, C. A. (1996). A combination of vocal f0 dynamic and summary features discriminates between three pragmatic categories of infant-directed speech. *Child Development*, 67, 205–217.
- Kitamura, C., & Burnham, D. (1998). The infant's response to vocal affect in maternal speech. In C. Rovee-Collier (Ed.), *Advances in Infancy Research*, 12, 221–236.
- Kitamura, C., & Burnham, D. (2003). Pitch and communicative intent in mother's speech: Adjustments for age and sex in the first year. *Infancy*, 4, 85–110.
- Kitamura, C., & Lam, C. (2009). Age-specific preferences for infant-directed affective intent. *Infancy*, 14, 1–24.
- Klee, T., Pearce, K., & Carson, D. (2000). Improving the positive predictive value of screening for developmental language disorder. *Journal of Speech, Language, and Hearing Research*, 43, 821–833.
- Kooijman, V., Hagoort, P., & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research*, 24, 109–116.
- Kuhl, P. (2007). Is speech learning 'gated' by the social brain? *Developmental Science*, 10, 110–120.
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, L. V., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277, 684–686.
- 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 of the USA*, 100, 9096–9101.

- Lasky, R. E., Syrdal-Lasky, A., & Klein, R. E. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology*, 20, 215–225.
- Liu, H. M., Kuhl, P. K., & Tsao, F. M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science* 6 F1–F10.
- Lloyd-Fox, S., Blasi, A., & Elwell, C. E. (2010). Illuminating the developing brain: The past, present and future of functional near-infrared spectroscopy. *Neuroscience Biobehavioral Review*, 34, 269–284.
- Locke, J. (2001). First communion: The emergence of vocal relationships. *Social Development*, 10, 294–308.
- Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, 11, F9–F16.
- Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, 10, 241–265.
- Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, 106, 1367–1381.
- Mattys, S. L., & Jusczyk, P. W. (2001). Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology: Human Perception & Performance*, 27, 644–655.
- Mattys, S., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General*, 134, 477–500.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, B101–B111.
- McCardle, P., Colombo, J., & Freund, L. (2009). Measuring language in infancy. In Colombo, J., McCardle, P., & Freund, L. (Eds.), *Infant pathways to language: Methods, models, and research disorders* (pp. 1–12). New York: Psychology Press.
- McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy*, 6, 203–229.
- McMurray, B., & Aslin, R. N. (2005). Infants are sensitive to within-category variation in speech perception. *Cognition*, 95, B15–B26.
- Mehler, J., Gervain, J., Endress, A. D., & Shukla, M. (2008). Mechanisms of language acquisition: Imaging and behavioral evidence. In C. A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (2nd ed., pp. 325–336). Cambridge, MA: MIT Press.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143–178.
- Moon, C., Panneton-Cooper, R., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior & Development*, 16, 495–500.
- Moore, D., Spence, M., & Katz, G. (1997). Six-month-olds' categorization of natural infant directed utterances. *Developmental Psychology*, 33, 980–989.
- Morgan, J. L., & Saffran, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development*, 66, 911–936.
- Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual prosody and speech intelligibility. *Psychological Science*, 15, 133–137.
- Naigles, L. R. (2002). Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition*, 86, 157–199.
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception & Performance*, 24, 756–766.
- Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, 43, 1–19.
- Nazzi, T., Paterson, S., & Karmiloff-Smith, A. (2003). Early word segmentation by infants and toddlers with Williams syndrome. *Infancy*, 4, 251–271.
- Nazzi, T., Dilley, L. C., Jusczyk, A. M., Shattuck-Hufnagel, S., & Jusczyk, P. W. (2005). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech*, 48, 279–298.

- Newman, R. S. (2005). The cocktail party effect in infants revisited: Listening to one's name in noise. *Developmental Psychology*, 41, 352–362.
- Newman, R. S. (2006). Perceptual restoration in toddlers. *Perception & Psychophysics*, 68, 625–642.
- Newman, R. S. (2009). Infant's listening in multitalker environments: Effect of the number of background talkers. *Attention, Perception, & Psychophysics*, 71, 822–836.
- Newman, R. S., & Hussain, I. (2006). Changes in infant preference for infant-directed speech in low and moderate noise by 4.5- to 13-month-olds. *Infancy*, 10, 61–76.
- Newman, R. S., & Jusczyk, P. W. (1996). The cocktail party effect in infants. *Perception & Psychophysics*, 58, 1145–1156.
- Newman, R. S., Bernstein Ratner, N., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology*, 42, 643–655.
- Nittrouer, S. (1996). Discriminability and perceptual weighting of some acoustic cues to speech perception by three-year-olds. *Journal of Speech & Hearing Research*, 39, 278–297.
- Nittrouer, S., & Lowenstein, J. H. (2010). Learning to perceptually organize speech signals in native fashion. *Journal of the Acoustical Society of America*, 127, 1624–1635.
- Nozza, R. J., Rossman, R. N. F., Bond, L. C., & Miller, S. L. (1990). Infant speech-sound discrimination in noise. *Journal of the Acoustical Society of America*, 87, 339–350.
- Nozza, R. J., Rossman, R. N. F., & Bond, L. C. (1991). Infant-adult differences in unmasked thresholds for the discrimination of consonant-vowel syllable pairs. *Audiology*, 30, 102–112.
- Nygaard, L. C., & Lunders, E. R. (2002). Resolution of lexical ambiguity by emotional tone of voice. *Memory & Cognition*, 30, 583–593.
- Panneton, R., Kitamura, C., Mattock, K., & Burnham, D. (2006). Slow speech enhances younger, but not older infants' perception of vocal emotion. *Research in Human Development, Special Issue: The ecology of emotion in parenting relationships: Perspectives across the lifespan*, 3, 7–19.
- Panneton-Cooper, R. (1993). The effect of prosody on young infants' speech perception. In C. Rovee-Collier & L.P. Lipsitt (Eds.), *Advances in Infancy Research* (pp. 137–167). Norwood, NJ: Ablex.
- Panneton-Cooper, R., & Aslin, R.N. (1990). Preference for infant-directed speech in the first months after birth. *Child Development*, 61, 1584–1595.
- Panneton-Cooper, R., & Aslin, R. N. (1994). Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Development*, 65, 1663–1677.
- Panneton-Cooper, R., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for motherese. *Infant Behavior and Development*, 20, 477–488.
- Pegg, J., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development*, 15, 325–345.
- Peña, M., Makí, A., Kovačić D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences of the USA*, 100, 11702–11705.
- Pike, K. (1945). *The intonation of American English*. Ann Arbor, MI: University of Michigan Press.
- Polka, L., & Bohn, O-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, doi:10.1016/j.wocn.2010.08.007.
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 421–435.
- Polka, L., Rvachew, S., & Molnar, M. (2008). Speech perception by 6- to 8-month-olds in the presence of distracting sounds. *Infancy*, 13, 421–439.
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, 28, 349–351.
- Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development*, 75, 1387–1401.
- Rost, G. C., & McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy*, 15, 680–635.



- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.
- Saffran, J. R., Werker, J. F., & Werner, L. A. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In W. Damon & R. M. Lerner (Eds.), *Handbook of Child Psychology* (pp. 58–108). Hoboken, NJ: John Wiley & Sons.
- Santesso, D. L., Schmidt, L. A., & Trainor, L. J. (2007). Frontal brain electrical activity (EEG) and heart rate in response to affective infant-directed (ID) speech in 9-month-old infants. *Brain and Cognition*, 65, 14–21.
- Sato, Y., Sogabe, Y., & Mazuka, R. (2010). Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology*, 46, 106–119.
- Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice*, 9, 235–248.
- Schmale, R., & Seidl, A. (2009). Accommodating variability in voice and foreign accent: Flexibility of early word representations. *Developmental Science*, 12, 583–601.
- Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science*, 16, 197–201.
- Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shifts in the discrimination of vowel contrasts in bilingual infants: Is the distributional account all there is to it? *Developmental Science*, 12, 874–887.
- Seidl, A., & Johnson, E. K. (2006). Infant word segmentation revisited: Edge alignment facilitates target extraction. *Developmental Science*, 9, 565–573.
- Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy*, 3, 365–394.
- Smith, N. A., & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy*, 13, 410–420.
- Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory & Language*, 49, 249–267.
- Soderstrom, M., Conwell, E., Feldman, N., & Morgan, J. L. (2009). Statistical learning in language acquisition: Beyond demonstrations, towards a theory. *Developmental Science*, 12, 409–411.
- Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature*, 388, 381–382.
- Stern, D. (1985). *The interpersonal world of the infant*. New York: Basic Books.
- Sundara, M., & Scutellaro, A. (2010). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics*, doi:10.1016/j.wocn.2010.08.006.
- Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition*, 108, 232–242.
- Swingle, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, 8, 432–443.
- Swingle, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, 71, 73–108.
- Teinonen, T., Aslin, R. N., Alku, P., & Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition*, 108, 850–855.
- Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, 39, 706–716.
- Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, 7, 53–71.
- Thorpe, K., & Fernald, A. (2006). Knowing what a novel word is not: Two-year-olds 'listen through' ambiguous adjectives in fluent speech. *Cognition*, 100, 389–433.
- Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin and Review*, 9, 335–340.
- Trainor, L. J., Austin, C. M., & Desjardines, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science*, 11, 188–195.
- Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development*, 97, 466–472.

- Trehub, S. E., Bull, D., & Schneider, B. A. (1981). Infants' detection of speech in noise. *Journal of Speech and Hearing Research*, 24, 202–206.
- Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *Journal of the Acoustical Society of America*, 120, 2285–2294.
- Vallabha, G. K., McClelland, J. L., Pons, F., Werker, J. F., & Amano, S. (2007). Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of the National Academy of Sciences of the USA*, 104, 13273–13278.
- van de Weijer, J. (1998). *Language input for word discovery*. Wageningen, the Netherlands: Ponsen & Looijen, BV.
- Vihman, M., Nakai, S., DePaolis, R. A., & Halle, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory & Language*, 50, 336–353.
- Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, 10, 159–164.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, 167, 392–393.
- Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastian-Galles, N., & Werker, J. F. (2007). Visual language discrimination in infancy. *Science*, 316, 1159.
- Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning & Development*, 1, 197–234.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, 7, 49–63.
- Werker, J. F., Polka, L., & Pegg, J. E. (1997). The conditioned head turn procedure as a method for assessing infant speech perception. *Early Development and Parenting*, 6, 171–178.
- Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, 34, 1289–1309.
- Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, 103, 147–162.
- White, L., Melborn, J. F., & Mattys, S. (2010). Segmentation by lexical subtraction in Hungarian speakers of second-language English. *Quarterly Journal of Experimental Psychology*, 63, 544–554.
- Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.
- Zangl, R., Klarman, L., Thal, D. J., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development*, 6, 179–208.