# REPORT

# Tuned to the signal: the privileged status of speech for young infants

## Athena Vouloumanos and Janet F. Werker

*Department of Psychology, University of British Columbia, Canada*

## Abstract

*Do young infants treat speech as a special signal, compared with structurally similar non-speech sounds? We presented 2- to 7-month-old infants with nonsense speech sounds and complex non-speech analogues. The non-speech analogues retain many of the spectral and temporal properties of the speech signal, including the pitch contour information which is known to be salient to young listeners, and thus provide a stringent test for a potential listening bias for speech. Our results show that infants as young as 2 months of age listened longer to speech sounds. This listening selectivity indicates that early-functioning biases direct infants' attention to speech, granting speech a special status in relation to other sounds.*

## Introduction

Is speech a privileged signal for infants? At birth, the newborn's perceptual system is already tuned to some of the dimensions of human speech that are exploited by the phonological and syntactic systems of language (Bertoncini, Bijeljac-Babic, Blumstein & Mehler, 1987; Christophe, Dupoux, Bertoncini & Mehler, 1994; Jusczyk, Bertoncini, Bijeljac-Babic, Kennedy & Mehler, 1990; Mehler, Jusczyk, Lambertz, Halsted, Bertoncini & Amiel-Tison, 1988; Nazzi, Bertoncini & Mehler, 1998; Ramus, Hauser, Miller, Morris & Mehler, 2000; Sansavini, Bertoncini & Giovanelli, 1997; Shi, Werker & Morgan, 1999). These initial sensitivities become increasingly tuned to the properties of the native language during the infant's first year (Werker & Tees, 1992), a refinement that is reflected both in the decline of infants' discrimination of contrasts or properties that are not informative for processing their native language (Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992; Werker & Tees, 1984), and in the gain of sensitivities that are pertinent to native-language processing (Jusczyk, Cutler & Redanz, 1993; Jusczyk, Hohne & Bauman, 1999; Myers, Jusczyk, Kemler Nelson, Charles Luce, Woodward & Hirsh Pasek, 1996). Moreover, towards the end of their first year, infants become able to integrate multiple cues, and thus perform more sophisticated analyses on linguistic input (Morgan & Saffran, 1995). These observations suggest that infants' processing of the speech signal undergoes radical reshaping in the first year of life, becoming more specific and more sophisticated as infants approach their first birthday. Though the perception of speech compared with non-speech has been shown to engage specific neural substrates in the left hemisphere of adults (Binder & Price, 2001; Scott, Blank, Rosen & Wise, 2000; Vouloumanos, Kiehl, Werker & Liddle, 2001) and infants (Dehaene-Lambertz, Dehaene & Hertz-Pannier, 2002; Peña, Maki, Kovacic, Dehaene-Lambertz, Koizumi, Bouquet & Mehler, 2003), the mechanisms underlying the changes in speech perception evident during infancy are as yet unknown. The late Peter Jusczyk (1997) suggested that one factor that may facilitate infants' enhanced language processing might be a bias for listening to speech compared with other sounds. Such a bias might shape the character of learning processes and guide the action of perceptual mechanisms (Bolhuis & Honey, 1998; Johnson, 1999; Jusczyk, 1997; Marler, 1990).

In this study, we investigate whether young infants demonstrate a bias for listening to speech. Specifically, we ask if infants listen preferentially to speech compared with other non-linguistic sounds. To date, no study has

Address for correspondence: Athena Vouloumanos, Department of Psychology, McGill University, 1205 Dr Penfield Avenue, Montreal, QC, H3A 1B1, Canada; e-mail: athena@psych.mcgill.ca

addressed this question directly, though two previous studies provide relevant data.[1] Colombo and Bundy (1981) found that 4.5-month-old infants fixated longer on a visual target when it was associated with continuous female speech compared with continuous unfiltered white noise. However, 2-month-olds failed to show a similar fixation bias when continuous female speech was contrasted with silence (Colombo & Bundy, 1981). Glenn, Cunningham and Joyce (1981) found that 9-month-olds pulled a lever more frequently to listen to a female voice singing *a cappella* compared with three solo musical instruments playing the same tune. These results suggest that infants favour speech over some sounds. However, our understanding of a potential listening bias remains incomplete. The non-speech counterparts (white noise and musical instruments) used in these studies differ greatly from speech in frequency and timing characteristics, shedding little light on how discriminating a potential listening bias for speech might be (Jusczyk, 1997). Moreover, the listening biases of younger infants, who have significantly less experience with speech, are as yet unknown.

To test the specificity of infants' listening preference for speech and control for infants' sensitivity to superficial acoustic dimensions that are characteristic of the speech signal, we sought to contrast speech with closely matched complex non-speech sounds that preserved many aspects of the spectral and timing dimensions of speech without actually sounding like speech to naïve listeners. To this end, we created complex non-speech analogues that are modelled on sine-wave analogues of speech (Remez, Rubin, Pisoni & Carrell, 1981; Vouloumanos *et al.*, 2001). These complex analogues consist of time-varying sinusoidal waves that track the resonant centre frequencies of natural speech and reproduce the changes in these frequency peaks across time (see Figure 1). To ensure that the two different signals were equally attractive to infant ears, we created analogues that retain information about the pitch contour of the speech counterparts, since pitch contour has been shown to underlie infants' preference for infant-directed speech (Fernald & Kuhl, 1987) and their ability to discriminate their native language (Mehler *et al.*, 1988). The similarities between the acoustic properties of the complex non-speech analogues and natural speech allow us to investigate whether infants are attracted to signals having a particular acoustic form, and to better delineate the range of signals that engage a potential listening bias.

To track the emergence of a listening bias for speech, we focused on the first half-year of life. Previous studies had shown that infants listen selectively to speech compared with acoustically dissimilar unpatterned sounds by 4.5 months (Colombo & Bundy, 1981). Since the perceptual similarities between speech and our non-speech analogues may render the task more difficult, we began by testing older infants of 6.5 months. To determine whether speech has a special status during early infancy, we tested younger infants of 4.5 months and 2.5 months.

## Method

### Participants

Infants were recruited at birth from the British Columbia Women's and Children's Hospital, in Vancouver, Canada, or through advertisements placed in the community section of various local newspapers. Parents were subsequently contacted by phone to participate. All infants were full-term deliveries and heard at least 20% English in their home environment.

Forty-eight infants of three ages were included in this study: 16 6.5-month-olds (mean age: 6, 14), 16 4.5-month-olds (mean age: 4, 19), and 16 2.5-month-olds (mean age: 2, 16). An additional thirty-eight infants were tested but were excluded from the analysis.[2]

### Materials

Auditory stimuli were of two types: a 'speech' set composed of nonsense words, and a 'non-speech' set composed of complex non-speech analogues (Figure 1). A subset of these stimuli has been used in previous studies with adults (Vouloumanos *et al.*, 2001).

Speech

Speech stimuli consisted of 12 tokens of two monosyllabic nonsense words (six 'lif' tokens and six 'neem' tokens)[3]

---

[1] The methodological studies of Butterfield and Siperstein (e.g. 1970) are sometimes cited as providing relevant evidence. However, the 'speech' condition in their experiments consisted of folk music, and as such doesn't directly bear on infants' preference for speech.

[2] Twelve 6.5-month-olds were excluded because of excessive fussiness (2), technical errors (3), experimenter error (5), failure to learn the contingency (1) or parental interference in the experiment (1). Thirteen 4.5-month-olds were excluded because of excessive fussiness (3), technical errors (2), parental interference during the experiment (1), failure to learn the contingency (3), experimenter error (3) or hiccups (1). Thirteen 2.5-month-olds were excluded because of fussiness or sleepiness (9), technical errors (1), failure to learn the contingency (1) or experimenter error (2).

[3] For ease of readability, we describe the speech stimuli using 'gloss'. The equivalent in IPA symbols is /lIf/ and /nim/.

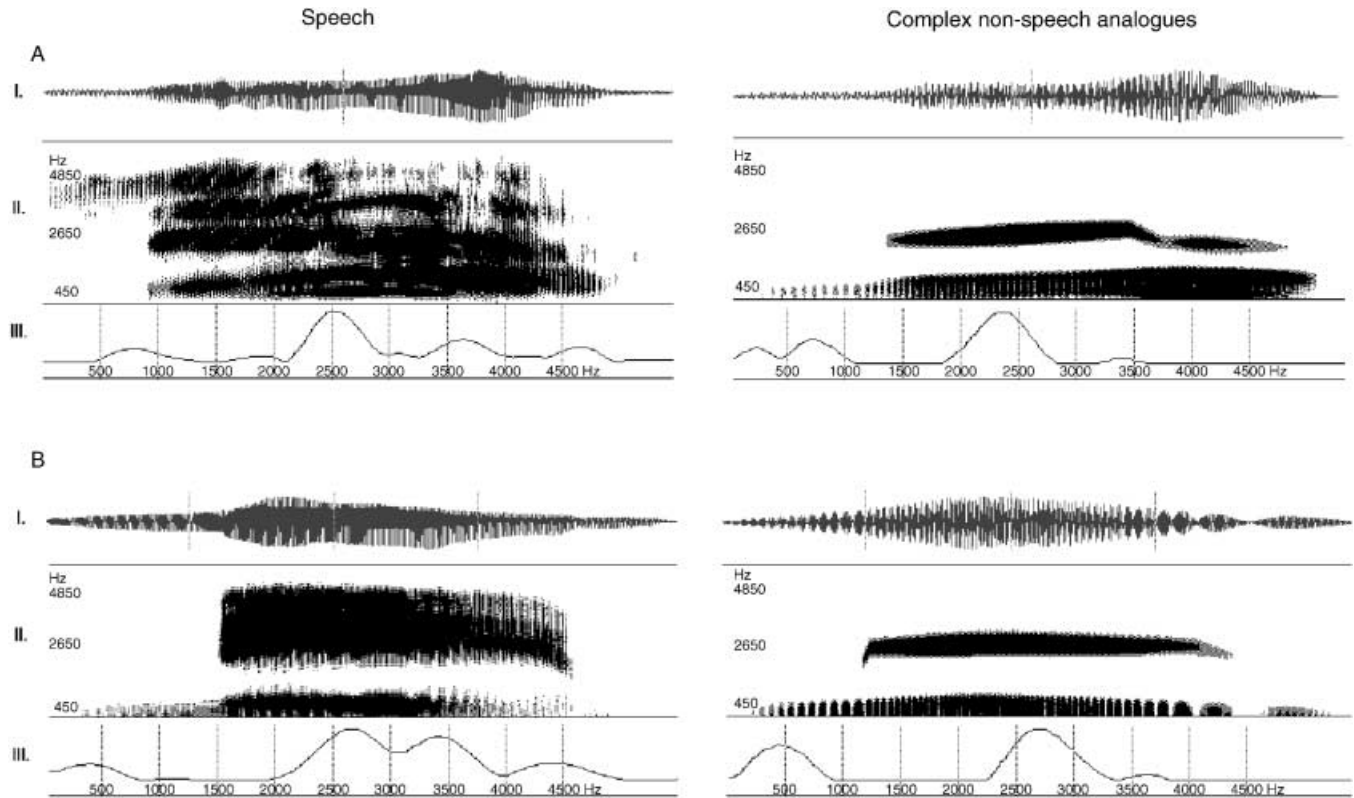Speech

Complex non-speech analogues



**Figure 1**  *Speech and complex non-speech stimuli for 'lif' (A), and 'neem' (B). Similarities between the two types of stimuli are illustrated in waveform diagrams (I), spectrograms showing changes in frequency across time (II) and spectra depicting the relative amplitudes of different frequencies (III).*

spoken by a female native English speaker. Tokens varied in intonational contour (average minimum and maximum pitch: 197 Hz and 350 Hz, respectively) and in duration (525–1155 ms).

Complex non-speech

Non-speech analogues were created by Sonya Bird and Guy Carden (Department of Linguistics, University of British Columbia, Canada). Non-speech stimuli consisted of time varying sinusoidal waves that tracked the main regions of significant energy in natural speech (namely the fundamental frequency and the first three formants). Sinusoidal waves tracking these energy peaks were created individually using Mathcad 3.1 (Mathsoft Inc., Cambridge, MA). Fundamental frequency (corresponding to pitch) was also tracked individually for each of the 12 speech tokens. Because the first three formants were virtually identical across the multiple natural repetitions of the two word types, one representative set of formants from a token of each word type was tracked. For the 'lif' tokens, this representative formant set was

composed of the first formant of the initial consonant segment ('l'), and the first three formants of the vocalic segment ('i'). The analogue for the fricative 'f' was created using a white noise generator and passed through a Butterworth filter (Pass Band cut-offs of 1700 and 4380 Hz, filter order 9) in Signalyze 3.12 (Agora Language Marketplace, Charlestown, MA). This representative set was then added onto a sinusoidal wave tracking the pitch contour of each of the six 'lif' segments using Signalyze 3.12 to create six different non-speech analogue 'lif' stimuli. For the 'neem' tokens, the representative formant set was composed of the first formant of the initial ('n') and final ('m') consonants, and the first three formants of the vocalic segment ('ee'). This representative set was then added onto a sinusoidal wave tracking the pitch contour of each of the six 'neem' segments using Signalyze 3.12 to create six different non-speech analogue 'neem' stimuli. Tokens were identical to speech foils in intonational contour (average minimum and maximum pitch: 197 Hz and 350 Hz, respectively) and in duration (525–1155 ms). Moreover, analogues retained the amplitude envelope, relative formant amplitude and

relative intensity of their speech counterparts. Crucially, the pitch contour of natural speech was preserved in the complex non-speech analogues.

### Design and procedure

Testing was conducted in a 7-ft × 9.5-ft sound-attenuated room. The walls were covered by matte black curtains, and the sole lighting source was a 60W floor lamp. Infants were seated on the lap of a parent or guardian, 4 feet away from a 27″ Mitsubishi CS-27205C television monitor (640 × 480 line vertical resolution), that protruded through a hole in the front curtain. Sounds were played at an average amplitude of 68dB (±3dB) using a BOSE 101 speaker placed directly above the television monitor. Infants were recorded with a Panasonic AG 180 video camera placed behind the front curtain with its lens positioned 10 inches below the television monitor. Parents were told that we were investigating 'how infants listen to different sounds', and wore Koss TD/65 or Peltor workstyle HT7A headphones playing music in order to mask the experimental sounds. Experimenters were blind to the specific condition being tested for any particular infant, and controlled the presentation of the stimuli from a separate room while monitoring the infants over a closed circuit using a Panasonic CT-13R12CT colour television. Stimulus presentation was controlled from a Power Mac 8500/1200 computer interfaced with a Sony LDP-1550 laser disc player using Habit 7.6 (Leslie Cohen, University of Texas at Austin).

Infants were tested using an infant-controlled sequential looking preference (SLP) procedure (Cooper & Aslin, 1990, 1994; Pegg, Werker & McLeod, 1992; Shi & Werker, 2001). In our version of this procedure, the program initially presents a red flashing light on the monitor to attract the infant's attention.[4] Once the infant fixates on the screen, testing begins. A stationary black and white checkerboard is displayed on the monitor at the same time as one set of experimental sounds is played from a hidden speaker placed on top of the monitor. In this SLP procedure, stimulus presentation is contingent on the infant's behaviour: infant fixation determines the onset and offset of every trial. The sound and checkerboard show continues for as long as the baby looks at the monitor. When the infant looks away continuously for longer than 1 second, stimulus presentation ceases. When the infant looks back at the monitor, the next trial begins: the checkerboard is displayed once again, but this time in tandem with the other set of experimental sounds. In any one experiment, the infant is presented with a total of 10 trials, five speech trials alternated with five non-speech trials. A full trial consists of 14 tokens chosen randomly from the set of 12 tokens, separated by 300 to 500 ms silence, for a maximum trial length of 20 seconds with the two older groups and 40 seconds with the youngest group.[5] For any given trial, speech or non-speech, tokens were ordered in a semi-random fashion so that every fixed window of four tokens included at least two 'lif's and two 'neem's. For half the infants, trial order was reversed.

## Results

As is standard with the SLP procedure, the first trial was excluded from the analysis (Cooper, Abraham, Berman & Staska, 1997; Cooper & Aslin, 1994; Shi & Werker, 2001). Since order of presentation was counterbalanced, an equal number of speech and non-speech trials was thus excluded. Using the remaining nine trials, we calculated each infant's total looking time for each type of sound, speech or non-speech. This was based on 15 frame-per-second, frame-by-frame coding of infant looks towards the screen during each sound trial. Because maximum trial length varied for different age groups and individual infants varied in their total looking time, we normalized the data by transforming total looking time to proportion looking time ($Prop_{speech} = M_{speech}/(M_{speech} + M_{nonspeech})$; for $Prop_{nonspeech} = 1 - Prop_{speech}$). A 3 (age: 6.5-, 4.5-, 2.5-month-olds) × 2 (sound type: speech vs. non-speech) × 2 (sex: female vs. male) × 2 (trial order) mixed analysis of variance indicated a main effect of sound type, $F_{1,36} = 15.048$, $p < .0005$, with proportionally longer looking times during speech trials ($M = .558$, $SE = .015$) than during non-speech trials ($M = .442$, $SE = .015$). No other main effects and no interactions were significant. We undertook a series of planned comparisons to determine if this effect was present for each age group. Individual two-tailed paired-sample $t$-tests on the average looking times revealed that a significant difference was present at each age: 2.5-month-olds: $t(15) = 2.143$, $p = .049$; 4.5-month-olds: $t(15) = 2.174$, $p = .046$; and 6.5-month-olds: $t(15) = 2.556$, $p = .022$. The average looking times of infants in different age groups are illustrated in Figure 2. A binomial test revealed that significantly more infants showed longer looking times for speech ($n = 34$) than for non-speech ($n = 14$), $p = .006$.

---

[4] In testing 2.5-month-olds, the red flashing light was also used between experimental trials to elicit infant attention.

[5] Pilot testing revealed that 20-second trials were too short for 2.5-month-olds to reliably learn the contingency, therefore 2.5-month-olds were tested using a maximum trial length of 40 seconds. Forty-second trials were created by simply repeating the 14 tokens presented during the first 20 seconds of that trial.
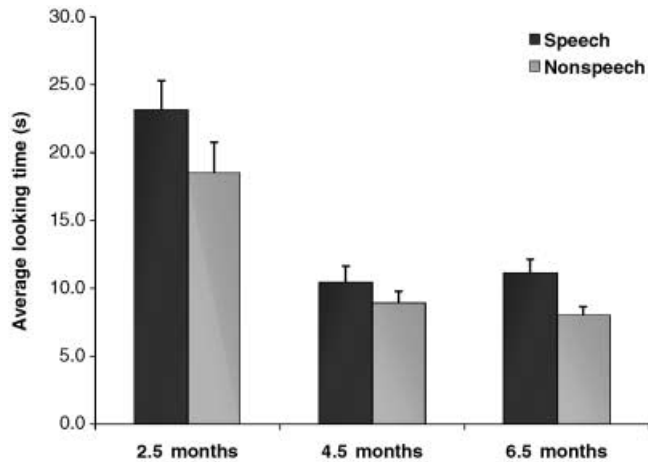
**Figure 2** *Average looking times for each age group: 2.5 months, 4.5 months and 6.5 months. Comparisons are significant at each age (see text).*

## Discussion

Infants between 2 and 7 months of age listened longer to speech compared with structurally similar complex non-speech sounds. Earlier results showed that infants listen preferentially to speech in comparison to white noise and musical instruments (Colombo & Bundy, 1981; Glenn *et al.*, 1981). Our results demonstrate that infants prefer speech even when it is contrasted with acoustically similar non-speech sounds. Moreover, we demonstrate that a bias for listening to speech is shown by infants as young as 2.5 months of age, 2 months earlier than the youngest age at which a bias for speech had previously been demonstrated (Colombo & Bundy, 1981).

Spectrally and temporally matched complex non-speech analogues fail to capture infant interest as effectively as speech, despite the structural similarities between the two types of sounds and despite the inclusion of the pitch contour in the non-speech analogues. The range of signals engaging a speech listening bias thus appears to be quite narrow. The non-speech stimuli used in this study preserve the time-varying frequency intervals that are characteristic of speech, as well as the relative formant amplitudes, the amplitude envelope and, importantly, the pitch contour of speech through the inclusion of the sinusoidal wave tracking the fundamental frequency. The fundamental frequency was included to preserve the pitch component known to be attractive to infants (Fernald & Kuhl, 1987), to add variability to the analogues, and to make the non-speech sounds more acoustically similar to their speech counterparts. The complex analogues did not, however, retain the characteristics of the voicing source, the broader band formant

information and parts of the harmonic spectrum, nor did they preserve the voice quality, biological quality or unified source characteristics of speech (Remez *et al.*, 1981). Thus, analogues lacked some of the qualities unique to speech as a biological sound produced by a human vocal tract. It remains to be seen which of these dimensions engages the speech bias that we observe.

In previous studies, the presentation of speech in the form of sentences has been shown to be important for investigating certain aspects of infant speech perception. Sentences, but not isolated words, elicited a preference for the mother's voice (Mehler, Bertoncini & Barriere, 1978). Similarly, infants' early discrimination of their native language is based on rhythmical information present in sentence form (Mehler *et al.*, 1988), and only later is evident for word-level segmental information (Jusczyk *et al.*, 1993). In order to create a stringent test for investigating infants' preference for speech, we presented speech as isolated words rather than sentences. That infants show a bias for speech when speech is presented as isolated words, indicates that this level of information is sufficient for eliciting a preference.

The young age at which infants listen preferentially to speech suggests that this bias is present very early in development. Such a listening bias could either be a *reflection* of the more extensive processing that speech undergoes, or the *cause* of subsequently more sophisticated processing, or both. Since the youngest infants tested in this study were 2.5 months old, it is not yet clear whether infants' bias for speech derives from exposure to language or whether it precedes it. We are currently investigating whether a listening bias for speech is present at birth (Vouloumanos & Werker, under review).

Regardless of the origin of a listening bias for speech, such a bias could benefit young language learners by allowing infants to separate and select speech out of their auditory environment in order to analyse the signal more completely (e.g. Jusczyk & Bertoncini, 1988; Jusczyk, 1997). Similar biases have been proposed in other domains, and may be pervasive in early development. For example, Johnson, Umiltà and their colleagues have proposed that an initial bias for orienting towards face-like stimuli plays a role in the development of face recognition (Johnson, Dziurawiec, Ellis & Morton, 1991; Valenza, Simion, Cassia & Umiltà, 1996).

The preference infants show for speech is reminiscent of the differential processing that some birds and primates demonstrate for calls of their own species (e.g. Gottlieb, 1997; Hauser, 1996; Marler, 1990). In some non-human primates, conspecific (same-species) vocalizations are preferentially processed by neurons along the superior temporal sulcus (Wang & Kadia, 2001; Wang, Merzenich, Beitel & Schreiner, 1995). This area may also be preferentially

recruited in humans for processing human vocalizations (Belin, Zatorre, Lafaille, Ahad & Pike, 2000; Binder & Price, 2001; Scott *et al.*, 2000). Studies using behavioural, electrophysiological and neuroimaging methods with newborns (Bertoncini *et al.*, 1989; Ecklund-Flores & Turkewitz, 1996; Peña *et al.*, 2003) and older infants (Dehaene-Lambertz *et al.*, 2002) suggest that young infants show a left hemisphere bias for processing speech. It is not yet known which parts of the human auditory system are tuned to conspecific sounds early in development, or whether neurons that respond specifically to conspecific *communication* sounds are present in the human cortex as they are in other species. The existence of conspecific biases implemented by specific neural substrates across different species suggests that these preferences may serve an adaptive function by helping organisms orient to conspecific information.

The results reported in this study provide the best evidence to date that very young infants listen selectively to speech compared with other sounds, in this case, sounds that mimic many of the physical characteristics of speech. This listening selectivity indicates that early-functioning biases direct infants' attention to speech, granting speech a special status in relation to other sounds. Such a bias might provide the basis for more sophisticated analysis of the speech signal which helps propel the infant into rapid language acquisition.

## Acknowledgements

## References

Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, **403** (6767), 309–312.

Bertoncini, J., Bijeljac-Babic, R., Blumstein, S.E., & Mehler, J. (1987). Discrimination in neonates of very short CVs. *Journal of the Acoustical Society of America*, **82** (1), 31–37.

Bertoncini, J., Morais, J., Bijeljac-Babic, R., McAdams, S., Peretz, I., & Mehler, J. (1989). Dichotic perception and laterality in neonates. *Brain and Language*, **37** (4), 591–605.

Binder, J.R., & Price, C. (2001). Functional neuroimaging of language. In R. Cabeza & A. Kingstone (Eds.), *Handbook of functional neuroimaging of cognition* (pp. 187–251). Cambridge, MA: MIT Press.

Bolhuis, J.J., & Honey, R.C. (1998). Imprinting, learning and development: from behaviour to brain and back. *Trends in Neurosciences*, **21** (7), 306–311.

Butterfield, E.C., & Siperstein, G.N. (1970). Influence of contingent auditory stimulation upon non-nutritional suckle. In J.F. Bosma (Ed.), *Third symposium on oral sensation and perception: The mouth of the infant* (pp. 313–334). Springfield, IL: Charles C. Thomas.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America*, **95** (3), 1570–1580.

Colombo, J.A., & Bundy, R.S. (1981). A method for the measurement of infant auditory selectivity. *Infant Behavior and Development*, **4** (2), 219–223.

Cooper, R.P., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for motherese. *Infant Behavior and Development*, **20** (4), 477–488.

Cooper, R.P., & Aslin, R.N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, **61**, 1584–1595.

Cooper, R.P., & Aslin, R.N. (1994). Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Development*, **65** (6), 1663–1677.

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, **298** (5600), 2013–2015.

Ecklund-Flores, L., & Turkewitz, G. (1996). Asymmetric headturning to speech and nonspeech in human newborns. *Developmental Psychobiology*, **29** (3), 205–217.

Fernald, A., & Kuhl, P.K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, **10** (3), 279–293.

Glenn, S.M., Cunningham, C.C., & Joyce, P.F. (1981). A study of auditory preferences in nonhandicapped infants and infants with Down's syndrome. *Child Development*, **52** (4), 1303–1307.

Gottlieb, G. (1997). *Synthesizing nature-nurture: Prenatal roots of instinctive behavior*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hauser, M.D. (1996). *The evolution of communication*. Cambridge, MA: MIT Press.

Johnson, M.H. (1999). Ontogenetic constraints on neural and behavioral plasticity: evidence from imprinting and face processing. *Canadian Journal of Experimental Psychology*, **53** (1), 77–91.

Johnson, M.H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, **40** (1–2), 1–19.

Jusczyk, P.W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Jusczyk, P.W., & Bertoncini, J. (1988). Viewing the development of speech perception as an innately guided learning process. *Language and Speech*, **31** (Pt 3), 217–238.

Jusczyk, P.W., Bertoncini, J., Bijeljac-Babic, R., Kennedy, L.J., & Mehler, J. (1990). The role of attention in speech perception by young infants. *Cognitive Development*, **5** (3), 265–286.

Jusczyk, P.W., Cutler, A., & Redanz, N.J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, **64** (3), 675–687.

Jusczyk, P.W., Hohne, E.A., & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception and Psychophysics*, **61** (8), 1465–1476.

Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, **255** (5044), 606–608.

Marler, P. (1990). Innate learning preferences: signals for communication. *Developmental Psychobiology*, **23** (7), 557–568.

Mehler, J., Bertoncini, J., & Barriere, M. (1978). Infant recognition of mother's voice. *Perception*, **7** (5), 491–497.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, **29** (2), 143–178.

Morgan, J.L., & Saffran, J.R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development*, **66** (4), 911–936.

Myers, J., Jusczyk, P.W., Kemler Nelson, D.G., Charles Luce, J., Woodward, A.L., & Hirsh Pasek, K. (1996). Infants' sensitivity to word boundaries in fluent speech. *Journal of Child Language*, **23** (1), 1–30.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, **24** (3), 756–766.

Pegg, J.E., 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** (3), 325–345.

Peña, M., Maki, A., Kovacic, 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 United States of America*, **100** (20), 11702–11705.

Ramus, F., Hauser, M.D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, **288** (5464), 349–351.

Remez, R.E., Rubin, P.E., Pisoni, D.B., & Carrell, T.D. (1981). Speech perception without traditional speech cues. *Science*, **212** (4497), 947–949.

Sansavini, A., Bertoncini, J., & Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology*, **33** (1), 3–11.

Scott, S.K., Blank, C.C., Rosen, S., & Wise, R.J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, **123** (12), 2400–2406.

Shi, R., & Werker, J.F. (2001). Six-month-old infants' preference for lexical words. *Psychological Science*, **12** (1), 70–75.

Shi, R., Werker, J.F., & Morgan, J.L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, **72** (2), B11–21.

Valenza, E., Simion, F., Cassia, V.M., & Umiltà, C. (1996). Face preference at birth. *Journal of Experimental Psychology: Human Perception and Performance*, **22** (4), 892–903.

Vouloumanos, A., Kiehl, K.A., Werker, J.F., & Liddle, P.F. (2001). Detection of sounds in the auditory stream: event-related fMRI evidence for differential activation to speech and nonspeech. *Journal of Cognitive Neuroscience*, **13** (7), 994–1005.

Vouloumanos, A., & Werker, J.F. (under review). An experience-independent bias for listening to speech in human newborns.

Wang, X., & Kadia, S.C. (2001). Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *Journal of Neurophysiology*, **86** (5), 2616–2620.

Wang, X., Merzenich, M.M., Beitel, R., & Schreiner, C.E. (1995). Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *Journal of Neurophysiology*, **74** (6), 2685–2706.

Werker, J.F., & Tees, R.C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life, *Infant Behavior and Development*, **7** (1), 49–63.

Werker, J.F., & Tees, R.C. (1992). The organization and reorganization of human speech perception. *Annual Review of Neuroscience*, **15**, 377–402.