Chapter 5. The role of prosody in early speech segmentation and word-referent mapping

Maria Teixidó, Clément François, Laura Bosch, Claudia Männel

CHAPTER 5

The role of prosody in early speech segmentation and word-referent mapping: Electrophysiological evidence

<u>Running title</u>: Prosody in word segmentation and mapping: ERP data

Maria Teixidó, University of Barcelona, Department of Cognition, Development and Educational Psychology

Clément François, University of Barcelona, Department of Cognition, Development and Educational Psychology, Cognition and Brain Plasticity Group (Bellvitge Biomedical Research Institute) IDIBELL, Institut de Recerca Pediàtrica Hospital Sant Joan de Déu

Laura Bosch, University of Barcelona, Department of Cognition, Development and Educational Psychology, Institut de Recerca Pediàtrica Hospital Sant Joan de Déu. Institut de Neurociències, University of Barcelona

Claudia Männel, Max Planck Institute for Human Cognitive and Brain Sciences, Department of Neuropsychology and Department of Neurology, University of Leipzig, Medical Faculty, Clinic for Cognitive Neurology

Abstract

This chapter reviews electrophysiological studies on early word-form segmentation and word-referent mapping, with a focus on the role of prosody in these early abilities closely related to vocabulary acquisition. First, we will review event-related brain potential (ERP) studies on word segmentation showing the impact of lexical stress cues, infant-directed speech (IDS) properties and melodic information on word-form extraction. Then, we will review research on word-referent mapping, revealing the scarcity of ERP studies specifically exploring the contribution of prosody in this domain. Throughout the chapter we will emphasize how electrophysiological methods offer a more fine-grained perspective of the brain processes supporting segmentation and mapping abilities, often revealing infants' sensitivities to auditory input before overt responses from behavioral methods can be obtained.

Introduction

Early in development, infants face the twin problems of segmenting the speech stream to extract linguistically relevant units and mapping these possible word-forms to their referents, the objects and events in the surrounding world. Developmental research has demonstrated the pivotal role of prosody in this process. Prosodic features, such as language rhythm, intonation, and lexical stress patterns, can provide infants with a starting point in early word learning (see Jusczyk, 1999; Goyet, Millotte, Christophe, and Nazzi, 2016; for reviews). Following the seminal work by Jusczyk and Aslin (1995), a large number of behavioral studies have revealed infants' gradually refined abilities to exploit prosodic cues, as well as other available cues from the input, to successfully segment the speech stream and build their first vocabulary.

Behavioral research methods involve overt attentional measures from video-recordings or eye-tracking procedures. In contrast, neuroimaging techniques, such as event-related brain potentials (ERPs), can register brain responses underlying early acquisition processes and even reveal infants' abilities where behavioral methods might not be sensitive enough. Following a brief introduction on the ERP methodology, this chapter offers an overview of infant ERP studies on the role of prosody in early word segmentation and word-referent mapping. More specifically, we will focus, in turn, on lexical stress patterns, some properties of infant-directed speech (IDS) such as accentuation and repetition, and melodic information, as cues to early word segmentation. Then, we will review word-referent mapping studies, pointing to the scarcity of ERP data specifically exploring the role of prosody in this domain. Throughout the sections of this chapter, we will emphasize how ERP data significantly extend our knowledge gathered from behavioral language acquisition research.

The method of event-related brain potentials (ERPs) in language acquisition research

Behavior-independent neuroimaging techniques, like the ERP methodology based on electroencephalography (EEG), offer insights into the mechanisms driving infants' behavior and their developmental progression. The ERP methodology has largely contributed to our understanding of early language acquisition processes and continues to be used as a powerful research tool complementing behavioral techniques (for an overview see Conboy, Rivera-Gaxiola, Silva-Pereyra, and Kuhl, 2008; Friederici, 2005; Friederici and Männel, 2013; Kuhl, 2010; Männel and Friederici, 2008).

The human brain constantly produces electrical activity. EEG captures this brain activity by recording ongoing scalp-level voltage fluctuations that originate from neurons' postsynaptic potentials. Electrophysiological research relies on the fact that external stimulation modulates EEG data. However, the effect of a single stimulation represents only a small part of the ongoing brain activity, such that the event-related portion of EEG recordings associated with a stimulus of interest needs to be extracted. Repeated stimulation and subsequent averaging across stimulus-time-locked EEG-epochs result in stable ERPs representing average stimulus processing and ruling out non-relevant brain activity.

The schematic ERP waveform in Figure 1 displays a sequence of positive- and negative-going stimulus-triggered voltage changes. The labeling of the different ERP components indicates their polarity and the latency of maximum amplitude relative to the onset of the stimulation. The N1(00) component, for example, refers to a negativity that can be observed around 100 ms after stimulus onset. In general, ERP components occurring early after stimulus onset (around 50–200 ms) are mainly modulated by sensory input features, while those occurring later (after 300 ms) reflect cognitive processes to a much larger degree.

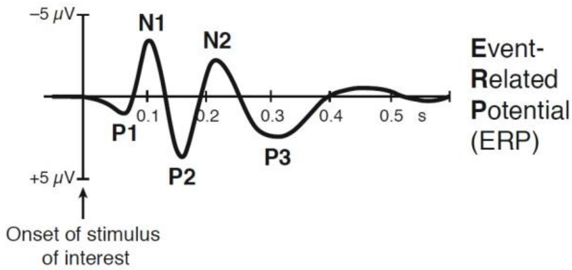

Figure 1. Schematic ERP waveform representing the average response to a class of stimuli. Positive- and negative-going voltage changes are labeled according to their amplitude polarity and latency from stimulus onset (adapted from Männel, 2008).

Different ERP paradigms have been developed to study early language acquisition during passive listening. For example, the Mismatch Negativity paradigm tests auditory discrimination of short stimuli, such as tones, phonemes or syllables, and involves the unexpected presentation of deviant stimuli in a series of repeated, standard stimuli. Familiarization-test paradigms are more suitable when longer or more complex stimuli are of interest, as is the case in word segmentation studies. The latter usually include a familiarization or learning phase, followed by a test phase with familiar and non-familiar elements, so that brain responses time-locked to the onset of the target stimuli can reveal familiarity recognition indicating successful word-form segmentation. Comparing ERPs elicited by different stimuli allows for inferences about the effect of experimental manipulations (Kutas, van Petten, and Kluender, 2006). Differences in amplitude can be the result of variations in processing demands or efficiency across experimental conditions, whereas latency differences can indicate the slowing down or speeding up of cognitive processing.

The ERP methodology, as a non-invasive and highly infant-suitable neuroimaging technique, has been successfully applied in early word segmentation and word-mapping studies. While its spatial resolution is weak, compared to hemodynamic techniques such as functional near-infrared spectroscopy (fNIRS), it has an excellent temporal resolution, comparable to whole-head magneto-encephalography (MEG). Thus, complementing behavioral data, ERP data have largely contributed to a better understanding of the time course of early word segmentation and mapping. The following sections will review available data from ERP studies, with a focus on research specifically exploring the role of prosodic factors on these crucial abilities for language acquisition.

Word segmentation

Finding the words in fluent speech is a fundamental ability in early language acquisition, linked to later vocabulary outcomes (Junge, Kooijman, Hagoort, and Cutler, 2012; Newman et al., 2006; Singh, Steven Reznick, and Xuehua, 2012). Although single-word utterances are not exceptional in the input to children, most words do not occur in isolation (Brent and Siskind, 2001). Infants thus need to segment the speech stream to extract the individual word-forms, exploiting any possible cues to word boundaries present in the speech input (see Thorson, this volume; de Carvalho, Dautriche, Millotte, and Christophe, this volume, for reviews). Behavioral studies indicate that word segmentation ability is present at around 7 to 10 months of age (Houston et al., 2000; Jusczyk, Houston, and Newsome 1999; Polka and Sundara, 2012). Word segmentation can be observed even earlier, by 6 months of age, for monosyllabic word-forms (Höhle and Weissenborn, 2003; Bosch, Figueras, Teixidó, and Ramon-Casas, 2013), and for words following an already familiar name or anchor word (Bortfeld, Morgan, Golinkoff, and Rathbun, 2005). In this context, ERP studies have provided relevant information on the segmentation process, not only by tracking the online processing, but also by offering a direct measure of the factors favoring or constraining early word-form segmentation. In the next sections, we first describe the electrophysiological indicators of word-form segmentation, followed by a review of ERP studies addressing (a) the role of lexical stress patterns in speech segmentation, (b) the effect of specific features of IDS in modulating the developmental onset of this ability, and (c) the contribution of melodic information (song) to the word segmentation process.

ERP indicators of word segmentation

Kooijman, Hagoort and Cutler (2005) designed the first ERP study on word segmentation with Dutch-learning 10-month-olds. They adapted the behavioral paradigm from Jusczyk and Aslin (1995), familiarizing infants with repetitions of strong-weak words and testing them on passages containing either the familiarized or novel words. Infants' ERPs in the familiarization phase showed a negative-going effect, with more negative ERP responses to the last repetitions compared to the first ones. This "familiarity" or repetition effect appeared from 200–500 ms post word onset at fronto-central recording sites. At test, a negative ERP effect over the left hemisphere was found from 350–500 ms, with larger amplitudes for familiarized target words than unfamiliarized words. The authors concluded that these negative ERP components (later termed N200-500, see Figure 2) were related to word-form familiarity through repetition, and recognition of the familiarized words, thus, indicating that segmentation of the target word-forms embedded in sentences had occurred.

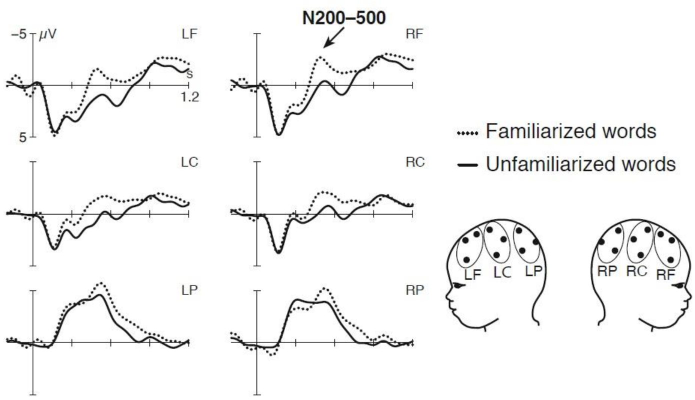

Figure 2. ERP responses of 12-month-olds to familiarized and unfamiliarized test words in a familiarization-test paradigm (modified from Männel and Friederici, 2013). L/RF = left/right frontal region, L/RC = left/right central region, L/RP = left/right posterior region.

Comparable negative-going ERP responses have been reported in studies on word recognition in older infants, both for known words (Mills, Coffey-Corina, and Neville, 1993; 1997; Thierry, Vihman, and Roberts, 2003) and for repetitions of pseudowords (von Koss Torkildsen et al., 2009). The N200–500 component can thus be taken as an ERP indicator of word segmentation. In the next sections, we review ERP evidence of the impact of different prosodic factors on infants' word segmentation abilities.

Language-specific prosodic cues to word segmentation: lexical stress patterns

Behavioral evidence has revealed that differences in the predominant stress pattern of languages may be responsible for the variation in the onset and the robustness of early word segmentation abilities (see Bhatara, Boll-Avetisyan, Höhle, and Nazzi, this volume). For instance, the metrical segmentation strategy, derived from the lexical stress properties of Dutch as the native language, would predict no difficulties in extracting disyllabic word-forms following the predominant stress pattern (e.g., strong-weak in Dutch) and less successful segmentation for the less frequent, weak-strong pattern. But behavioral evidence had failed to yield positive evidence of the expected strong-weak segmentation in 7.5-month-old Dutch-learning infants (Kuijpers, Coolen, Houston, and Cutler, 1998). A follow-up ERP study, however, adopting the same familiarization-test paradigm previously used with 10-month-olds (Kooijman, Hagoort, and Cutler, 2005), clarified this issue and confirmed successful segmentation of strong-weak words in 7-month-olds (Kooijman et al., 2013). Here, the N200–500 component emerged during familiarization, and, in addition, two different ERP patterns indicating familiarity recognition appeared at test. Some infants showed a right-frontal positive ERP effect at 350–450 ms to familiarized versus unfamiliarized words, while others showed a left-posterior negative ERP effect at 430–530 ms. When children's language outcomes were followed up at 3 years, those previously showing a negative-going ERP response obtained higher scores in vocabulary measures. Thus, polarity differences in the ERP could be taken to reflect different strategies underlying infants' word segmentation, linked to maturational variation and related to later vocabulary outcomes.

Evaluating the effect of lexical stress on word segmentation, Kooijman, Hagoort and Cutler (2009) examined whether Dutch-learning 10-month-olds would also be able to segment words that did not follow the predominant stress pattern, namely weak-strong words. ERPs of the familiarization revealed similar N200–500 effects as reported by Kooijman, Hagoort, and Cutler (2005), a familiarity effect that started while infants were still hearing the weak (first) syllable. Results of the test phase revealed a more variable pattern linked to the strong syllables of the target words. While infants were able to recognize repetitions of whole words occurring in isolation, they still relied on the strong syllable in more complex sentence contexts, failing to correctly segment the weak-strong words. These results thus differ from the ones obtained by Kooijman, Hagoort, and Cutler (2005), showing that 10-month-old infants recognized the whole strong-weak word and not only the strong syllable. Together, these findings highlight the important role of the predominant lexical-stress pattern of the native language in facilitating word segmentation.

We now turn to cross-linguistic differences in word segmentation, looking at French as a syllable-timed language that does not show word-initial stress, but phrase-final lengthening. Goyet, Schonen and Nazzi (2010) ran an ERP study to further evaluate the delayed disyllabic weak-strong word segmentation reported in a previous behavioral study with French 12-month-olds (Nazzi et al., 2006). Again, 12-month-old infants were familiarized with weak-strong words and then tested on sentences containing familiarized and novel words. Results from the familiarization phase showed an enhanced N200–500 occurring at 300–450 ms post word onset for later repetitions. Similar results were found at test, with the N200–500 effects of word-form familiarity resembling the ones previously reported by Kooijman, Hagoort, and Cutler (2005). Importantly, however, here they indicate successful word segmentation of weak-strong words by French-learning infants, while the 10-month-old Dutch-learning infants in Kooijman, Hagoort, and Cutler (2009) had failed at segmenting words with this specific stress pattern. While confirming French-learning infants' ability to segment disyllabic weak-strong words at the tested age of 12 months, it remains to be shown whether similar results (i.e., whole-word disyllabic segmentation and no independent final-syllable segmentation) can be found at an earlier age. Moreover, ERP data from infants acquiring other syllable-timed languages with variable lexical stress patterns (e.g., Catalan and Spanish) are still missing, limiting the overall conclusions about the function of cross-linguistically different prosodic features in infants' emerging segmentation abilities.

In summary, the described ERP studies show that word-level prosody affects early word segmentation in several significant ways. First, the fact that the language of interest has a more or less predominant word-stress structure seems to affect the onset of word segmentation. Second, words following the predominant pattern of a given language are segmented earlier than words that do not follow this pattern. Third, cross-linguistic rhythm differences can set some constraints on infants' ability to segment words longer than the smallest, monosyllabic elements. Beyond any cross-linguistic differences reported in the literature on infant word segmentation, it is important to highlight that 10-month-old infants show stable ERP responses indicating successful word segmentation; the N200–500 component, which has been previously reported for word-form familiarity. This finding suggests that infants at this early age are able to build up word-form representations from only a few word repetitions in the speech input. Moreover, ERP studies have identified a developmental transition between 7 and 10 months of age, from a positive-going to a negative-going familiarity effect in the ERP, suggesting different underlying mechanisms modulated by advancing brain maturation and sustained language exposure.

Properties of infant-directed speech (IDS) supporting word segmentation: Prosodic emphasis on sentence-embedded words

Across languages, adults have been shown to use infant-directed speech (IDS) when communicating with infants (Fernald et al., 1989). IDS is described as a speech register with especially salient prosodic features, such as more exaggerated pitch contours and emphasis on specific words, produced with higher pitch and intensity levels and longer duration. Moreover, IDS features a slower speech rate, shorter mean length of utterance, more prosodic repetition, higher frequency of specific words and syntactic patterns, and simpler grammatical structures (Fernald, 1992). The specific acoustic features of IDS have been found to facilitate language acquisition, and, importantly, word segmentation (Thiessen, Hill, and Saffran, 2005; see also Cristia, 2013; Soderstrom, 2007, for reviews).

While infants show an initial strong preference for IDS over ADS, this preference for prosodically enhanced speech seems to reduce during the first year of life (Newman and Hussain, 2006). Moreover, parents have been shown to adjust their IDS register to infants' age, such that IDS features are especially pronounced during infants' first months of life, but become less salient later on (Kitamura and Burnham, 2003; Stern, Spieker, Barnett, and MacKain, 1983). Thus, it seems that both infants' use of IDS in language learning and IDS features provided in the speech input change as a function of infants' age. In this context, Männel and Friederici (2013, 2010), using electrophysiological measures, systematically explored the age-dependent role of different speech cues that are typically used by caregivers in word-learning situations: prosodic accentuation and repetition (see Aslin, Woodward, LaMendola, and Bever, 1996; Bernstein-Ratner, 1996; Fernald and Mazzie, 1991; Fernald and Morikawa, 1993). The authors tested German-learning infants at 6, 9 and 12 months of age in a familiarization-test paradigm and manipulated the degree of prosodic modulation and number of repetitions of sentence-embedded target words. Infants were first familiarized with different sentences containing the same target word, presented with or without an additional, naturally produced prosodic emphasis (for acoustic analyses, see Männel and Friederici, 2013). The test phase consisted of tokens of the familiar target word and an unfamiliar control word. ERP responses to sentence-embedded targets collected during familiarization differed depending on infants' age. At 6 months, infants showed a fronto-central positive shift at 200–400 ms post word onset for accentuated versus unaccentuated words, but no effect of repetition. At 9 months, infants also showed an early occurring positive-going ERP effect of prosodic realization and, in addition, an N200–500 word-repetition effect. Finally, 12-month-old infants did not show an effect of prosodic realization on target word processing, but a sustained N200–500 as an effect of repetition. These results suggest that at younger ages, prosodic input cues drive infants' word processing, while with increasing age repetition takes more effect. When infants were subsequently tested on their word recognition, ERP responses to familiarized and unfamiliarized words were compared as a function of prosodic realization during familiarization. As displayed in Figure 3, test results revealed that 6-month-olds were only able to recognize words that had received an additional prosodic emphasis during familiarization, showing a positive-going familiarity response in the ERP, starting at around 500 ms post word onset. Nine-month-old infants showed two familiarity effects in the ERP: an N200–500 to familiarized versus unfamiliarized words independently of previous prosodic realization; and a late negativity, only for words that were accentuated in the familiarization phase. Finally, 12-month-old infants also showed a sustained N200–500 effect that occurred for all familiarized compared to unfamiliarized words, independently of accentuation. Taken together, these data suggest that infants recognize familiar words, relying on different speech input cues provided by the speaker when *teaching* infants new words. While younger infants primarily benefit from the additional prosodic emphasis on sentence-embedded words, older infants primarily rely on word repetition, and recognize familiar words independently of their prosodic realization during the learning situation.

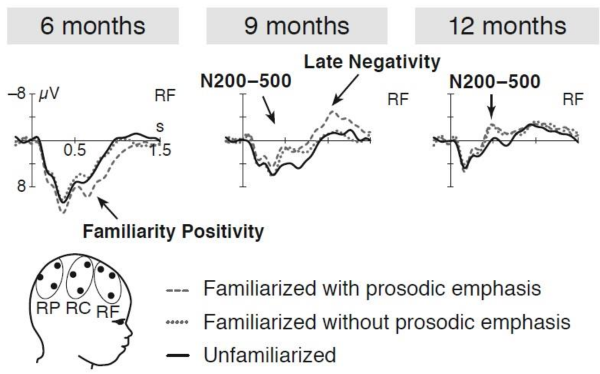

Figure 3. Developmental changes of ERP responses indicating word-form recognition as a function of prosodic manipulation during familiarization (adapted from Männel and Friederici, 2013). RF = right frontal region.

The ERP study by Männel and Friederici (2013) further contributes to the idea that prosodic information in the speech input enables infants' initial word segmentation, and confirms previous behavioral research showing that infants at 6 months of age make use of prosodic cues aligned with word boundaries to segment disyllabic words (Shukla, White, and Aslin, 2011). Moreover, Männel and Friederici (2013) observed a developmental positivity-negativity transition in the ERP responses indicating word recognition, with the latter indicating a more mature process (see also Kooijman Hagoort, and Cutler, 2005; Kooijman et al., 2013). This transition in the ERP markers of infant word segmentation can be attributed to infants' growing linguistic experience and parallel brain maturation. More specifically, at around 6 months of age, myelination and neurogenesis of the auditory cortex undergo tremendous changes (Moore and Guan, 2001; Moore and Linthicum, 2007) that can influence the latency and polarity of early sensory ERP components (see Eggermont and Moore, 2012). These maturational changes may also affect the polarity of infants' word recognition responses. Thus, the current example demonstrates how ERP studies additionally deliver information about the underlying brain mechanisms that trigger infants' word segmentation abilities.

From IDS to song: How early can melodic cues contribute to word segmentation?

Very early in life, infants are exposed to both language and music (Brandt, Gebrian, and Slevc, 2012), and both domains share important similarities in terms of structural information and involved cerebral resources (Heffner and Slevc, 2015; Patel, 2008). But, is music, and specifically song, relevant for speech segmentation and word learning? Song and IDS registers exhibit several similar acoustic and structural characteristics (Fernald, 1992; Papousek, Papousek, and Bornstein 1985). For example, many of the IDS features, such as exaggerated pitch contours, emphasis on specific words, higher pitch and intensity levels, longer duration, slower rate, shorter utterances and repetition, are also found in songs for infants. Infants show a clear preference for ID songs over AD songs (Unyk, Trehub, Trainor, and Schellenberg, 1992), and ID melodic information has been shown to specifically benefit speech processing as well as language acquisition (Gervain and Werker, 2013; Lebedeva and Kuhl, 2010; Schön et al., 2008; Thiessen and Saffran, 2009).

Assessing a beneficial effect of song on word segmentation, Schön and colleagues (2008) compared segmentation of a spoken, flat-contour stream and a sung stream in adults. While the level of performance was at chance after exposure to the spoken stream, successful word recognition was observed following exposure to the sung stream. These results suggest that redundancy in statistical musical and linguistic structures benefits segmentation. Based on these findings in adults (Schön et al., 2008) and evidence showing that prosodic cues facilitate segmentation in 7-month-old infants (Thiessen, Hill, and Saffran, 2005) and even neonates (Bosseler, Teinonen, Tervaniemi, and Huotilainen, 2016), it has been recently tested whether melodic information of ID song may benefit speech segmentation in newborns (François et al., 2017). The authors recorded EEG in a group of 2-day-old sleeping neonates during a familiarization-test paradigm with continuous artificial speech streams that were either spoken or sung. In the spoken condition, the continuous streams consisted of 100 repetitions of four pseudowords, with transitional probabilities between adjacent syllables as the only cues to word boundaries. In the sung condition, the language streams were identical except that each syllable had a unique pitch from the diatonic scale. Therefore, the resulting sung streams contained two types of overlapping cues signaling word boundaries: the statistical cues (drop in transitional probabilities between adjacent syllables) together with the pitch interval changes that systematically occurred at word boundaries only. In each condition, spoken and sung, newborns were first exposed to a 3.5 min familiarization stream followed by a test phase in which test items that violated the statistically-defined syllabic order of the words were inserted at random positions allowing for an evaluation of the implicit detection of structural violations.

During familiarization, a comparison between the ERPs of the first and the second half revealed the presence of a fronto-central negative deflection at 400–600 ms post word onset in both spoken and sung conditions. However, while in the spoken condition the fronto-central negativity reached maximum amplitude during the second half, its amplitude was already at maximum during the first half in the sung condition, suggesting a faster segmentation process. To determine whether familiarization was sufficient to establish pseudoword representations and detect violations of syllabic order in the test phase, ERPs to the newly learned pseudowords were compared to those elicited by the violation items. Evidence of successful detection of structural violations was found in the sung condition, but not in the spoken condition. Compared to the newly learned pseudowords, test items violating the syllabic structure elicited an early positivity in the 250–400 ms time window. This positivity resembled the mismatch response consistently observed in infants for speech and non-speech stimuli (see Cheour-Luhtanen et al., 1995; Mueller, Friederici, and Männel, 2012; Stefanics et al., 2009). Thus, the test results suggest that sung streams facilitated the building of robust word-form memory traces and triggered strong enough expectations to correctly detect online structural violations. In addition, these results confirm the importance of the fast-emerging fronto-central negativity during the learning phase for subsequent word-form recognition, as neonates were sensitive to structural violations in the sung condition, but not in the spoken one. These results provide direct evidence of the benefit of songs during the initial steps of language acquisition, thus highlighting the relevance of both speech prosody and melodic information to break the speech code and begin identifying basic linguistic units. Moreover, the benefit of song on speech segmentation supports the idea that musical interventions might be a powerful tool to facilitate language acquisition (Flaugnacco et al., 2015; François, Grau-Sánchez, Duarte, and Rodriguez-Fornells, 2015; Zhao and Kuhl, 2016).

Word-referent mapping

So far, we have focused on research evaluating how infants extract word-forms from fluent speech, greatly aided by prosodic cues. Word learning, however, goes beyond word segmentation, because newly segmented word-forms will eventually become labels for objects, people, and events in children's environment. Thus, establishing word-referent associations and storing these as memory representations are crucial steps children have to accomplish in building a first receptive vocabulary. Behavioral studies have shown that starting at 6–8 months of age, infants are able to successfully establish word-to-world mappings in laboratory settings (Gogate, Bolzani, and Betancourt, 2006; Matatyaho and Gogate, 2008), and even at 2 months of age when presented with distinct syllable-object motion pairings (Gogate, Prince, and Matatyaho, 2009). Regarding semantic representations, Bergelson and Swingley (2012) even found that 6-month-old infants already know the meaning of very frequent words, having established some word-object associations from their daily experience. These abilities have also been explored with the ERP method in infants at different ages, offering further insight into the underlying brain mechanisms and their developmental progression. In the following sections, we will first review ERP evidence of infants' early word knowledge and their ability to establish novel word-object associations. We will then turn to the discussion of the role of prosody in infants' word-referent mapping skills, revealing the scarcity of ERP studies in this field and the need for further neuroimaging research.

ERP indicators of word-referent mapping

Friedrich and Friederici (2004) were the first to investigate infants' brain mechanisms underlying the processing of words in meaningful contexts. To tackle the link between words and their referents by means of ERPs, the authors used a cross-modal picture-word priming paradigm with German-speaking 19-month-olds and adults (see also Friedrich and Friederici 2005a; 2005b; 2008; 2011; 2015; for review, see Männel, 2008). Participants were presented with pictures of objects and accompanying IDS words that correctly or incorrectly labeled the objects. Crucially, the objects' presentation preceded the label onset, such that the picture content could prime subsequent word processing. Thus, if participants knew the objects (from previous experimental familiarization or existing memory representations), this knowledge could trigger the expectation of the corresponding word and thus ease phonological word-form processing and lexical-semantic processing upon the label's presentation.

The adult ERP data collected in the picture-word priming paradigm revealed two processes indicating successful word processing by priming from the preceding picture context (Friedrich and Friederici, 2004): First, facilitated phonological word-form processing as reflected in a frontal N200–500 effect (i.e., word-form familiarity) for picture-matching words compared to non-matching ones, proving established word-object associations. And second, successful lexical-semantic word comprehension as reflected in an N400 effect (i.e., semantic integration efforts) for words not matching the picture compared to matching ones, due to semantic expectancy violations given referential connections between words and the semantic picture context. The N400 describes a negativity that occurs at around 400 ms relative to stimulus onset and is more pronounced the more semantically unfamiliar, unexpected, or non-matching an event is, given the current semantic context or semantic knowledge (see Holcomb, 1993). Evidence of the N400 in infancy research would thus be a valid indicator of infants' emerging ability to understand the referential meaning of words. It is important to note that although both N200–500 and N400 effects are partly overlapping in their time of occurrence, they can still be differentiated based on their different scalp distributions: The N200–500 typically occurs fronto-centrally, while the N400 occurs broadly distributed over the scalp, with a centro-parietal focus.

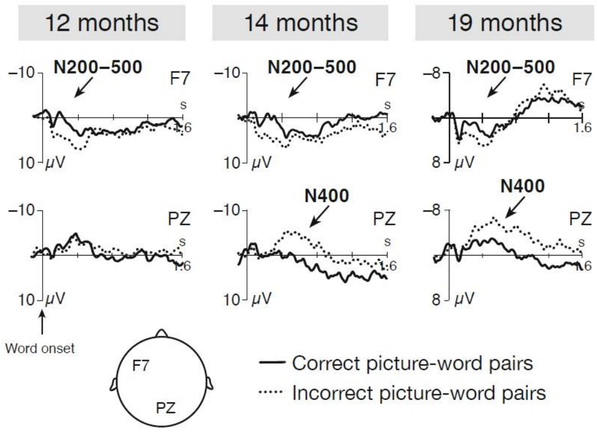

Figure 4. ERP responses for word processing in a cross-modal picture-word priming paradigm at electrodes F7 and PZ (adapted from Friederici, 2005; Friedrich and Friederici 2004; Friedrich and Friederici, 2005a, b).

Crucially, using the same cross-modal picture-word priming paradigm in 19-month-olds, Friedrich and Friederici (2004) observed comparable ERP effects in infants as reported for adults. Infants did not only show ERP indicators of word-form familiarity from cross-modal phonological priming (N200–500), but also the establishment of referential meaning (N400) (Figure 4). When analyzing these ERP effects in light of children's lexical knowledge (as reported by their parents), there was no correlation of children's receptive vocabulary and the N200–500 effect, but a correlation with the N400-like effect. In other words, only infants with higher comprehension abilities showed an N400 similar to the one found in adults, confirming the interpretation of this infant ERP component as indicating referential processes.

Following this seminal work, Friedrich and colleagues conducted a series of ERP experiments in German exploring the developmental progression of the emerging ability to form referential connections during early infancy. In these follow-up ERP studies, the authors were able to show that even 14-month-olds displayed recognition of word-forms (N200–500) and were able to form referential connections between objects and labels (N400) similarly to older infants and adults (Friedrich and Friederici, 2005b; 2008; Figure 4). However, younger infants at 12 and 6 months of age, while initiating word-object associations both phonologically (N200–500) and semantically (N400) during the familiarization phase, only showed phonologically-driven associations (N200–500) when tested later, suggesting memory limitations preventing full mapping representations (Friedrich and Friederici, 2005a; 2011; Figure 4). Finally, 3-month-old infants were found to generate word-form representations during familiarization (N200–500), as well as an early perceptually-driven form of association between objects and words (Friedrich and Friederici, 2015). The latter ERP effect appeared as centro-parietal negativity at 500–1000 ms post word onset to matching versus non-matching picture-word pairs. This late centro-parietal negativity might indicate a precursor learning mechanism that enables infants to establish first connections between the perceptual representations of objects and words, forming their first protowords (see Nazzi and Bertoncini, 2003). However, 3-month-olds did not reveal any sign of retention of these associations when tested one day later, making evident the difficulties infants at this age have in retaining this information beyond the immediate encoding stage. Together, these results suggest that the neural mechanisms underlying the building of a first vocabulary are present from early on, while the capacity for creating full referential connections and their successful retention is still highly constrained below 14 months of age.

In summary, the existing ERP studies on word-referent mapping across infancy have reliably revealed two ERP components: First, a fronto-centrally distributed N200–500 that indicates word-form familiarity and evidences word-object associations at a phonological level. Second, a centro-parietal N400 that indicates semantic integration efforts and evidences word-object mapping at a semantic level. Importantly, the association and reference processes underlying these components seem to work semi-independently, and show a different developmental onset. Thus, the reported findings impressively illustrate how the use of electrophysiological measures at early ages not only enables us to successfully delineate the developmental onset and progression of infants' mapping abilities, but also to differentiate the underlying processes.

The contribution of prosody to word-referent mapping: From behavioral to ERP studies

Comparable to behavioral research exploring the role of prosodic features in infants' successful word segmentation, there are an impressive number of behavioral studies addressing the effect of prosody in word-object mapping. Factors such as accentuation (Grassmann and Tomasello, 2007; 2010), lexical stress (Curtin, Campbell, and Hufnagle, 2012; Graf-Estes and Bowen, 2013), specific pitch accents (Thorson and Morgan, 2015), and the role of IDS (Graf-Estes and Hurley, 2013; Ma, Golinkoff, Houston, and Hirsh-Pasek, 2011) have been analyzed and connected to word mapping abilities (see Thorson, this volume). For example, Graf-Estes and Hurley (2013) examined whether the pronounced prosodic features of IDS would enhance word-object associations in 17.5-month-olds. The authors habituated infants to pairs of novel objects and novel words, either in IDS or ADS. Looking-time measures revealed that infants were only able to map words onto objects for labels produced in IDS, but not in ADS. Crucially, the advantage of IDS only took effect when the realization of IDS labels involved prosodic variability. Focusing on a different prosodic property, Curtin, Campbell and Hufnagle (2012) found that lexical stress cues can also guide label-referent associations. In their study, English-learning 16-month-olds were already showing a bias to associate weak-strong patterns with action labels, a language-specific bias consistent with the stress pattern properties of their native language. Thus, behavioral research has delivered significant indications of prosodic cues also driving infants' initial word mapping skills. However, a systematic examination of how prosodic input features affect referential processes by means of electrophysiological measures is still missing. As exemplified above, ERP studies on infants' word-referent mapping have used IDS material (Friedrich and Friederici, 2004; 2005a; 2005b; 2006; 2008; 2011). Yet, these studies have never directly contrasted the effect of IDS and ADS, or manipulated other prosodic speech input features, despite promising behavioral evidence. Thus, ERP studies aiming at exploring the impact of prosody on infants' developing mapping skills are clearly needed.

Another line of ERP research that needs to be strengthened is the investigation of word segmentation and mapping as a joint process, also assessing the impact of speech input features. Naturally, when infants start to segment word-forms from fluent speech, they eventually conceive them as labels to referents in their surrounding world. Recent behavioral research has shown that from an early age, infants can simultaneously segment and map simple, statistically defined word-forms, but only if prosody favors this process (Shukla, White, and Aslin, 2011). Investigating simultaneous segmentation and mapping in a more ecological learning situation, Teixidó and Bosch (2014) tested 4-, 6- and 9-month-old infants on natural language stimuli. Interestingly, the authors observed successful word segmentation and mapping only for infants at 9 months, but not at younger ages. Here, a follow-up ERP study has set out to explore the underlying neural mechanisms driving the success (or failure) of infants' segmentation and mapping skills at young ages (Männel, Teixidó, Bosch, Friederici, & Friedrich, 2017). Importantly, this study manipulated the prosodic realization of target words by placing them at different sentence positions. Preliminary ERP results indicate that while positioning did not affect infants' segmentation ability, it impacted on infants' mapping success, because only targets at prosodically salient edge positions enabled mapping. This ERP study represents a relevant first research step, yet there are clearly more studies needed that assess the neural mechanisms supporting infants' emerging mapping skills and the role played by prosody at different developmental stages.

Conclusions

In this chapter, we have reviewed the role of prosody in infants' word learning, evaluating their emerging abilities of word segmentation and word-referent mapping. Here we have focused on electrophysiological studies that capture the underlying neural mechanisms of infants' emerging abilities, accessible from an early age, when behavioral assessments can be challenging. Current evidence from ERP studies suggests that there are at least two different electrophysiological indicators related to infant word learning. First, an N200–500 component, that is, a word-form familiarity effect characterized by a fronto-lateral negativity in the ERP at 200–500 ms post word onset, elicited when a given phonological sequence is familiar. Second, an N400 component, that is, a semantic (violation) effect characterized by a centro-parietal negativity at around 400 ms post word onset, elicited when the meaning of a word does not match the semantic (picture) context. Importantly, the developmental onset and the amplitude of these two ERP components can predict later language development, particularly children's vocabulary. Crucial for developmental research, both ERP components can be utilized as indicators of infants' emerging word learning abilities to draw a more fined-grained picture of the progression of word-form segmentation and the establishment of referential meaning across infancy. Over and above the illustration of infants' advancing word learning skills by means of ERPs, the data presented here have highlighted the idea that prosody is a key factor in getting infants started on detecting word forms in fluent speech and conceiving them as referents to their surrounding world.

Acknowledgments

The present chapter was supported by the Spanish *Ministerio de Ciencia e Innovación* (projects PSI-2014-55105-P to L.B. and PSI-2015-69132-P to C.F.), the German Research Foundation (project MA 6897/2-1 to C.M.) and the Max Planck

Society (C.M.). We would like to thank the editors and two anonymous reviewers for their helpful comments and editorial improvements.

References

Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In J. L. Morgan, & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 117-134). Mahwah, NJ: Erlbaum.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, 109(9), 3253-3258.

Bernstein-Ratner, N. (1996). From Signal to syntax: But what is the nature of the signal. In J. L. Morgan, & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 135-150). Mahwah, NJ: Erlbaum.

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me. Familiar names help launch babies into speech-stream segmentation. *Psychological Science*, 16(4), 298-304.

Bosch, L., Figueras, M., Teixidó, M., & Ramon-Casas, M. (2013). Rapid gains in segmenting fluent speech when words match the rhythmic unit: Evidence from infants acquiring syllable-timed languages. *Frontiers in Psychology*, 4(106). Retrieved from https://doi.org/10.3389/fpsyg.2013.00106.

Bosseler, A. N., Teinonen, T., Tervaniemi, M., & Huotilainen, M. (2016). Infant-directed speech enhances statistical learning in newborn infants: An ERP study. *PloS one*, 11(9), e0162177.

Brandt, A. K., Slevc, R., & Gebrian, M. (2012). Music and early language acquisition. *Frontiers in Psychology*, 3(327).

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, 81(2), B33-B44.

Cheour-Luhtanen, M., Alho, K., Kujala, T., Sainio, K., Reinikainen, K., Renlund, M., Aaltonen, O., Eerola, O., & Näätänen, R. (1995). Mismatch negativity indicates vowel discrimination in newborns. *Hearing Research*, 82(1), 53-58.

Commented [u1]: PR: Please write all authors.

Conboy, B. T., Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2008). Event-related potential studies of early language processing at the phoneme, word, and sentence levels. *Early Language Development*, 5, 23-64.

Cristia, A. (2013). Input to language: The phonetics and perception of infant-directed speech. *Language and Linguistics Compass*, 7(3), 157-170.

Curtin, S., Campbell, J., & Hufnagle, D. (2012). Mapping novel labels to actions: How the rhythm of words guides infants' learning. *Journal of Experimental Child Psychology*, 112(2), 127-140.

Eggermont, J. J., & Moore, J. K. (2012). Morphological and functional development of the auditory nervous system. In L. A. Werner, R. R. Fay, & A N. Popper (Eds.), *Human Auditory Development* (pp. 61-105). New York: Springer.

Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In J. H. Berkow (Ed.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (pp. 391-428). London: Oxford University Press.

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, 27(2), 209-221.

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development*, 64(3), 637-656.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16, 477-501.

Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., & Schön, D. (2015). Music training increases phonological awareness and reading skills in developmental dyslexia: a randomized control trial. *PloS one*, 10(9), e0138715.

François, C., Grau-Sánchez, J., Duarte, E., & Rodriguez-Fornells, A. (2015). Musical training as an alternative and effective method for neuro-education and neuro-rehabilitation. *Frontiers in Psychology*, 6(475).

François, C., Teixidó, M., Agut, T., Bosch, L. & Rodriguez-Fornells, A. (2017). Enhanced neonatal brain responses to sung streams predict vocabulary outcomes by age 18 months. *Scientific Reports*, 7(1), 12451.

Friederici, A. D. (2005). Neurophysiological markers of early language acquisition: from syllables to sentences. *Trends in Cognitive Sciences*, 9(10), 481-488.

Friederici, A. D., & Männel, C. (2013). Neural correlates of the development of speech perception and comprehension. In K. N. Ochsner, & S. M. Kosslyn (Eds.), *The Oxford Handbook of Cognitive Neuroscience* (Vol. 1, pp. 171-192). New York: Oxford University Press.

Friedrich, M., & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience*, 16(8), 1465-1477.

Friedrich, M., & Friederici, A. D. (2005a). Phonotactic knowledge and lexical-semantic processing in one-year-olds: Brain responses to words and nonsense words in picture contexts. *Journal of Cognitive Neuroscience*, 17(11), 1785-1802.

Friedrich, M., & Friederici, A. D. (2005b). Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. *Neuroreport*, 16(6), 653-656.

Friedrich, M., & Friederici, A. D. (2008). Neurophysiological correlates of online word learning in 14-month-old infants. *Neuroreport*, 19(18), 1757-1761.

Friedrich, M., & Friederici, A. D. (2011). Word learning in 6-month-olds: Fast encoding–weak retention. *Journal of Cognitive Neuroscience*, 23(11), 3228-3240.

Friedrich, M., & Friederici, A. D. (2015). The origins of word learning: Brain responses of three-month-olds indicate their rapid association of objects and words. *Developmental Science*, 20(2), e12357.

Gervain, J., & Werker, J. F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature Communications*, 4, 1490-1496.

Gogate, L. J., Bolzani, L. H., & Betancourt, E. A. (2006). Attention to maternal multimodal naming by 6- to 8-month-old infants and learning of word–object relations. *Infancy*, 9(3), 259-288.

Gogate, L. J., Prince, C. G., & Matatyaho, D. J. (2009). Two-month-old infants' sensitivity to changes in arbitrary syllable–object pairings: The role of temporal synchrony. *Journal of Experimental Psychology: Human Perception and Performance*, 35(2), 508.

Goyet, L., de Schonen, S., & Nazzi, T. (2010). Words and syllables in fluent speech segmentation by French-learning infants: An ERP study. *Brain Research*, 1332, 75-89.

Goyet, L., Millotte, S., Christophe, A., & Nazzi, T. (2016). Processing continuous speech in Infancy: From major prosodic units to isolated word forms. In W. Snyder, & J. Pater, (Eds.), *The Oxford Handbook of Developmental Linguistics*, (Vol. 1, pp. 133- 156). New York: Oxford University Press.

Graf-Estes, K., & Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. *Infancy*, 18(5), 797-824.

Graf-Estes, K., & Bowen, S. (2013). Learning about sounds contributes to learning about words: Effects of prosody and phonotactics on infant word learning. *Journal of Experimental Child Psychology*, 114(3), 405-417.

Grassmann, S., & Tomasello, M. (2007). Two-year-olds use primary sentence accent to learn new words. *Journal of Child Language*, 34(03), 677-687.

Grassmann, S., & Tomasello, M. (2010). Prosodic stress on a word directs 24-month-olds' attention to a contextually new referent. *Journal of Pragmatics*, 42(11), 3098-3105.

Heffner, C. C., & Slevc, L. R. (2015). Prosodic structure as a parallel to musical structure. *Frontiers in Psychology*, 6(1962).

Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, 6(2), 122-127.

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, 30(1), 47-61.

Houston, D. M., Jusczyk, P. W., Kuijpers, C., Coolen, R., & Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin and Review*, 7(3), 504-509.

Junge, C., Kooijman, V., Hagoort, P., & Cutler, A. (2012). Rapid recognition at 10 months as a predictor of language development. *Developmental Science*, 15(4), 463-473.

Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences*, 3(9), 323-328.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29(1), 1-23.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39(3), 159-207.

Kitamura, C., & Burnham, D. (2003). Pitch and communicative intent in mother's speech: Adjustments for age and sex in the first year. *Infancy*, 4(1), 85-110.

Kooijman, V., Hagoort, P., & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research*, 24, 109-116.

Kooijman, V., Hagoort, P., & Cutler, A. (2009). Prosodic structure in early word segmentation: ERP evidence from Dutch ten-month-olds. *Infancy*, 14(6), 591-612.

Kooijman, V., Junge, C., Johnson, E. K., Hagoort, P., & Cutler, A. (2013). Predictive brain signals of linguistic development. *Frontiers in Psychology*, 4(25).

Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron*, 67(5), 713-727.

Kuijpers, C., Coolen, R., Houston, D., & Cutler, A. (1998). Using the head-turning technique to explore cross-linguistic performance differences. In C. Rovee-Collier, L. Lipsitt, & H. Hayne (Eds.), *Advances in Infancy Research* (Vol. 12, pp. 205–220). London: Ablex.

Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II (1994–2005). In M. Traxler, & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics*. (Vol. 1, pp. 83-143). San Diego, CA: Academic Press.

Lebedeva, G. C., & Kuhl, P. K. (2010). Sing that tune: Infants' perception of melody and lyrics and the facilitation of phonetic recognition in songs. *Infant Behavior and Development*, 33(4), 419-430.

Ma, W., Golinkoff, R. M., Houston, D. M., & Hirsh-Pasek, K. (2011). Word learning in infant- and adult-directed speech. *Language Learning and Development*, 7(3), 185-201.

Männel, C. & Friederici, A. D. (2008). Event-related brain potentials as a window to children's language processing: From syllables to sentences. In I. A. Sekerina, E. M. Fernández, H. Clahsen (Eds.), *Developmental Psycholinguistics: On-line Methods in Children's Language Processing* (pp. 29-72). Amsterdam: John Benjamins.

Männel, C. & Friederici, A. D. (2010). Prosody is the key: ERP studies on word segmentation in 6- and 12-month-old children. *Journal of Cognitive Neuroscience*, Supplement, 261.

Männel, C., & Friederici, A. D. (2013). Accentuate or repeat? Brain signatures of developmental periods in infant word recognition. *Cortex*, 49(10), 2788-2798.

Männel, C. (2008). The method of event-related brain potentials in the study of cognitive processes. A tutorial. In A. D. Friederici, & G. Thierry (Eds.), *Early*

Language Development: Bridging Brain and Behaviour, Bd. Trends in language acquisition research (TiLAR) (pp. 1-22). Amsterdam/Philadelphia: John Benjamins.

Männel, C., Teixidó, M., Bosch, L., Friederici, A. D. & Friedrich, M. (November, 2017). Sentence prosody cues object category learning at 6 months. Paper presented at the Ninth Annual Meeting of the Society for the Neurobiology of Language (SNL), Baltimore, USA.

Matatyaho, D. J., & Gogate, L. J. (2008). Type of maternal object motion during synchronous naming predicts preverbal infants' learning of word–object relations. *Infancy*, 13(2), 172-184.

Mills, D. L., Coffey-Corina, S. A., & Neville, H. J. (1993). Language acquisition and cerebral specialization in 20-month-old infants. *Journal of Cognitive Neuroscience*, 5(3), 317-334.

Mills, D. L., Coffey-Corina, S. A., & Neville, H. J. (1997). Language comprehension and cerebral specialization from 13 to 20 months. *Developmental Neuropsychology*, 13(3), 397-445.

Moore, J. K., & Guan, Y. L. (2001). Cytoarchitectural and axonal maturation in human auditory cortex. *Journal of the Association for Research in Otolaryngology*, 2(4), 297-311.

Moore, J. K., & Linthicum Jr, F. H. (2007). The human auditory system: a timeline of development. *International Journal of Audiology*, 46(9), 460-478.

Mueller, J. L., Friederici, A. D., & Männel, C. (2012). Auditory perception at the root of language learning. *Proceedings of the National Academy of Sciences*, 109(39), 15953-15958.

Nazzi, T., & Bertoncini, J. (2003). Before and after the vocabulary spurt: two modes of word acquisition? *Developmental Science*, 6(2), 136-142.

Nazzi, T., Iakimova, G., Bertoncini, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, 54(3), 283-299.

Newman, R. S., & Hussain, I. (2006). Changes in preference for infant-directed speech in low and moderate noise by 4.5- to 13-month-olds. *Infancy*, 10(1), 61-76.

Newman, R., Ratner, N. B., 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(4), 643-655.

Papousek, M., Papousek, H., & Bornstein, M. (1985) The naturalistic vocal environment of young infants: On the significance of homogeneity and variability in parental speech. In T.M. Field, & N.A. Fox (Eds.), *Social Perception in Infants* (pp. 269-297). Norwood, NJ: Ablex.

Patel, A. D. (2008). Science & music: talk of the tone. *Nature*, 453(7196), 726-727.

Polka, L., and Sundara, M. (2012). Word segmentation in monolingual infants acquiring Canadian English and Canadian French: native language, cross-dialect, and cross-language comparisons. *Infancy*, 17(2), 198-232.

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition*, 106(2), 975-983.

Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences*, 108(15), 6038-6043.

Singh, L., Steven Reznick, J., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: a longitudinal analysis. *Developmental Science*, 15(4), 482-495.

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, 27, 501-532.

Stefanics, G., Háden, G. P., Sziller, I., Balázs, L., Beke, A., & Winkler, I. (2009). Newborn infants process pitch intervals. *Clinical Neurophysiology*, 120(2), 304-308.

Stern, D. N., Spieker, S., Barnett, R. K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language*, 10(01), 1-15.

Teixidó, M., & Bosch, L. (2014, July). *Early segmentation and word mapping in a natural language: evidence from 4-, 6- and 9-months-olds*. Poster presented at the XIX Biennial International Conference on Infant Studies (ICIS), Berlin, Germany.

Thierry, G., Vihman, M., & Roberts, M. (2003). Familiar words capture the attention of 11-month-olds in less than 250 ms. *Neuroreport*, 14(18), 2307-2310.

Thiessen, E. D., & Saffran, J. R. (2009). How the melody facilitates the message and vice versa in infant learning and memory. *Annals of the New York Academy of Sciences*, 1169(1), 225-233.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, 7(1), 53-71.

Thorson, J. C. & Morgan, J. L. (2015). Acoustic correlates of information structure in child and adult speech. In E. Grillo & K. Jepson (Eds.), *Proceedings of the 39th*

Annual Boston University Conference on Language Development (pp. 411-423). Somerville, MA: Cascadilla Press.

Unyk, A. M., Trehub, S. E., Trainor, L. J., & Schellenberg, E. G. (1992). Lullabies and simplicity: A cross-cultural perspective. *Psychology of Music*, 20(1), 15-28.

von Koss Torkildsen, J., Hansen, H. F., Svangstu, J. M., Smith, L., Simonsen, H. G., Moen, I., & Lindgren, M. (2009). Brain dynamics of word familiarization in 20-month-olds: Effects of productive vocabulary size. *Brain and Language*, 108(2), 73-88.

Zhao, T.C., & Kuhl, P.K. (2016). Musical intervention enhances infants' neural processing of temporal structure in music and speech. *Proceedings of the National Academy of Sciences*, 113(19), 5212-5217.