39 Prosody as an Entry Point into Language Structure in Early Language Acquisition

GURO STENSBY SJULS, MILA DIMITROVA VULCHANOVA, AND JUDIT GERVAIN

39.1 Introduction

To acquire language, infants need to extract information from speech and develop an understanding of the relationship between sounds and their meaning. When observing infants and young children in everyday life, this seems like a gradually developing and effortless task, but the underlying process is likely more complex and currently not completely understood. For the language learner, some form of tracking of the stream is arguably needed for language acquisition to take off, for word and grammar learning to proceed. The question arises as to what features of the phonetically rich, but lexically and grammatically opaque input, provide an entry point into, and facilitate, the acquisition of the structure of language in its full complexity, including words and grammar. Here we take a closer look at prosody and its role in early development of auditory neural oscillations, focusing on a model in which synchronization to the slow fluctuations associated with the prosodic phrase level scaffold grammar learning in infancy (Nallet and Gervain, 2021).

39.2 The Prenatal Prosodic-Shaping Model

According to the prenatal prosodic-shaping model (Nallet and Gervain, 2021), infants' prenatal experience with the speech signal lays the foundation for subsequent grammar learning after birth. Developing fetuses are exposed to speech as early as from week 24 to 28 of gestation (Eggermont and Moore, 2011). However, due to the intrauterine environment, sounds are low-pass filtered, essentially providing the fetus with the prosodic contour of the speech signal (Gerhardt and Abrams, 2000; Menn et al., 2023).

Prosodic cues contribute to the parsing of the speech stream in the form of intonational phrase boundaries (Thompson and Balkwill, 2006), dynamic pitch changes (Watson and Gibson, 2005), metrical information (Liu et al., 2015), and so on. In terms of its function, prosody may be used as a grammatical marker, for example, of focus or interrogatives, and can

https://doi.org/10.1017/9781009295888.046 Published online by Cambridge University Press

also provide meaning to an utterance above and beyond the lexical and grammatical content by nuancing the speaker's intent in communicating affect, emphasis, irony, and so on (Coutinho and Dibben, 2013; Scherer, 1986; Zentner et al., 2008). As such, the above-mentioned linguistic phenomena associated with prosodic cues provide an anchor to the underlying structure of language and, importantly to the topic of this chapter, may thus bootstrap the development of grammar (Gervain and Werker, 2013; Nazzi and Ramus, 2003; Soderstrom et al., 2003).

How such regularities are processed by the brain has been a much-debated topic (e.g., Ding et al., 2016; Giraud and Poeppel, 2012). Recent advances have established that adults' brain responses simultaneously track the different timescales of the speech signal. Bottom-up processing of units in the speech signal is supported by neural oscillations in the delta (0.5–3.5 Hz), theta (4–8 Hz), and low-gamma (>35 Hz) frequency bands (Ding et al., 2016; Giraud and Poeppel, 2012). These bands, respectively, underlie the processing of prosodic phrases, syllables, and (sub-)phonemic units of speech (Giraud and Poeppel, 2012), as their frequencies match those of the relevant linguistic units. For further details, we direct the reader to Chapters 3 and 5. However, it is still unclear how the neural tracking of the speech and, in particular, the oscillatory hierarchy develop during the first year of life.

The prenatal prosodic-shaping model proposes that this development starts already prior to birth. In utero, the fine-grained phonemic information (i.e., the gamma band) is mostly suppressed in the low-passed auditory signal that reaches the fetus, while syllabic and prosodic phrase information is preserved, as the spectral content fluctuates at slower frequencies (corresponding to theta and delta frequency bands) (Gerhardt and Abrams, 2000). Given this prenatal experience with the speech signal, the neural tracking of larger linguistic units may already start prenatally, while oscillations tracking (sub-)phonemic information may not be operational until after birth (Nallet and Gervain, 2021). Postnatally, infants are exposed to the full-band speech signal, at which point the neural tracking of fine-grained phonemic elements may start being shaped by experience with the (unfiltered) speech signal.

Due to their prenatal exposure to parts of the speech signal, fetuses are born with a certain familiarity with language. It has been attested that neonates prefer sounds to which they have been exposed in the womb (DeCasper and Fifer, 1980; Mehler et al., 1988; Moon, 2017; Moon et al., 1993), which suggests that they do have the ability to learn from the low-passed signal available to them, and that some shaping of the language system takes place already in utero. For example, newborns show a preference for their mother's voice compared to an unknown female voice (DeCasper and Fifer, 1980; Moon, 2017), and for their native language over unfamiliar languages (Mehler et al., 1988; Moon et al., 1993). Following these results, a growing body of research suggests that prenatal experience, with the prosodic features preserved in the low-passed speech signal, might lay the foundations for even more complex language acquisition.

Newborns can, for example, distinguish well-formed from ill-formed prosodic sequences based on duration, pitch, or intensity, but only if the varying element is contrastive in the language they heard before birth (Abboub et al., 2016). Specifically, French newborns can discriminate between short-long and long-short sequences (variation in duration), which mark contrastive distinctions, but not between loud-soft and soft-loud (variation in intensity) or high-low and low-high sequences (variation in pitch), which are not markers of contrastive distinctions in French prosody (Nespor et al., 2008; Nespor and Vogel, 2007). In addition, even

https://doi.org/10.1017/9781009295888.046 Published online by Cambridge University Press

though consonants are likely not perceivable by the fetus, some information about vowels might be available, because vowels, which are the main carriers of prosodic information, are high-energy events in the speech signal. Accordingly, Moon et al. (2013) observed opposite preferences between American and Swedish newborns for the vowel with which they had prenatal experience (the American /i:/ versus the Swedish /y/ vowel).

It thus becomes evident that infants are born with a certain familiarity with the prosodic features of their native language. Importantly, as prosody also carries lexical, morphosyntactic, and pragmatic information, it is highly relevant to language development overall. In older infants, prosody provides cues to, for example, word boundaries (Shukla et al., 2011) and word order (Gervain and Werker, 2013), and is thus an important bootstrapping mechanism for lexical and grammatical development. However, already newborns display the ability to utilize prosodic features to gain access to more complex aspects of their native language. For instance, they can discriminate between function words (marking morphosyntactic structure) and content words (carrying lexical meaning) (Shi et al., 1999), and they are sensitive to word order and its violations (Benavides-Varela and Gervain, 2017). In addition, newborns are sensitive to prosodic violations at the utterance level, in that they discriminate between well-formed and ill-formed prosodic contours (Martinez-Alvarez et al., 2023).

According to the prenatal prosodic-shaping model (Nallet and Gervain, 2021), prosody provides the earliest experience with language, and is one of the mechanisms that links innate predispositions and soon-to-be relevant input from the environment. In other words, as prosodic features of spoken language fluctuate at slower frequencies, they are likely to be preserved in utero, as suggested by newborns' familiarity with these features. This prenatal prosodic experience is hypothesized to shape the neural architecture, meaning that neural entrainment to prosody is already operational at birth (Menn et al., 2023; Ortiz Barajas et al., 2021). When newborns get exposed to the full-band speech signal after birth, which includes the fine-grained acoustic information at the phonemic level, oscillations in the delta and theta bands are already fine-tuned, at least to some extent, to the rhythm of the prenatally heard language. After months of exposure to the full speech signal, phoneme perception becomes attuned to the native language, and neural activity in its corresponding frequency band, gamma, is fine-tuned and hierarchically embedded in the delta- and theta-band oscillations. This model can thus offer a theoretical account in which prenatal experience with prosody is the foundation on which subsequent language development is built, in that, with postnatal exposure, oscillations in the gamma band are gradually embedded in the prenatally acquired delta- and theta-band oscillations. In other words, the developmental chronology of experience with speech, first filtered then full-band, explains the hierarchical organization of oscillations. However, see also Menn et al. (2023) for a related perspective in which electrophysical maturation and the emergence of gamma-band activity shapes the acquisition of phonological knowledge.

39.3 Evidence for the Model

Recent research on the development of neural tracking suggests that delta-band tracking is operational in the first year of life. Infants at six and nine months of age, presented with streams

https://doi.org/10.1017/9781009295888.046 Published online by Cambridge University Press

of rhythmic stimuli in the form of the syllable "ta" or a drumbeat at a presentation rate of 2 Hz, displayed local peaks in electroencephalography (EEG) power at the presentation rate for both stimulus types, compared to a silent baseline condition (Choisdealbha et al., 2022). The response was entrained to the stimuli, namely time-locked to the onset of the stimulus, as observed through a relative increase in phase consistency. The response in the 4 Hz (harmonic frequency of the presentation rate) and 7 Hz (not-harmonic frequency of the presentation rate) was also assessed and compared to the response at the presentation rate. An increase in power at the harmonic frequency was observed, regardless of stimulus type, but a time-locked response similar to the 2 Hz response was only observed to the speech stimulus at the 4 Hz harmonic frequency. The infants were tested longitudinally, but, importantly, no evidence of improved tracking as a function of age was observed (Choisdealbha et al., 2022).

Similarly, when longitudinally following four-, seven-, and 11-month-old infants' tracking of sung speech (nursery rhymes), a peak in power in the delta band (at ~2.2 Hz) and the theta band (at ~4.3 Hz) was observed (Attaheri et al., 2022). However, delta-band tracking stayed significantly higher compared to theta-band tracking at each age, and was specifically strong at four months, while theta-band tracking increased over the course of the first year of life. The alpha band (8–12 Hz) was used as a control condition, in which no reliable tracking was observed (Attaheri et al., 2022). Given the results of these studies, infants appear to faithfully track auditory speech-related stimuli in the theta and delta bands, a mechanism that is operational from at least four months of age (Attaheri et al., 2022; Choisdealbha et al., 2022).

Interestingly, in terms of the predictions made by the prenatal prosodic-shaping hypothesis, Attaheri et al. (2022) also observed delta- and theta-band-driven phase amplitude coupling with higher-frequency amplitudes. Namely, the phase of delta- and theta-band activity acted as a carrier of amplitude in higher-frequency bands, both beta and gamma frequencies, although greatest for gamma-band activity. This finding is consistent with the model as it suggests that the prenatally unavailable higher frequencies associated with phonemes will become embedded in the delta-band oscillations when the infant is exposed to these frequencies after birth. In other words, the prenatal experience with the phases of delta- and theta-band frequencies is likely to play an important role in the temporal organization of the amplitude of higher-frequency bands in the infant brain, that is, in their nesting within the slower bands.

Another branch of research on early neural tracking of speech focuses on infant-directed speech (IDS) compared to adult-directed speech (ADS) (Kalashnikova et al., 2018; Menn et al., 2022). IDS, a type of speech often used by adults when speaking to infants, is characterized by a higher pitch, more variability in intonation, and a slower tempo compared to ADS. This is reflected in amplified slow-frequency modulations as compared to ADS, and has been shown to be preferred by infants, offering potential benefits in language development (Song et al., 2010). Seven-month-old infants have been found to track both naturally produced IDS and ADS, as measured in the increases in power in the theta band (Kalashnikova et al., 2018). Furthermore, significant correlations were found between the pattern of neural activity and the envelope of the speech signal for IDS, but not for ADS, meaning that the envelope of the speech signal was more strongly reflected in the neural signal when infants were presented with IDS compared to ADS (Kalashnikova et al., 2018).

https://doi.org/10.1017/9781009295888.046 Published online by Cambridge University Press

Menn et al. (2022) extend on these results by estimating whether the effect for IDS is driven by the syllabic rate or the prosodic stress for IDS compared to ADS in nine-month-olds, namely, which frequency band causes the effect (theta versus delta band, respectively). The infants listened to their mothers describe items in either an IDS- or ADS-like way. In addition to significant speech–brain coherence at the syllabic and prosodic stress rate for both IDS and ADS, a significantly higher coherence was found for IDS at the prosodic stress rate, but not at the syllabic rate, a difference driven by a left-central cluster. As such, their results indicate that prosodic stress (greater coherence at delta-band frequencies), but not syllable rhythm, might be the facilitator of greater neural tracking of IDS (Menn et al., 2022). These results might arise as a consequence of the differences in attentional salience between IDS and ADS, as the prosodic differences may contribute to increased attention to the speech sounds.

The results reviewed above come from somewhat older infants (between four and 11 months of age). However, are these abilities already present at birth? When assessing tracking of syllables, that is, activity in the theta band, no differences were found between newborns and six-month-old infants presented with short sentences read in IDS: Both groups similarly track the phase and amplitude of the envelope of familiar (native language) and unfamiliar (different rhythmic class and same rhythmic class as the first language [L1]) languages (Ortiz Barajas et al., 2021). Interestingly, while phase tracking continues to be universal, amplitude tracking is only kept up for the unfamiliar language, especially the rhythmically different one. As such, phase and amplitude tracking might be differentially modulated over the course of the first year, which may be reflective of a perceptual narrowing following infants' experience with their native language (Kuhl, 2004; Ortiz Barajas et al., 2021). More recent results with newborns (Ortiz Barajas et al., 2023) suggest that newborns show enhanced power in the delta and theta bands in response to the language heard prenatally and, to some extent, to a rhythmically similar unfamiliar language, as compared to a rhythmically different unfamiliar language, whereas no such language differences were present in the gamma band, where no enhanced power was found for speech in any of the languages tested. These results also confirm the hypothesis that slower oscillations are fine-tuned already at birth (Nallet and Gervain, 2021; see also Menn et al., 2023).

39.4 General Discussion

To acquire their native language, infants need to develop sensitivity to the phonological properties and contrasts characteristic of that language. These skills are a necessary first step on the path to word learning and grammar acquisition. Several models have been proposed to account for the relatively fast and seemingly effortless accomplishment of this challenging task. One such model, the native language model (Kuhl, 2004; Kuhl et al., 1992), focuses on developmental changes in early infant auditory perception and aims to incorporate social and communicative factors in recent versions of the model (Kuhl et al., 2008). The prenatal prosodic-shaping model offers a novel perspective on perceptual narrowing from the point of view of recent advances in the neurobiology of language and its alignment with neural

https://doi.org/10.1017/9781009295888.046 Published online by Cambridge University Press

structures and mechanisms that support its development in human infants. Development of synchronization of neural oscillations to speech at different levels of granularity offers a possible format for such an account.

Seen together, the current literature indicates that infants from four to 11 months of age reliably track speech in both the delta (Attaheri et al., 2022; Choisdealbha et al., 2022; Menn et al., 2022) and theta (Attaheri et al., 2022; Kalashnikova et al., 2018; Menn et al., 2022) frequency bands. Syllabic tracking, as reflected in the theta band, shows a developmental increase between four and 11 months (Attaheri et al., 2022). However, in terms of the phase of the signal, tracking appears to remain relatively stable from birth until around six months of age for both familiar and unfamiliar languages (Ortiz Barajas et al., 2021). Amplitude tracking, on the other hand, is universal at birth, but only manifests for unfamiliar languages at six months of age (Ortiz Barajas et al., 2021). Between four and 11 months, prosodic tracking in the delta band has been found to remain greater than syllabic tracking throughout these ages. In addition, IDS may facilitate neural tracking (Kalashnikova et al., 2018; Menn et al., 2022), an effect primarily driven by prosodic stress as reflected in stronger delta-band coherence for IDS compared to ADS (Menn et al., 2022).

In terms of the prenatal prosodic-shaping model, the low-frequency parts of the speech signal that are available prenatally are successfully tracked by the infant brain during the first year of life, although some developmental changes can also be observed. Based on extant research, both syllabic tracking and tracking of larger prosodic units are present from birth. Interestingly, in terms of nested oscillations, the phases of both delta- and theta-band oscillations do act as carriers of amplitude in higher-frequency bands, especially in the gamma band (Attaheri et al., 2022). While the currently available evidence supports the prenatal prosodic-shaping model, the number of available studies is still very small. In particular, few studies have tested newborn infants and fetuses. Studying newborns and fetuses poses a challenge compared to adults. Specifically with EEG, newborns' data tend to have a lower signal-to-noise ratio, which limits the analysis methods that can be used. During the first months of life, the electrophysiological activity is less structured than in later development and in adulthood, and evidence can as such be of a more indirect nature. Applying the same oscillatory models as with adults and older infants can therefore be somewhat challenging.

Taken together, and considering future empirical word in the field, the prenatal prosodic-shaping model has the potential to explain how available brain mechanisms interface with the infant environment, both at prenatal and neonatal/postnatal stages. As such, this model offers predictable hypotheses concerning the hierarchical nesting of neural oscillations in concert with increased complexity of the acquired language skills, from word learning to advanced grammar.

39.5 Acknowledgements

ERC Consolidator Grant "BabyRhythm" nr. 773202 to Judit Gervain and a FARE grant nr. R204MPRHKE from the Italian Ministry for Universities and Research to Judit Gervain.

https://doi.org/10.1017/9781009295888.046 Published online by Cambridge University Press

Box 39.1 Chapter Overview

Summary

The current chapter reviews recent findings of infants' neural tracking of speech and relates these findings to subsequent grammar acquisition. Specifically, we discuss the potential role of prenatal exposure to speech for speech-tracking abilities at birth and its potential as an entry point into language structure in early language acquisition, in light of the prenatal prosodic-shaping model.

Implications

There is a gap in the literature when it comes to newborns' tracking of speech in the gamma band, corresponding to (sub-)phonemic elements of speech. Although several recent results are consistent with the predictions of the prenatal prosodic-shaping model, it can be empirically approached by addressing this gap.

Gains

Understanding the neural mechanisms that support grammar development is highly relevant to psycholinguistics/neurolinguistics, as much is yet unknown. The model presented in this chapter represents a potential framework for interpretation of the growing body of research on the role of neural oscillations in early speech processing.

REFERENCES

Abboub, N., Nazzi, T., and Gervain, J. (2016). Prosodic grouping at birth. *Brain and Language*, 162, 46–59. https://doi.org/10.1016/j.bandl.2016.08.002

Attaheri, A., Choisdealbha, Á. N., Di Liberto, G. M., et al. (2022). Delta- and theta-band cortical tracking and phase-amplitude coupling to sung speech by infants. *NeuroImage*, 247, 118698. https://doi.org/10.1016/j.neuroimage.2021.118698

Benavides-Varela, S., and Gervain, J. (2017). Learning word order at birth: A NIRS study. *Developmental Cognitive Neuroscience*, 25, 198–208. https://doi.org/10.1016/j.dcn.2017.03.003

Choisdealbha, Á. N., Attaheri, A., Rocha, S., et al. (2022). Cortical oscillations in pre-verbal infants track rhythmic speech and non-speech stimuli. Proceedings of the 46th Annual Boston University Conference on Language Development. https://doi.org/10.17863/CAM.86265

Coutinho, E., and Dibben, N. (2013). Psychoacoustic cues to emotion in speech prosody and music. *Cognition & Emotion*, 27(4), 658–684. https://doi.org/10.1080/02699931.2012.732559

DeCasper, A. J., and Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, 208(4448), 1174–1176. https://doi.org/10.1126/science.7375928

Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, 19(1), 158–164. https://doi.org/10.1038/nn.4186

https://doi.org/10.1017/9781009295888.046 Published online by Cambridge University Press

Eggermont, J. J., and Moore, J. K. (2011). Morphological and functional development of the auditory nervous system. In L. Werner, R. Fay, and A. Popper (eds.), *Human Auditory Development* (Vol. 42) (pp. 61–105). New York: Springer.

Gerhardt, K. J., and Abrams, R. M. (2000). Fetal exposures to sound and vibroacoustic stimulation. *Journal of Perinatology*, 20(1), S21–S30. https://doi.org/10.1038/sj.jp.7200446

Gervain, J., and Werker, J. F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature Communications*, 4(1), 1490. https://doi.org/10.1038/ncomms2430

Giraud, A.-L., and Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511–517. https://doi.org/doi.org/10.1038/nn.3063

Kalashnikova, M., Peter, V., Di Liberto, G. M., Lalor, E. C., and Burnham, D. (2018). Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. *Scientific Reports*, 8(1), 1–8. https://doi.org/10.1038/s41598-018-32150-6

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5(11), 831–843. https://doi.org/10.1038/nrn1533

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255(5044), 606–608. https://doi.org/10.1126/science.1736364

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., et al. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493), 979–1000. https://doi.org/10.1098/rstb.2007.2154

Liu, F., Jiang, C., Wang, B., Xu, Y., and Patel, A. D. (2015). A music perception disorder (congenital amusia) influences speech comprehension. *Neuropsychologia*, 66, 111–118. https://doi.org/10.1016/j.neuropsychologia.2014.11.001

Martinez-Alvarez, A., Benavides-Varela, S., Lapillonne, A., and Gervain, J. (2023). Newborns discriminate utterance-level prosodic contours. *Developmental Science*, 26(2), e13304. https://doi.org/10.1111/desc.13304

Mehler, J., Jusczyk, P., Lambertz, G., et al. (1988). A precursor of language acquisition in young infants. *Cognition*, 29(2), 143–178. https://doi.org/10.1016/0010-0277(88)90035-2

Menn, K. H., Männel, C., and Meyer, L. (2023). Does electrophysiological maturation shape language acquisition? *Perspectives on Psychological Science*, 18(6), 1271–1281. https://doi.org/10.1177/17456916231151584

Menn, K. H., Michel, C., Meyer, L., Hoehl, S., and Männel, C. (2022). Natural infant-directed speech facilitates neural tracking of prosody. *NeuroImage*, 251, 118991. https://doi.org/10.1016/j.neuroimage.2022.118991

Moon, C. (2017). Prenatal experience with the maternal voice. In M. Filippa, P. Kuhn, and B. Westrup (eds.), *Early Vocal Contact and Preterm Infant Brain Development* (pp. 25–37). Cham: Springer.

Moon, C., Cooper, R. P., and Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, 16(4), 495–500. https://doi.org/10.1016/0163-6383(93)80007-u

Moon, C., Lagercrantz, H., and Kuhl, P. K. (2013). Language experienced in utero affects vowel perception after birth: A two-country study. *Acta Paediatrica*, 102(2), 156–160. https://doi.org/10.1111/apa.12098

Nallet, C., and Gervain, J. (2021). Neurodevelopmental preparedness for language in the neonatal brain. *Annual Review of Developmental Psychology*, 3, 41–58. https://doi.org/10.1146/annurev-devpsych-050620-025732

Nazzi, T., and Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, 41(1), 233–243. https://doi.org/10.1016/s0167-6393(02)00106-1

https://doi.org/10.1017/9781009295888.046 Published online by Cambridge University Press

Nespor, M., and Vogel, I. (2007). *Prosodic Phonology: With a New Foreword* (Vol. 28). Berlin: de Gruyter.

Nespor, M., Shukla, M., van de Vijver, R., et al. (2008). Different phrasal prominence realizations in VO and OV languages. *Lingue e Linguaggio*, 7(2), 139–168. https://doi.org/10.1418/28093

Ortiz Barajas, M. C. O., Guevara, R., and Gervain, J. (2021). The origins and development of speech envelope tracking during the first months of life. *Developmental Cognitive Neuroscience*, 48, 100915. https://doi.org/10.1016/j.dcn.2021.100915

Ortiz Barajas, M. C. O., Guevara, R., and Gervain, J. (2023). Neural oscillations and speech processing at birth. *iScience*, 26(11), 108187. https://doi.org/10.1016/j.isci.2023.108187

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, 99(2), 143. https://doi.org/10.1037/0033-2909.99.2.143

Shi, R., Werker, J. F., and Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, 72(2), 11–21. https://doi.org/10.1016/s0010-0277(99)00047-5

Shukla, M., White, K. S., and Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-month-old infants. *Proceedings of the National Academy of Sciences*, 108(15), 6038–6043. https://doi.org/10.1073/pnas.1017617108

Soderstrom, M., Seidl, A., Nelson, D. G. K., and Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, 49(2), 249–267. https://doi.org/10.1016/s0749-596x(03)00024-x

Song, J. Y., Demuth, K., and Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *Journal of the Acoustical Society of America*, 128(1), 389–400. https://doi.org/10.1121/1.3419786

Thompson, W. F., and Balkwill, L.-L. (2006). Decoding speech prosody in five languages. *Semiotica*, 2006(158), 407–424. https://doi.org/10.1515/SEM.2006.017

Watson, D., and Gibson, E. (2005). Intonational phrasing and constituency in language production and comprehension. *Studia Linguistica*, 59(2–3), 279–300. https://doi.org/10.1111/j.1467-9582.2005.00130.x

Zentner, M., Grandjean, D., and Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, 8(4), 494. https://doi.org/10.1037/1528-3542.8.4.494

https://doi.org/10.1017/9781009295888.046 Published online by Cambridge University Press