Infants’ perception of sound patterns in oral language play

Laura E. Hahn

International Max Planck Research School (IMPRS) for Language Sciences
The educational component of the doctoral training was provided by the International Max Planck Research School (IMPRS) for Language Sciences. The graduate school is a joint initiative between the Max Planck Institute for Psycholinguistics and two partner institutes at Radboud University – the Centre for Language Studies, and the Donders Institute for Brain, Cognition and Behaviour. The IMPRS curriculum, which is funded by the Max Planck Society for the Advancement of Science, ensures that each member receives interdisciplinary training in the language sciences and develops a well-rounded skill set in preparation for fulfilling careers in academia and beyond. More information can be found at www.mpi.nl/imprs.

The MPI series in Psycholinguistics
Initiated in 1997, the MPI series in Psycholinguistics contains doctoral theses produced at the Max Planck Institute for Psycholinguistics. Since 2013, it includes theses produced by members of the IMPRS for Language Sciences. The current listing is available at www.mpi.nl/mpi-series.

Baby & Child
Research Center

ISBN: 978-94-92910-32-5
Cover design by Linda Weiß
layout by Publiss | www.publiss.nl
Printed and bound by Ridderprint, www.ridderprint.nl

The research reported in this thesis was conducted at the Baby and Child Research Centre and the Centre for Language Studies of Radboud University, in Nijmegen, the Netherlands

Infants' perception of sound patterns in oral language play

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

dinsdag 1 februari 2022
om 16.30 uur precies

door

Laura Elisabeth Hahn
geboren op 9 februari 1988
te Rüdersdorf, Duitsland

Promotor:

prof. dr. J.P.M. Fikkert

Copromotoren:

dr. T. Benders (Macquarie University, Australië)
dr. T.M. Snijders

Manuscriptcommissie:

prof. dr. J.M. McQueen
prof. C.F. Rowland
prof. dr. M. van Oostendorp
prof. dr. N. Boll-Avetisyan (Universität Potsdam, Duitsland)
dr. C.M.M. Junge (Universiteit Utrecht)

Contents

1

Chapter 1

General introduction

> “Children need art and stories and poems and music as much as they need love and food and fresh air and play.”
>
> Philipp Pulmann

Parents and caregivers all over the world make use of language play in the form of songs, nursery rhymes, bouncing games and finger plays (Ilari, 2005; Ilari, Moura, & Bourscheidt, 2011; Stern, 1974). These intuitive and ritualized types of poetic verbal behaviour contribute to social cohesion and attachment (Dissanayake, 2000; Markova, 2018), as well as emotion regulation (Cirelli, Trehub, & Trainor, 2018; Corbeil, Trehub, & Peretz, 2016; Mehr, Song, & Spelke, 2016; Trehub, Ghazban, & Corbeil, 2015). With regards to infant-caregiver interaction, poetic language play is ever-present as well (Markova, 2018; Stern, 1974; Trehub, 2019): infants are exposed to singing and spoken nursery rhymes from their caregivers, e.g., during diaper change or bath time, prior to meals or before being put to bed (Ilari, 2005; Mehr, 2014; Trehub et al., 1997).

Language play with infants serves specific communicative functions. For example, more vivid and dynamic renditions of songs (so called *playsongs*) arouse infants, while *lullabies*, sung rather quietly, with low pitch and slow tempo, calm them down (Cirelli, Jurewicz, & Trehub, 2019). Lullabies in particular are recognized cross-culturally and represent a specific song genre with a clear form-function relationship (Mehr, Singh, York, Glowacki, & Krasnow, 2018). Infants certainly are captivated by language play. Slow and rhythmic songs and spoken nursery rhymes, typically produced with a smiling and loving tone of voice, attract long gazes by infant listeners (Leong, Byrne, et al., 2017; Tsang, Falk, & Hessel, 2017) and delay infant distress (Corbeil et al., 2016; Trehub et al., 2015). They also alter cortisol (Shenfield et al., 2003) and oxytocin levels (Markova et al., 2018) in both caregiver and infant-listeners, hormones typically associated with arousal and social engagement.

This dissertation extends the previous research focus on the social-emotional effect of oral language play between infants and caregivers to include a linguistic perspective. The main research question was whether infants perceive the prosodic-phonological sound patterns provided in oral language play. These processing abilities in oral language play were then associated with concurrent and later vocabulary size. If found, such associations could imply that infants’ processing of language play might support everyday speech processing and acquisition.

Detecting sound patterns in language play and its relationship with language development

Why should poetic language play support first language acquisition of young infants? Poetry and language are naturally tightly connected (Jakobson, 1960). “The linguistic sames which are potentially relevant in poetry are just those which are potentially relevant in grammar. [...] The intrinsic structure of language, the raw material of poetry, is carried over into poetry.” (Kiparsky, 1973, p. 235). More specifically, phonological rules and constraints of natural language cannot be disregarded in poetry. Fabb (2010) proposed the *Development Hypothesis*: “the form of literary language and the rules and constraints which hold of it are developments of the form, rules, and constraints of ordinary language.” (Fabb, 2010). Yet, so far, research into the role of poetic oral language play in infants’ acquisition of their native language is sparse. This is in stark contrast with second language teaching, where songs and spoken nursery rhymes are common and extensively studied with respect to their role in vocabulary acquisition (Bebout & Belke, 2017; Cook, 1997; Davis, 2017; Engh, 2013; Ma, Fiveash, Margulis, Behrend, & Thompson, 2020).

Oral language play is a clearly structured type of auditory stimulus, which could support infants’ perception of linguistic patterns therein. Verse lines and melodic phrases provide a scaffold for words and rhymes, while pauses mark boundaries between successive lines of each verse. Caregivers highlight this hierarchical structure in their renditions of language play using a multitude of kinetic, visual and auditory cues (see https://esforum.de/sfr10/trehub.html, section “chapter 18” by Trehub (2013) for an example video). Especially the rhythmic and phrasal structure of songs and nursery rhymes is made salient in infant-directed renditions of oral language play. For example, caregivers reliably differentiate between major and minor beats in their singing and produce longer notes and pauses at song phrase boundaries (Longhi, 2009; Nakata & Trehub, 2011), while being more variable in the dynamics of their expression (Nakata & Trehub, 2011). Caregivers also make use of soft rhythmic bouncing or head-nodding to highlight rhythmic structure during singing (Cirelli et al., 2018). During renditions of rhyming poems and stories to preschoolers, they also alter their speech intensity and word duration to express variations in metrical structure and rhyme predictability (Breen, 2018; Fitzroy & Breen, 2019). This salient rhythmic scaffold of oral language play might provide infants with a means to track linguistic regularities embedded within songs and nursery rhymes, the research question of this dissertation.

The current dissertation aims to answer the question of whether infants perceive sound patterns in oral language play, namely rhymes, rhythm and phrase boundaries. Recognizing the rhyme-relationship between words is a predictor for the development of reading and writing. Here we entertain the idea that oral language play during infancy already allows for a first encounter with rhymes, which might benefit infants in structuring their mental lexicon and in turn enable later explicit rhyme judgments. Phrasal and rhythmic cues in speech provide infants with a hint towards the underlying linguistic structure. For example, intonational phrases indicate which words form a syntactic group and rhythm allows to differentiate stressed from unstressed syllables, a reliable cue to word boundaries. Potentially, the sound patterns of rhythm, rhymes and phrase boundaries are also recognized in oral language play. This processing ability would provide infants with an additional source of linguistic learning alongside the speech input they receive outside of playful contexts.

Secondary research questions pertain to i) a processing benefit for language play compared to everyday infant-directed speech. The clear acoustic structure that is provided in oral language play might make recognition of sound patterns therein easier than in everyday speech and previous studies indeed suggest linguistic processing might be boosted in such an enriched context. Another research question ii) refers to a relationship between phonological processing abilities for language play and vocabulary size. Recently, a number of studies suggested a promising link between informal caregiver language play and infants' and toddlers' language development (Franco, Suttora, Spinelli, Kozar, & Fasolo, 2021; Politimou, 2018; Suppanen, Huotilainen, & Ylinen, 2019; Virtala & Partanen, 2018). For example, infants who received more sung input during their first months of life have larger vocabularies at 14 months old (Franco et al., 2021). The current dissertation extends this research, by linking gains in lexical development (vocabulary size) with the detection of specific phonological/prosodic patterns in oral language play, namely rhyme and rhythm.

Sound patterns in oral language play

Oral language play is a subcomponent of poetry (Jakobson & Waugh, 1979). For the present dissertation it could be summarized as "children's lore, used by and for children" (Opie & Opie, 1959). Just like other poetic genres, oral language play contains prototypic stylistic features which determine their acoustic shape (Burling, 1966; Opie & Opie, 1959; Rubin, 1995; Sullivan, 1999). It is structured into lines and verses with a regular number of syllables, which carry a simple rhythm and often rhyme. In its sung form, arch shaped melodic contours are dominant (Huron, 2006). The prototypic structure of songs and nursery rhymes probably eases their memorization (Rubin, 1995) and production (Tierney, Russo, & Patel, 2011), which in turn supports the transmission of oral language play from one generation of children and caregivers to the next (DeCastro-Arrazola & Kirby, 2018; Rubin, 1995).

The present dissertation concerns infants' first language acquisition. Consequently, the poetic genres of interest here are the ones typically used with preverbal infants, namely child songs and spoken nursery rhymes (Ilari, 2005; Markova, 2018). A **song** can be understood as a text set onto a tune. It consists of melodies and rhythms sung to a steady beat (Lomax, 1968; Savage, Brown, Sakai, & Currie, 2015) and can be found universally, across the world's cultures (Mehr et al., 2019, 2018). The merge of lyrics and melody in songs happens in a non-arbitrary way. For example, there is a common tendency in Western oral traditions to place stressed syllables at stressed metrical positions in the melody (DeCastro-Arrazola, Kranenburg, & Janssen, 2015; Dell & Halle, 2009). The term **nursery rhyme** is often used interchangeably with child songs (Oxford Learners Dictionairy, 2021). This dissertation will only concern spoken nursery rhymes and we will refer to their sung versions as (child) songs. Nursery rhymes have the same stylistic features as songs, with verse structure, rhyme, regular meter, but are not bound to a specific melody (Lerdahl, 2001). The sound patterns inherent to nursery rhymes and songs that are of relevance for this dissertation are rhyme, rhythm, meter, melodic phrases and their corresponding poetic unit verse lines (see Figure 1.1 for all). Each of these sound patterns will now be discussed in more detail.

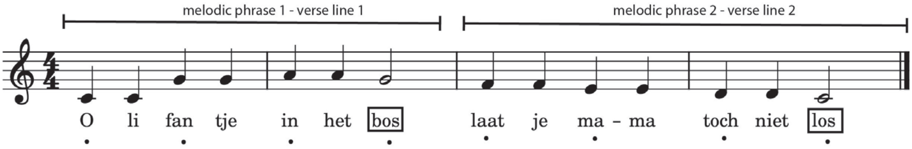


Figure 1.1 Example of the beginning of a proto-typical child song. Stressed beats and syllables indicated with black dots (under the text); line-final rhyme highlighted with black rectangle; phrase/verse structure indicated with horizontal brackets on top of the musical notes.

Verse lines are a rather large structural unit in poetry (indicated with horizontal lines in Figure 1.1). Being a subcomponent of each verse (Fabb, 2015), they can be considered carrier chunks for smaller poetic units like words, rhymes and alliteration. Verse lines are imposed on linguistic or melodic phrase structure (B. Hayes & Kaun, 1996; Lehrdahl & Jackendoff, 1985), which in the example in Figure 1.1 are aligned. The phrase structure of oral language play thus often resembles syntactic structure (see further chapter 2). For infants, sensitivity to the units of verse lines and melodic phrases might be a steppingstone into parsing of their native language. This possibility is further explored in chapter 2 of this dissertation, where we provide evidence for 6-month-olds' segmentation of infant-directed song and speech into phrases.

Rhythm can be considered a "series of temporal intervals" with no inherent structure (Hannon & Johnson, 2005). In the example song in Figure 1.1, rhythm is conveyed by the length of each musical note, i.e., 6 quarter-notes and 1 longer half-note at the end of each melodic phrase. **Meter,** in turn, denotes an abstract temporal structure, with multiple hierarchical layers, that listeners impose on a rhythm. In poetry, meter refers to the pattern of a fixed number of stressed and unstressed syllables laid on top of each verse line (Arleo, 2006; Burling, 1966; Kelly & Rubin, 1988; Selkirk, 1980). Stressed and unstressed syllables together form prosodic feet, a repeating rhythmic unit (B. Hayes, 2010; Nespor & Vogel, 2007). In a song or melody, one salient level of meter is represented by the beat, i.e. the note listeners would tap their fingers to (Hannon, Nave-Blodgett, & Nave, 2018). In the example in Figure 1.1 poetic and melodic meter are aligned, meaning that stressed syllables in the lyrics coincide with prominent beats in the melody. This parallelism is indicated by dots beneath the lyrics. Meter in poetry tightly interacts with prosodic-phonological features and phrasing of natural languages (Hanson & Kiparsky, 1996; B. Hayes, 1989, 2010; Kelly & Rubin, 1988). By recognizing the metrical structure in nursery rhymes and songs, infants might thus get a hint onto the rhythmic-prosodic structure of their native language, for example with regards to stress-based word segmentation (Snijders et al., 2020). In chapters 5 and 6 metrical rhyming verse (spoken nursery rhymes) will be used to investigate whether adult and infant listeners benefit from a regular meter in their rhyme perception and their neural alignment with the rhythm of speech.

Rhyme refers to phonological identity of substrings of verse lines (Peust, 2014; Stallworthy, 1996). The recurring phonological material occurs mostly line-final in child songs and nursery rhymes, although it can also be found line-initially or line-internally. This dissertation will only concern line-final rhyme, as this type is most prevalent in West-Germanic languages (Peust, 2014). In Figure 1.1 a typical line-final rhyme is indicated in the black rectangle, where *bos* (forest) rhymes with *los* (let loose). More specifically, the combination of syllable nucleus and coda, -*os*, is repeated across verse lines. This combination of syllable nucleus and coda in phonological theory is commonly referred to as the *rime* unit (Treiman, 1983). Here we will avoid the term *rime* where possible to avoid confusion with *rhyme*. The meta-linguistic ability to segment syllable onsets from rimes, thus *b* and *l* from the rime -*os* in the rhyme example in Figure 1.1, is of profound importance for the development of literacy (Ziegler & Goswami, 2005). Rhyme awareness enables children to reflect on the phonological structure of words, which in turn supports their reading and spelling. The ability to recognize rhyming words develops during preschool years (Vloedgraven & Verhoeven, 2007) and previous research considered oral language play during early childhood as an informal contributor to the early precursors of rhyme awareness (Bryant, Bradley, MacLean, & Crossland, 1989; Dunst, Meter, & Hamby, 2011; MacLean, Bryant, & Bradley, 1987). In the current dissertation, we apply this hypothesis to even younger children, namely preverbal infants. In chapters 3, 4, 5 and 6 infants' (and adults') implicit perception of rhyme in songs and nursery rhymes is studied.

Together, rhyme, rhythm, meter and verse lines let listeners perceive poetry as globally homogeneous auditory objects, resulting in a perceptual prevalence of phonological/prosodic form over semantic content (Jakobson & Waugh, 1979; Menninghaus et al., 2015a). The focus on form over semantic content renders oral language play ubiquitous in interactions with pre-verbal infants. Crucially, oral language play in a given language amplifies phonological and typological tendencies of the respective language, but also resembles crosslinguistic universals (Kiparsky 1973; Fabb 2010). For infants, there is thus potentially a lot to learn from oral language play, about language in general and about their native language in particular. Yet, our knowledge about whether and how infants process linguistic information in language play is sparse. It also remains unclear whether phonological and prosodic processing of language play is indicative of concurrent and later linguistic abilities. This dissertation reports evidence for infants' perception of phrasal (verse line) structure in song and speech (chapter 2) and rhythm and rhyme perception in nursery rhymes and songs (chapters 3, 4 and 6). The latter results are extended with data from an adult study (chapter 5), where listeners passively listened to nursery rhyme stimuli, to assess whether regular meter influences adults' implicit rhyme perception in poetry. In addition to these experimental results, we relate our findings to concurrent and future vocabulary size, to elucidate a potential contribution of oral language play to infants' language development.

Methodological considerations

Stimuli

Throughout the studies reported here, naturalistic song and nursery rhyme stimuli were employed. This strengthens the ecological validity of our findings and extends previous research, which used rather simplistic and impoverished song stimuli (see introductions to chapters 2, 3 and 4). In chapters 2, 3 and 4 song stimuli were used. The songs used in chapter 2 were created in parallel to the speech stimuli from a previous study (Johnson & Seidl, 2008). They resemble typical child songs in their melodic structure, with arch-shaped melodic contours and salient pauses between successive melodic phrases. Chapters 3 and 4 feature a collection of novel songs with lyrics in nonsense Dutch made of phonologically legal pseudowords. The usage of nonsense Dutch facilitated song creation and granted more control over, for example, lexical frequency of syllable rimes used within the critical words in the song lyrics (i.e., *rhyming* versus *non-rhyming* line-final pseudowords, see respective chapters for detail). In chapters 5 and 6 spoken nursery rhymes were used. The switch to another genre of language play was motivated by previous research on prosodic-phonological cues present in rhythmic infant-directed speech (Leong & Goswami, 2015) and the interaction of rhyme and meter in adults' perception of poetry (Obermeier et al., 2016). Our nursery rhyme stimuli had a typical 4-line structure, regular iambic meter and monosyllabic rhyming words at the end of each verse line (Burling, 1966). Since the stimuli were used with both adult (chapter 5) and infant participants (chapter 6), regular Dutch lyrics were created for each nursery rhyme, instead of the nonsense-Dutch used in some of the song stimuli. Nevertheless, the critical target words at the end of each verse line were pseudowords, again to facilitate stimulus creation.

Age groups studied

Chapters 2, 3, 4 and 6 are based on data from monolingual Dutch infants. In chapters 2 and 3, 6- and 9-month-old infants were tested in accordance with previous research assessing the same capacity in spoken language (Johnson & Seidl, 2008; Jusczyk, Goodman, & Baumann, 1999). For chapters 4 and 6, slightly older infants were tested, 10.5-month-olds. This age group was sampled in an attempt to gain more robust effects than in a previous study (chapter 3) and again in accordance with previous research testing similar processing abilities in spoken language (Junge, Cutler, & Hagoort, 2014; Junge, Kooijman, Hagoort, & Cutler, 2012; Kooijman, Hagoort, & Cutler, 2005; Kooijman, Junge, Johnson, Hagoort, & Cutler, 2013). Chapter 5 features data from an adult study, while using the exact same nursery rhyme stimuli as in the infant study reported in chapter 6. This allows for a comparison between mature and developing neurophysiological correlates of rhyme and rhythm processing.

Methods

The current set of studies resembles a combination of behavioural and neurophysiological methods supplemented with parental questionnaires about infants' vocabulary size. Chapters 2 and 3 make use of the Headturn Preference Procedure (HPP), which measures infants' listening preferences to two types of acoustic stimuli. This method requires an explicit response (headturn) by the infant listener, as it measures looking times for different types of stimuli each presented from one side of a testing booth. Developing processing abilities however, are not always visible in overt behavior such has head-turns. Moreover, the underlying time course of stimulus processing remains oblique (Kooijman, Johnson, & Cutler, 2008). Therefore, the subsequent studies employed electroencephalography

(EEG). This passive measure requires no kinetic response by the infant and allows for more detailed investigations into the time course of auditory perception. In chapters 4, 5 and 6, event-related-potentials (*ERPs*) for rhyming versus non-rhyming pseudowords in songs and nursery rhymes were collected from infants (chapters 4 and 6) and adults (chapter 5) and previous ERP studies on infant word segmentation were used as roadmap for the design of these studies (Junge et al., 2014, 2012; Kooijman et al., 2005, 2013). Two types of rhyme ERPs were investigated: *rhyme sensitivity* and *rhyme repetition*. Rhyme sensitivity was measured at the first point of diversion between rhyming and non-rhyming songs/ nursery rhymes, i.e., where the rhyme is repeated for the first time in rhyming but not in non-rhyming stimuli. The rhyme repetition effect was measured as the averaged response to rhymes occurring at the end of the remaining phrases of the song or nursery rhyme. The rhyme repetition measure rendered a more reliable estimate of rhyme perception (see chapter 3 for detail), and was therefore correlated with infant vocabulary size, again following previous research (Junge & Cutler, 2014; Junge et al., 2012).

In addition to ERPs, an EEG measure of neural synchrony was analyzed as well. Recent advances in cognitive neuroscience indicate that speech rhythm is mirrored in neural activity, as phonological units such as phrases, stress, syllables and phonemes, each occurring at different time scales within the acoustic stimulus, are reflected in neural activity in specific oscillatory frequency bands (Giraud & Poeppel, 2012; Obleser & Kayser, 2019; Peelle & Davis, 2012). This relationship between speech and brain rhythms might be particularly strong for rhythmic infant-directed speech (Leong & Goswami, 2015), the type of stimulus nursery rhymes provide. In chapter 6 infant speech-brain coherence (*SBC*) was employed as an EEG index for synchrony between infants' brain oscillations and the speech amplitude envelope. Specifically, SBC was measured for the stress and syllable rhythm of spoken nursery rhymes to infer whether these rhythms are reflected in infants' cortical activity. Infants' cortical tracking of the stress- and syllable rhythm was then again related to their vocabulary size. Assessing coherence between speech and infants' neural activity resembles a relatively novel measure and few previous studies report on rhythmic cortical tracking in infants (Choi, Batterink, Black, Paller, & Werker, 2020; Jessen, Fiedler, Münte, & Obleser, 2019; Kalashnikova, Peter, Di Liberto, Lalor, & Burnham, 2018; Lang, Ott, Giudice, & Schabus, 2020; Menn et al., 2020; Ortiz Barajas, Guevara, & Gervain, 2021; Snijders, 2020).

Estimates of infants' vocabulary size were obtained from parental questionnaires and were linked to infants' perception of rhyme and rhythm in language play. Previous research has established a strong link between early phonological and prosodic processing abilities and later general linguistic development (for review see: Cristia, Seidl, Junge, Soderstrom, & Hagoort, 2014; Wang, Seidl, & Cristia, 2021). So far, such a relationship has been solely established for processing of (ID)-speech. The current dissertation extends this research base to another domain of infant-caregiver communication, namely oral language play.

Research questions and dissertation outline

Chapter 2 reports the results of a HPP study investigating **whether infants segment songs into phrasal units.** The study is a replication and extension of previous studies (Johnson & Seidl, 2008; Nazzi, Kemler Nelson, Jusczyk, & Jusczyk, 2000), which attested phrasal segmentation in the speech domain. **Chapter 3** asks **whether infants are sensitive to rhymes in child songs**, again using the HPP. In **chapter 4**, infants' early rhyme sensitivity will be studied again, now in a passive listening EEG paradigm using the same song stimuli as in chapter 3. Individual differences in rhyme perception will be related to concurrent and future vocabulary size in order to test for an early relationship between the precursors of phonological awareness (as indexed by the rhyme ERPs) and lexical development. **Chapter 5** investigates **whether regular meter influences adults' implicit rhyme perception in poetry** (nursery rhymes) using a passive-listening EEG paradigm. **Chapter 6** features the same nursery rhyme stimuli used in chapter 5, to ask **whether infants' oscillatory activity follows the stress rhythm of nursery rhymes**. Moreover, implicit rhyme perception in poetry reported in the previous chapters will be replicated. In **chapter 7** the main findings of this dissertation are summarized and discussed.

2

Chapter 2

Six-month-old infants recognize phrases in song and speech

This chapter is a slightly modified version of: Hahn L. E., Benders T., Snijders T. M., Fikkert P., 2020. Six-month-old infants recognize phrases in song and speech. *Infancy*.25(5):699-718.

Data, materials and analysis scripts for this study are available at:
https://osf.io/4zvad

Abstract

Infants exploit acoustic boundaries to perceptually organize phrases in speech. This prosodic parsing ability is well-attested (Nazzi et al., 2000; Johnson & Seidl, 2008) and is a cornerstone to the development of speech perception and grammar. However, infants also receive linguistic input in child songs. This study provides evidence that infants parse songs into meaningful phrasal units and replicates previous research for speech. Six-month-old Dutch infants (n = 80) were tested in the song or speech modality in the Headturn Preference Procedure. First, infants were familiarized to two versions of the same word sequence: one version represented a well-formed unit, the other contained a phrase boundary halfway through. At test, infants were presented two passages, each containing one version of the familiarized sequence. The results for speech replicated the previously observed preference for the passage containing the well-formed sequence, but only in a more fine-grained analysis. The preference for well-formed phrases was also observed in the song modality, indicating that infants recognize phrase structure in song. There were acoustic differences between stimuli of the current and previous studies, suggesting that infants are flexible in their processing of boundary cues while also providing a possible explanation for differences in effect sizes.

Acknowledgements

We thank Lisa Rommers, Maaike ten Buuren, Katharina Menn and Mareike Geiger for their help in scheduling and testing the infants, Katharina Menn and Emma van Kampen for their reliability coding and Annelies van Wijngaarden for being our speaker and singer for the stimuli. The manuscript also benefited from the critical comments of three anonymous reviewers.

Introduction

Across the globe, caregivers sing for their infants (Trehub & Trainor, 1998). Infant-directed (ID) song has a universal acoustic shape (Mehr et al., 2019) and is a distinct communicative modality that is recognized across cultures (Mehr et al., 2019; Trehub, Unyk, & Trainor, 1993a). Young children are exposed to songs in a highly ritualized and repetitive fashion (Bergeson & Trehub, 2002; Custodero & Johnson-Green, 2008; Custodero, Rebello Britto, & Brooks-Gunn, 2003; Ilari, 2005), and ID-song clearly serves social-emotional functions: to soothe or stimulate the infant and to strengthen the bond between caregiver and infant (Cirelli et al., 2019; Corbeil et al., 2016; Shenfield et al., 2003).

Yet, songs often contain language in the song lyrics and thereby entail the possibility of language learning for the infant listener. Previous research provides evidence that infants are sensitive to small phonological units like syllables and words in song lyrics: for example, already before their first birthday infants recognize changes in the syllable order in songs (François et al., 2017; Lebedeva & Kuhl, 2010; Suppanen et al., 2019; Thiessen & Saffran, 2009), differentiate between rhyming and non-rhyming songs (chapter 3), and learn novel words recurring in the song lyrics (Snijders, Benders, & Fikkert, 2020).

Language learning, however, entails more than the recognition of single syllables or words. Infants also need to establish hierarchical relationships between smaller units of an utterance. One way for infants to deduce syntactic structure from the input is to tune into phrasal prosody (the melody and rhythm of speech), as boundaries between prosodic constituents typically overlap with boundaries between syntactic constituents (even though the reverse is not always true) (Nespor & Vogel, 2007; Shattuck-Hufnagel & Turk, 1996). Speakers signal prosodic boundaries by altering prosodic cues like pitch, duration and pauses at constituent edges (Wagner & Watson, 2010). Recognizing these prosodic cues in their input aids infants in determining the edges of syntactic constituents (e.g., de Carvalho, Dautriche, & Christophe, 2017; Hawthorne & Gerken, 2014) and stimulates their morpho-syntactic development (Morgan & Demuth, 1996). The current study asks whether infants also exploit the melodic phrase structure of ID-songs to perceptually organize the linguistic input in songs.

Infants' recognition of phrase structure in speech and song

Over the course of the first year of life, infants develop sensitivity to the instantiation of prosodic boundaries in their native language (Johnson & Seidl, 2008; Wellmann, Holzgrefe, Truckenbrodt, Wartenburger, & Höhle, 2012). For infants, just like for adults, the prosodic packaging of speech provides a perceptual filter (Jusczyk et al., 1992) that eases recognition, segmentation and memorization of linguistic elements (Frazier, Carlson, & Clifton, 2006; Hochmann, Langus, & Mehler, 2016; Johnson, Seidl, & Tyler, 2014; Mandel, Jusczyk, & Kemler Nelson, 1994; Seidl & Johnson, 2006; Shukla, White, & Aslin, 2011).

To date it is not known whether infants also recognize phrasal units of caregiver singing. However, many acoustic cues to phrase structure are the same in melodies and speech (Deutsch & Feroe, 1981; Heffner & Slevc, 2015; Lehrdahl & Jackendoff, 1985; Riemann, 1912; Trainor & Adams, 2000) and prosodic phrase segmentation is not bound to the listeners' native language or spoken modality: English-speaking adults do segment words from unfamiliar languages if these words are placed at prosodic phrase boundaries (Endress & Hauser, 2010; Langus, Marchetto, Bion, & Nespor, 2012) and American-English infants can exploit the prosody of non-native languages, e.g., Japanese (Hawthorne, Mazuka, & Gerken, 2015), Polish (Jusczyk, 2003) and even American Sign Language (Brentari, González, Seidl, & Wilbur, 2010), to recognize phrases. This combination of observations informs the hypothesis that infants might also be able to perceptually organize songs into phrases.

Various aspects of ID-singing may be particularly beneficial for infants’ recognition of phrasal song structure: ID-songs have a predictable canonical form (Mehr et al., 2019) and are produced at a rather slow tempo, with multiple visual and auditory cues to phrasal boundaries (Delavenne, Gratier, & Devouche, 2013; Falk & Kello, 2017; Leong & Goswami, 2015; Longhi, 2009), and with particularly salient acoustic boundary cues (Falk & Kello, 2017; Trainor, Clark, Huntley, & Adams, 1997). Songs thus provide infants with ample time to process an acoustic stimulus that is rich in structural cues. Moreover, songs grab infants' attention at least as effectively as ID-speech (Corbeil, Trehub, & Peretz, 2013; Costa-Giomi, 2014), and possibly more (Nakata & Trehub, 2004; Tsang et al., 2017).

The arguments provided so far all suggest that infants might be able to parse phrasal units from ID-song. Yet, there are also differences in the acoustic instantiation of boundary cues between song and speech (see e.g., references in Merrill et al., 2012). For infants to recognize phrase structure in songs, they thus need flexible representations of phrase boundaries which adjust to the song modality. So far, the available literature does not provide conclusive evidence for this flexibility.

Investigating infants' ability to recognize phrase structure in songs is also relevant in light of recent evidence that the recognition of phrasal structure in linguistic or musical play is related to grammar development in typically developing preschoolers (Politimou, Dalla Bella, Farrugia, & Franco, 2019) and children with developmental language disorder (Richards & Goswami, 2019). As prosodic parsing is a pre-cursor to syntactic development (Morgan & Demuth, 1996), these studies raise the possibility that caregivers' language play, including ID-singing, contributes to the development of prosodic parsing. However, the work suggesting a relationship between phrase perception and grammar development focused on (pre)school children (Politimou et al., 2019; Richards & Goswami, 2019), whereas prosodic parsing already develops within the first year of life (Carvalho, Dautriche, Millotte, & Christophe, 2018), and a direct test of children's ability to segment phrases from songs or other forms of language play is still poignantly lacking. The current study thus aims to provide evidence for infants' recognition of the phrasal building blocks of ID-songs.

Infants' recognition of phrase structure in music

Currently, literature on infants' perception of melodic phrase structure in music and song is sparse. In a seminal study by Jusczyk and Krumhansl (1990), 6-month-olds differentiated between excerpts from Mozart Minuets with pauses at natural (phrase boundary) positions and excerpts with unnatural pauses (within phrases). A follow-up study with American-English infants extended this finding to melodies from non-western (Japanese) child songs (Jusczyk, 2003), indicating that melodic phrase structure perception does not require extensive experience with a musical tradition. Crucially though, none of these previous studies required infants to encode and process melodic phrase structure. Instead, infants were provided with a pre-segmented stimulus that was reminiscent of their daily musical experience (naturally segmented) or rather odd (unnaturally segmented). It thus remains unclear whether infants chunk native songs into meaningful units and recognize subcomponents of the songs, despite these being the type of musical stimulus infants are exposed to on a daily basis. Recent evidence from Dutch infants (chapter 3 of this thesis) also only indirectly supports the notion of song structure being accessible: 9-month-olds differentiated rhyming (and thus more natural) songs from non-rhyming (and thus less natural) songs, but the study did not test whether this differentiation has implications for the processing of the linguistic content. In the current study, we will provide infants only with natural native child songs and will explicitly test their ability to recognize familiarized song phrases.

Extending the prosodic parsing paradigm

A good starting point for an investigation of infants' encryption of the inherent structure of ID-song is transferring a reliable paradigm from infant speech perception research to the song modality while also replicating previous research for ID-speech. Such a paradigm was provided by Nazzi and colleagues (2000), showing that 6-month-olds used prosodic phrase structure to segment clauses from continuous speech. In this Headturn Preference study, infants were familiarized to two versions of the word sequence *leafy vegetables taste so good* (Nazzi et al, 2000, Experiment 1). One version of the sequence was prosodically well-formed, carrying phrase boundaries at the edges, and sounded like a coherent clause: *[Leafy vegetables taste so good].* The other version of the word-sequence contained a phrase boundary halfway through, sounding more like snippets of two adjacent clauses: *leafy vegetables] [Taste so good*. In the subsequent test phase, infants heard two spoken passages of three sentences each. The well-formed sequence from the familiarization phase reoccurred as a coherent clause of one passage. The ill-formed sequence from the familiarization phase reoccurred as a subcomponent of two adjacent clauses of the other passage. Infants listened longer to the passage containing the well-formed compared to the ill-formed word sequence, indicating that they capitalized on the prosodic structure of the passage to recognize the familiarized well-formed word sequence therein. This paradigm has been adopted in numerous subsequent studies (Seidl, 2007; Seidl & Cristia, 2008; Soderstrom, Kemler Nelson, & Jusczyk, 2005). Critically for the present study, Dutch 6-month-olds also showed the same preference for the passage containing the well-formed word sequence (Johnson & Seidl, 2008).

Whether infants' prosodic parsing ability extends to the musical modality has already been explored in two short reports (Hawthorne & Gerken, 2013; Nazzi et al., 2000, see discussion at the end of this chapter). Both studies applied the paradigm described above to melodies from a musical instrument. The preliminary results suggest that infants recognized the familiarized well-formed tone-sequence within a longer musical piece.

The current study

The current study investigates infants' recognition of the phrasal building blocks of ID-song and replicates earlier studies on infants' recognition of phrases in ID-speech. We will use the paradigm described above that has successfully revealed infants' phrase segmentation of ID-speech (for Dutch: Johnson & Seidl, 2008; the original study for English: Nazzi et al., 2000) with a new sample of Dutch 6-month-olds and a new version of the Dutch stimuli and extend the paradigm to ID-song, using natural song material that matches the ID-speech stimuli in content and syntactic structure. Our approach significantly extends previous work in two ways: First, infants' processing of song lyrics has so far been limited to smaller phonological and lexical building blocks (François et al., 2017; Hahn et al., 2018 (chapter 3); Lebedeva & Kuhl, 2010; Snijders et al., 2020; Suppanen et al., 2019;

Thiessen & Saffran, 2009). We will extend the scope of this research to phrases, cognitive units which are relevant not only for the perception of song structure but also for lexical and syntactic development in infants' native language. Secondly, we will build upon the previous work on infants' auditory grouping in polyphonic instrumental music (Jusczyk & Krumhansl, 1993; Krumhansl & Jusczyk, 1990), monophonic melodies (Nazzi et al., 2000; Hawthorne & Gerken, 2013) and non-native child songs (Jusczyk, 2003), employing the type of musical stimulus that possibly best represents infants' musical input (Volkova, Trehub, & Schellenberg, 2006), namely native child songs. By extending the paradigm of Nazzi and colleagues (2000) we will also move beyond mere preferences for naturally phrased melodies. Instead, our study requires infants to incrementally process and organize ecologically valid native song input and match this input to memorized song fragments.

Method

Participants

A sample of 95 6-month-old infants (mean age in days: 184, range: 167-209 days, *SD* = 9.02, 53 girls) from monolingual Dutch households was tested of which 12 infants were excluded, because they fussed or cried during the experiment (*n* = 11) or grew up in a bilingual household (*n* = 1). Three more infants were excluded from part of the analysis because they did not contribute trials in both experimental conditions for the critical data set (see Analysis Section), resulting in a final dataset of n = 80 or n = 83 infants depending on the respective analysis. Participants were recruited from the Baby and Child Research Center at Radboud University, Nijmegen, the Netherlands. According to their caregivers, infants were born full-term, had normal hearing, and no familial history of language or reading problems. The present study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each infant before any assessment or data collection. Ethical approval for the study was obtained from the Ethiek Commissie Faculteit der Sociale Wetenschappen (ECSW) at Radboud University in Nijmegen, Netherlands. Caregivers had the choice between 10€ or a book as a reward for their participation. The results of a questionnaire on musical exposure confirmed that all participants were regularly exposed to songs and music from electronic devices and human singers (the results of the questionnaire are summarized in the online materials).

A power analysis using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007), based on Experiment 1 of Johnson and Seidl (2008) (estimated correlation between groups set to 0.5, Cohen's $d_{Experiment\ 1}$ = 0.35) resulted in a required minimum sample of 52 infants to detect the phrase-segmentation effect in each modality (80% power in one-sided *t*-test with α = .05). We thus aimed for usable data of 104 infants in total. Due to time and resource limitations, however, data collection was terminated after 95 participants.

Materials

Materials, design and procedure of the current study closely followed the study by Johnson and Seidl (2008, Experiment 1, henceforth “J&S”). Stimuli were novel spoken and sung recordings of the J&S stimuli, complemented by a second stimulus set. The basis of all stimulus materials was a pair of text passages from J&S, both consisting of three sentences, separated by two phrase boundaries (see Table 2.1, Pair 1).

Table 2.1 Texts from Passage Pair 1 and 2.

	Pair 1 (Melody 1) copied from Johnson and Seidl (2008)	**Pair 2 (Melody 2) created in analogy to pair 1 passage 1**
passage 1	Tante vraagt zich af wat de jongens eten. **Koude pizza smaakt niet zo goed**. Hun opa vindt dat wel erg lekker. *Aunt wonders what the boys are eating.* ***Cold pizza doesn’t taste so good.*** *Their grandpa really likes that.*	Oma krijgt meteen wat de meisjes maken. **Hete koffie wordt heel snel oud.** Hun vriendje heeft toch nog geen beker. *Grandma soon gets what the girls are making.* ***Hot coffee gets old really quickly.*** *Their friend has no cup yet.*
passage 2	Het staat in de oven. **De jongens eten koude pizza.** Smaakt niet zo goed in de vroege ochtend. *It’s (placed) in the oven.* ***The boys are eating cold pizza.*** *(It) doesn’t taste so good in the early morning.*	Ze zijn in de keuken. **De meisjes maken hete koffie.** Wordt heel snel oud, als je hem niet op drinkt. *They are in the kitchen.* ***The girls make hot coffee.*** *Gets old really quickly, if you don’t finish it.*

Note. Phrase-internal sequences in **bold**, straddling sequences underlined. Within one session (song or speech), one infant would hear sequences and passages from one pair of passages. During the familiarization phase of the experiment, an infant would only hear the internal sequence from one passage and the straddling sequence from the other passage within the same pair (crucially, both sequences consist of the same words but differ in phrasal structure). During the test phase infants hear the full passages of each pair. The experimental condition of the test passages (internal or straddling) is determined by the respective sequences heard during the familiarization phase.

Within both passages, the same sequence of words occurred (e.g., *koude pizza smaakt niet zo goed*), but one passage contained the sequence as a single phrase, i.e., *phrase-internal* (e.g., [*koude pizza smaakt niet zo goed*]) and the other passage contained the sequence with a phrase boundary in the middle, i.e., phrase-straddling (e.g., … *koude pizza] [smaakt niet zo goed* …). The two passages were used for the test phase of the experiment. The *phrase-internal* and *phrase-straddling* sequences extracted from the passages were used for familiarization. All stimuli were recorded in a spoken as well as a sung version. Passage pair 1 was based on the Dutch stimuli of J&S, with a slight change to fit the melody[1]. Passage pair 2 was created in analogy to pair 1: The number of syllables, word stress, and phrase structure were identical, and the lyrics had the same assumed familiarity (all content words of both pair 1 and pair 2 appeared in the Dutch N-CDI (Zink & Lejaegere, 2002) and the words had a similar mean log raw frequency of 3.6 (pair 1) and 3.8 (pair 2) in the Dutch Celex corpus (Baayen, Piepenbrock, & Gulikers, 1995). Both passage pairs are provided in Table 2.1.

Song stimuli. Both pairs of passages were set onto the melodies of child songs (Figure 2.1). Passage 1 was set onto melody 1 (“Sea Saw Margery Daw”, originally from England) and passage 2 was set onto melody 2 (“Vine Melcul Suparat”, originally from Romania), with one syllable per musical note and stressed syllables on strong metrical positions within each melody. The position of sentence boundaries in the passages was aligned with the position of melodic phrase boundaries in the melodies.

Three listeners (two amateur – one professional musician, two Dutch – one English native speaker) who were kept naïve to the purpose of the study, judged the quality of the resulting melodies. All three found them to resemble typical child songs.

Recording. The same female Dutch speaker was recorded for the spoken and sung stimuli and was kept naïve to the purpose of the study. Only after the recording it became apparent that the same person’s voice had been recorded for the original J&S stimuli. The singer/speaker was instructed to speak and sing in a lively, child-directed manner while looking at the photo of a toddler from her family. She chose a speaking and singing tempo and a pitch height that were convenient to her. Recording took place in a sound attenuated booth and further processing was done using Praat (5.3.49, (Boersma & Weenink, 2014)) and Audacity (2.1.0): Pauses between phrases were set to silence but kept at their original duration. Two sequences were cut from each passage: one internal and one straddling (see Figure 2.2), resulting in 8 sequences for the sung and 8 sequences for the spoken modality.

---

[1] We slightly modified one phrase: original wording: “Hun zus vindt dat lekker.” ‘*Their sister likes that’.* Novel wording: “Hun opa vindt dat wel erg lekker.” ‘*Their grandpa really likes that.’*

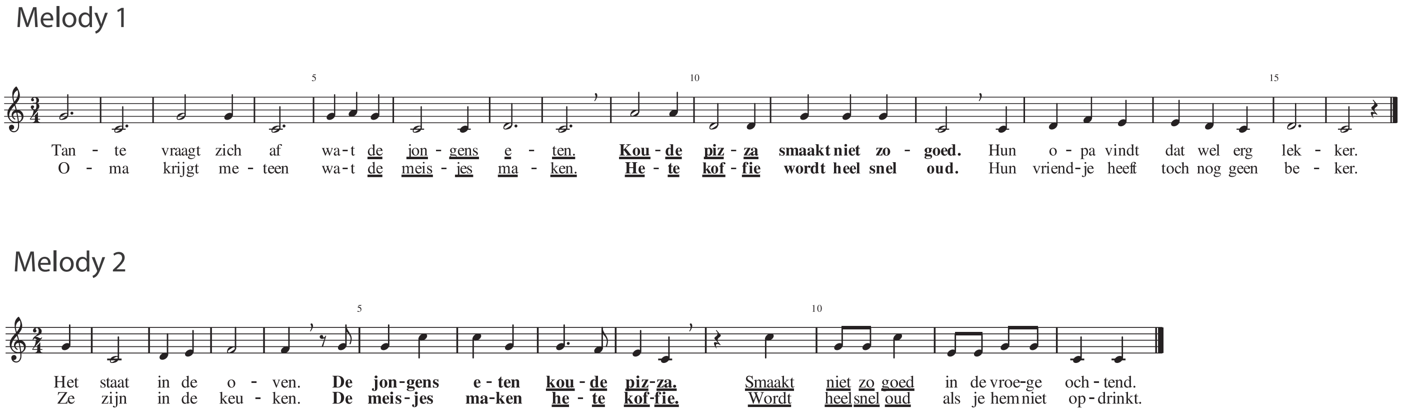
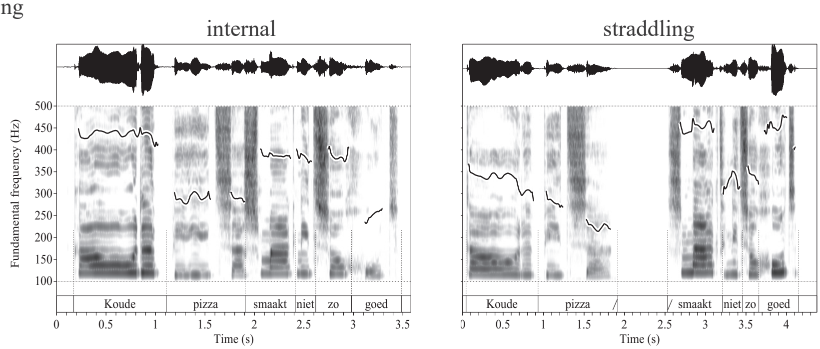


Figure 2.1 Melody 1 and 2 as used for the song stimuli. Phrase-internal sequences in bold, straddling sequences underlined. See the online materials for a sound example.

Song



Speech

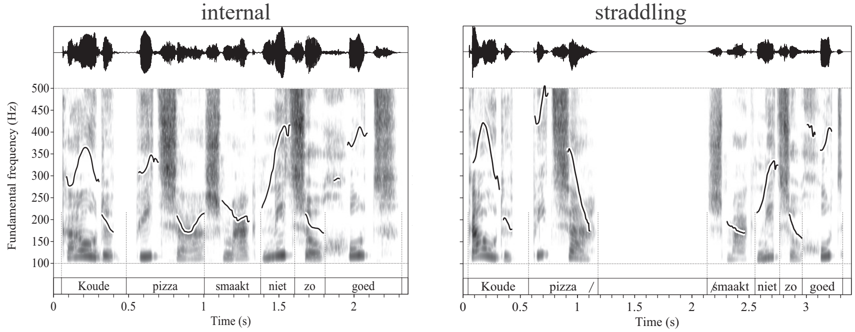


Figure 2.2 Example stimulus. All subfigures denote waveform, sonogram, spectrogram and text-grid annotation and time on the y-axis.

Acoustic analysis. Acoustic measures were obtained around the internal boundary in the straddling sequence (e.g., … pizza] [smaakt …) and compared to the same sequence without a boundary in the phrase-internal sequence (e.g., [… pizza smaakt ….]) using Praat (sound files and corresponding text grids can be found in the online materials).

Comparison of song and speech stimuli within the current study. Phrase boundaries in both song and speech stimuli were expressed by longer pauses and longer pre-boundary vowels at the phrase boundary in the straddling sequence compared to the corresponding internal sequence. In the song stimuli, the pitch rose after the boundary in the straddling sequences. In the spoken sequences the opposite pattern was found: pitch increased at the final vowel and then decreased at the first vowel of the following phrase. The speaker thus used a rising boundary tone to mark the end of her spoken phrases.

Comparison of stimuli from Johnson and Seidl (2008) and the stimuli for the current study. Despite the fact that we used the same words and the same speaker, the speech stimuli of the current study were substantially slower than the stimuli by J&S (see Table 4 in the online materials). The longer pauses between sentences together with the overall slower speech rate of the current stimuli resulted in large differences in onset time of the critical sequence (see Table 5 in the online materials). These differences will be taken into account in the analysis of the looking-time data (see section Mixed-effect model analysis below). Furthermore, the pitch reset at the phrase boundaries of the straddling sequences of the J&S study was less pronounced in the stimuli of the current study (see Table 6 in the online materials), probably due to the slightly different intonation and the rising boundary tone the speaker used for the current study.

Stimulus pre-test. Three Dutch native speakers judged the intelligibility of the sung and spoken sequences and were asked to judge the *straddling/internal* manipulation of the sung and spoken sequences. The three judges were first asked to listen once to each of the 16 sequences and immediately write out the text orthographically, as they understood it. All three participants wrote down the correct texts without mishearing. They were then presented with the *phrase-internal* and *phrase-straddling* versions of every sequence as a pair and were asked to indicate which of the two sequences sounded more coherent. All three participants judged all *phrase-internal* sequences to sound more coherent than their *phrase-straddling* counterparts.

Procedure

The experiment was run using the Headturn Preference Procedure. Three lights were placed within a three-sided booth at infant eye-level: a blue light in the center and red lights on the right and left walls of the booth. A camera was hidden below the center light to observe infant behavior from outside. Stimuli were presented via loudspeakers below the red lights. The infant and caregiver were seated in the middle of the booth, directly opposite the blue center light, exactly in between the left and right red lights. Stimulus presentation was controlled from outside the test booth by the experimenter, using the stimulus presentation software Look! (Meints & Woodford, 2008). The experimenter was blind to trial number and trial condition and coded the looking behavior of the infant (left, right, center) using assigned keys. The same procedure was used for both familiarization trials and test trials. The entire session was video-recorded for offline reliability coding (see section “reliability coding” in the online materials).

Design. Infants took part in either the song or the speech version of the experiment, and were tested on their ability to segment either songs or speech into phrases (effect of modality, between subjects). Following Nazzi's (2000) Headturn Preference Procedure, infants were first familiarized with two sequences of the same words, one uttered as phrase internal, carrying phrase boundaries at the edges, e.g., *[Koude pizza smaakt niet zo goed]* ("Cold pizza doesn't taste so well"), the other uttered as phrase straddling, carrying a phrase boundary halfway, e.g., *koude pizza] [smaakt niet zo goed* ("cold pizza. Doesn't taste so well"). The internal sequence thus represented a well-formed acoustic unit, the straddling sequence was ill-formed. Apart from this acoustic difference, the exact same words were occurring in the sequences used in the familiarization phase. In the test phase, infants were presented with two passages of three sentences each: one passage contained the phrase straddling sequence, the other the phrase internal sequence (Table 2.1). For the analysis, looking times to the passages were assessed. Which passage functioned as the internal and which as the straddling passage was determined by the content of the respective sequence used during familiarization (effect of condition, within-subjects).

Counterbalancing and randomization. The four pairs of passages (Table 2.1) were distributed across eight lists (four lists for speech, four lists for song). Within each list, one pair of passages was used and presentation side of the first stimulus (left, right) was counterbalanced and the same presentation side and the same condition were restricted to occur maximally two times in a row.

Experimental session. Caregivers were first briefed about the experimental procedure and filled out the music exposure survey (see the online materials for an English translation of the questionnaire). At the start of the experiment, infants were seated on their caregiver's lap in the center of a three-sided test booth. Both caregiver and experimenter wore headphones throughout the experiment and listened to masking music (samba music played simultaneously with spoken text from various female speakers). Testing started with a familiarization phase during which infants heard alternations of the phrase-internal and phrase-straddling sequence and accumulated a minimum of 30 sec of looking time for each sequence (in accordance with J&S). Within the test phase, each infant was presented with two passages. One passage contained the *phrase-internal*, the other the *phrase-straddling* sequence from the familiarization phase. Which passage acted as phrase-straddling or phrase-internal during test depended on which sequence a particular infant was familiarized to. A single test trial consisted of repetitions of a passage for the same condition (internal/straddling). Trials alternated in condition (internal/straddling). Passages were presented in 12 trials distributed over three blocks. Within every block, each passage was presented once from the left and once from the right side.

The full experimental session lasted about five minutes, depending on the number of familiarization trials an infant required to reach the 30 sec familiarization criterion. Sessions were aborted earlier if the infant fussed. Data from aborted test sessions were not analyzed. After the experiment, caregivers were debriefed about the research question of the experiment.

Results

All data preprocessing and analyses have been performed using R for windows (R Development Core Team, 2020). All raw data and analysis scripts are available in the online materials.

Mixed-effect model analysis

Linear mixed-effect models were used to analyze differences in looking times between the internal and straddling passages in the test phase of the experiment. Two models were fit, one to the full dataset of all trials from all children (N = 83, n = 41 in song; 996 trials, 492 trials in song) and a second model starting from trials during which infants had attended long enough to be presented with the first 500 ms of the critical sequence within the test passage (n = 80, n = 39 in song, 680 trials, 295 trials in song). The second model on this Critical Sequence dataset was considered warranted given the overall slower speech rate and longer pauses in the present compared to the J&S stimuli, as described above. Note that three subjects were excluded from the second model because they did not contribute trials for both conditions in this dataset. The remaining 80 infants contributed an average of 4 trials per condition (range: 1-6 for both conditions). The fixed effects of both models were 1) boundary condition (internal vs. straddling, coded as an orthogonal contrast), 2) modality (song vs. speech, coded as an orthogonal contrast), 3) test-trial number linear (1 to 12, coded as the linear polynomial), 4) test-trial number quadratic (1 to 12, coded as the quadratic polynomial) and 5) the interaction of boundary condition and modality[2]. The random effects structure of both models was specified to include random intercepts for participant and by-participant random slopes for the effect of experimental condition (internal

---

[2] Note that we ran additional analyses on the full dataset with a version of model 1 that included Experiment Version as a fixed effect (4 Versions for each Modality). This model was rank-deficient and had higher AICs and BICs than the same model without this fixed effect. We therefore decided to remove this effect from model 1, and did not include it in model 2. The results for model 1 with and without the factor Experiment Version were qualitatively the same.

/ straddling). We deliberately chose not to specify the maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013), because the use of only two pairs of passages for speech and song did not warrant specification of item-related random effects. Both model 1 and 2 were fit onto box-cox transformed looking times ($\lambda = 0.12$ for model 1, $\lambda = -0.02$ for model 2 (Csibra, Hernik, Mascaro, Tatone, & Lengyel, 2016)). The R-package ”lmerTest” was used to run the models and evaluate significance of the effects (Kuznetsova, Brockhoff, & Christensen, 2016).

Results of mixed-effect model analysis

When only considering trials during which infants listened long enough to reach the critical sequence within the test passage, infants preferred to listen to the passage that contained the phrase-internal sequence in both song and speech (Figure 2.3). The second linear mixed-effect model (Table 2.2) run on this Critical Sequence data set (model 1, $n = 80$, 680 trials) revealed significant main effects of Condition ($t = 2.21$, $ß = 0.05$, $p = .03$), Modality ($t = 2.78$, $ß = 1.0$, $p = .007$) and the linear and quadratic polynomial of Test Trial Number ($t = -4.39$, $ß = -0.28$, $p < .001$; $t = 3.42$, $ß = 0.22$, $p < .001$). There was no significant interaction between Condition and Modality ($t = 0.24$, $ß = 0.01$, $p = .81$) and thus no evidence that segmentation is easier in song than in speech.

Considering all trials from all children ($N = 83$, 996 trials), we did not find evidence for a preference for passages with the phrase-internal or the phrase-straddling sequence nor did we find evidence that looking times differed between song and speech (Figure 1 in the online materials). The linear mixed-effect model 2 (Table 2.2) only indicated significant main effects of the linear ($t = -7.34$, $ß = -1.84$, $p < .001$) and quadratic ($t = 2.19$, $ß = 0.55$, $p = .03$) polynomial of Test Trial Number, indicating that overall looking times decreased over the course of the experiment, but to a lesser degree towards the end of the experiment.

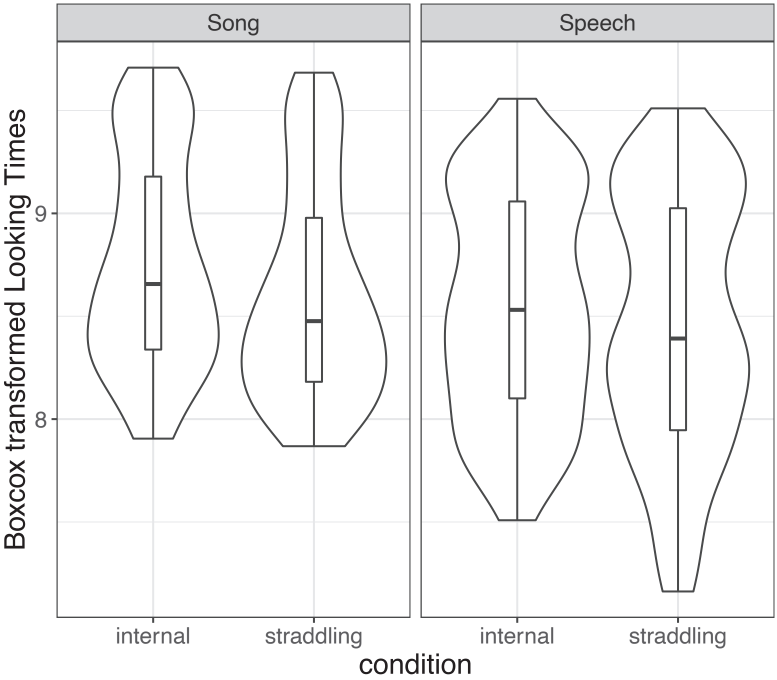


Figure 2.3 Violin- and boxplot of box-cox transformed looking times in the critical sequence dataset. Center lines represent grand medians, boxes entail first and third quartiles.

Table 2.2 Parameters of linear mixed-effect model 1 and 2.

		Model 1, Critical Sequence data set *n* = 80, 680 trials				**Model 2, Full data set *N* = 83, 996 trials**			
Predictor	Contrast Coding	β	*SE*	*t*(z)	*p*	β	*SE*	*t*(z)	*p*
Intercept		8.18	0.04	231.73	<0.001	16.25	0.17	94.90	<0.001
Condition	Internal (1), Straddling (-1)	0.05	0.02	2.21	0.03	-0.01	0.07	-0.12	0.90
Modality	Speech (-1), Song (1)	0.10	0.04	2.78	0.007	0.05	0.17	0.29	0.78
TrialLin	Linear polynomial	-0.28	0.06	-4.39	<0.001	-1.84	0.25	-7.34	<0.001
TrialQuad	Quadratic polynomial	0.22	0.06	3.42	<0.001	0.55	0.25	2.19	0.03
Condition*Modality		0.01	0.02	0.24	0.81	0.06	0.07	0.77	0.44

t-test analysis

To adhere to more standard analyses of infant looking time data we also report results of *t*-tests within both modalities using the aggregated looking times within the Critical Sequence data set (*n* = 80). Given the number of previous studies that found a preference for the internal sequence, we decided to run one-sided *t*-tests to test the hypothesis that looking times for the internal sequence are longer than for the straddling sequence. Two-sided *t*-tests will be reported for the sake of completeness. Averaged looking times were also box-cox transformed, using λ = 0.2 for song and λ = 0.36 for speech data. Levene's test indicated equal variance among groups in both the song and speech dataset. A Shapiro-Wilk test indicated that the song data deviates from normality even after transformation (*p* = 0.02). Therefore, the results of the *t*-test for song have to be interpreted with caution. Effect sizes Cohen's d_z, and Hedge's g_{av} were calculated for the untransformed looking times in the *t*-test datasets, according to recommendations given by Lakens (2013) and formulas introduced by Cohen (1988) and Hedges and Olkin (1985). A spreadsheet by Lakens (2013), available under https://osf.io/ixGcd/, was used for the calculation.

t-test results

The *t*-tests run on the averaged looking times for both song and speech trials within the Critical Sequence dataset indicated a significant preference for the internal sequence for the song modality only (Table 2.3). In both modalities, about half of the infants tested showed a preference for the internal sequence, i.e., longer listening times for the internal compared to the straddling passage.

Table 2.3 Results of the t-tests on transformed aggregated looking times for both modalities

	Song	**Speech**
n, # n preference Internal	39, 23	41, 25
Mean (SD) internal, in sec	18.40 (7.43)	13.78 (5.69)
Mean (SD) straddling, in sec	15.88 (7.51)	13.37 (6.17)
cor	0.29	0.72
Cohen's d_z, Hedges' g_{av}	0.28, 0.33	0.09, 0.07
t-value, λ (box-cox transformation)	1.86, 0.2	0.76, 0.36
p-Value, Conf Int (one-sided)	0.04 [0.10, ∞]	0.2 [-1.45, ∞]
p-Value, Conf Int (two-sided)	0.07 [-0.01, 2.30]	0.4 [-2.00, 4.54]

Note. λ = value used for the box-cox transformation.

Discussion

The current study set out to replicate 6-month-old Dutch infants' auditory grouping abilities based on intonational phrase boundaries in ID-speech (Johnson & Seidl, 2008) and assess whether this ability extends to ID-song.

Infants in the current study were tested in a paradigm first developed by Nazzi and colleagues (2000), which has successfully revealed phrase segmentation in an earlier study with Dutch 6-month-old infants (Johnson & Seidl, 2008). We replicated this latter study in the same lab, using the same speaker for the stimuli, and extended it to the ID-song modality. To this end, infants were first familiarized to two critical sequences of the same words in either song or speech (e.g., /koude pizza smaakt niet zo goed/ ("cold pizza does not taste so well")). One sequence was uttered with a well-formed phrase structure, with phrase boundaries at the edges: [koude pizza smaakt niet zo goed] while the other sequence was uttered with an ill-formed phrase structure, straddling a phrase boundary in the middle: koude pizza] [Smaakt niet zo goed. Infants were then presented with two three-sentence test passages, one containing the well-formed word sequence and the other the ill-formed word sequence. In both song and speech, infants listened longer to the passage containing the well-formed sequence. This indicates that infants were able to segment the passages of song and speech into their underlying phrasal constituents and recognized the well-formed familiarized sequence therein. Infants' known ability to recognize the phrase structure of ID-speech thus extends to ID-song.

Contribution

The current study is the first to provide evidence that 6-month-old infants segment native child songs into well-formed phrases. Infants thus capitalize on the acoustic boundary cues within song melodies to organize a continuous song into structurally relevant constituents and recognize phrases while the song unfolds. The present results significantly extend previous research on infants' musical grouping abilities by using ecologically valid musical stimuli and by requiring infants to group native song melodies into perceptual chunks while the song unfolds (Jusczyk & Krumhansl, 1990; Nazzi et al., 2000; Hawthorne & Gerken, 2013). This study also extends our knowledge on infants' recognition of phonological units in song lyrics from syllables (François et al., 2017; Lebedeva & Kuhl, 2010; Suppanen, Huotilainen, & Ylinen, 2019; Thiessen & Saffran, 2009), rhymes (chapter 3) and single words (Snijders et al., 2020) to larger prosodic units, namely phrases. The potential functional relevance of these findings will be discussed below.

The current results also contribute to two more general issues in the field of first language acquisition: the first is the question about shared cognitive mechanisms underlying the processing of music, song and speech; the second pertains to the optimal acoustic stimulus for infant language learning. Concerning the first question: infants' mental organization of speech and song into phrases observed in the current study may be grounded in a modality-general processing mechanism (Conway, Pisoni, & Kronenberger, 2009; Schön et al., 2010; Trehub & Hannon, 2006): a conceivable account would be that the salient acoustic structure of instrumental music, ID-song and ID-speech attracts infants' attention to utterance edges (De Diego Balaguer, Martinez-Alvarez, & Pons, 2016; Drake, Jones, & Baruch, 2000; Falk & Kello, 2017; Leong & Goswami, 2015). Alternatively, infants' phrase recognition in ID-song might stem from transfer of a speech-specific or even native language-specific prosodic parsing strategy to the song modality (Morgan & Demuth, 1996). Future research should identify the exact mechanisms underlying phrase segmentation and clarify to what extent these are bound to a specific developmental stage, input modality or language. Our contribution to this open issue is the finding that at six months, infants' perception of phrase structure is not limited to speech-specific boundary cues (Johnson & Seidl, 2008; Seidl, 2007; Seidl & Cristia, 2008; Wellmann et al., 2012) but encompasses a more generic phrase boundary percept in song melodies, a finding that needs to be incorporated into future accounts of infant speech segmentation.

The second general contribution of the current study concerns the question about the kind of acoustic stimulus from which infants learn best. Infants' astonishing learning success in their first year of life has been attributed to the exaggerated acoustic shape of ID-speech (Kuhl et al., 1997). If this were the case, then infants should learn even better from ID-song, a type of stimulus that is even more exaggerated when compared to ID-speech in terms of pitch, rhythm and tempo (Trehub et al., 1997). In the current study, the pre-test confirmed the naturalness of the song and speech stimuli and infants showed increased attention to the songs versus speech stimuli. Also, the effect sizes of the speech and song modality were in the predicted direction (speech Cohen's $d_z = 0.09$; song Cohen's $d_z = 0.28$). Nevertheless, the current study provided no evidence for easier segmentation in ID-song than ID-speech. This is contrary to previous studies which reported a song benefit for infants' linguistic processing (François et al., 2017; Lebedeva & Kuhl, 2010; Thiessen & Saffran, 2009), but is in line with other work where no processing benefit for songs was observed (Snijders et al., 2020; Suppanen et al., 2019). In the following we will discuss possible reasons for the lack of a song advantage in the current study.

Understanding the absence of a modality effect

Absence of evidence for easier segmentation from songs compared to speech might reflect the relative acoustic similarity between our song and speech stimuli, resulting from the fact that the stimuli in both modalities were created to be analogous. As a result of this necessary experimental control, the speech stimuli may have been slower while the song stimuli may have displayed less repetition compared to their respective real-life counterparts. Alternatively, the hypothesized processing benefits of ID-song might have been present in the current study but counteracted by the higher familiarity of ID-speech, resulting in overall similar segmentation outcomes from song and speech. Also, it may simply be that more statistical power is needed to provide evidence for an interaction between modalities (speech/song) and phrase segmentation. As the data of the present study are inconclusive regarding the cause of the absence of a modality effect, future studies should elucidate in how far ID-song boosts, hinders, or truly has no impact on infants' segmentation abilities.

Limitations of the replication

Infants' preference for the passages containing the well-formed sequence in both song and speech was only evident in an analysis that differed from the study we aimed to replicate and extend (Johnson & Seidl, 2008). To understand the first difference between the analyses, one should remember that the test passages consisted of three sentences. The familiarized sequences occurred within the second sentence (see Table 2.1). Our analysis only included looks after infants had attended long enough to be presented with the first 500 ms of the critical sequence within the test passages (316 of 996 trials and 3 infants excluded). Johnson and Seidl (2008), on the other hand, analyzed data from all test trials. The change in analysis seemed warranted given that our stimuli were substantially slower than those in the previous study (Johnson & Seidl, 2006). As a second difference, we made use of mixed-effect models on single-trial data in addition to t-tests on data averaged over trials within children. Using this more sophisticated analysis technique might have been necessary because of the relatively small effect sizes in the present study (Cohen's d_z = 0.09 and 0.28 in the aggregated data of speech and song, respectively) compared to the somewhat larger effect observed in the study by Johnson and Seidl (2008, Experiment 1; Cohen's d_z = 0.35).

In an attempt to understand why the effect size in the current study was smaller we can outright disregard a number of factors: language, age, experimental set-up, and even the speaker of the stimuli, and the lab in which the study was conducted were all the same as in the original study. A factor that might have impacted the effect sizes is the tempo of the experimental stimuli. For one, the critical sequences occurred within three seconds from the start of the test passages of the original study (Johnson & Seidl, 2008; range: 1.26-2.99 sec) but only up to 4 sec into the passages of the present study (range: 1.66-4.23 sec, see Table 5 in the online materials). Consequently, infants in the current study had to listen longer before they encountered the critical sequences. Secondly, the comparatively long pauses between the consecutive sentences of the test passages of the present study (range: 400-850 ms in Johnson & Seidl, 2008; range: 923-1541 ms in the current study; Table 4 in the online materials) might have created a less coherent auditory percept of the passages, resulting in overall more challenging listening conditions and hence smaller effect sizes. Despite these differences, the present study nevertheless provides moderate support for infants' processing of prosodic structure in ID-speech (Johnson & Seidl, 2008; Nazzi et al., 2000).

Future research

Previous research has considered ID-song first and foremost as a means of stimulating affiliation and mood regulation (e.g., Cirelli et al., 2018). Consistent with this view, songs are not always included in descriptions of infants' speech input (e.g., Cristia, 2013; Golinkoff et al., 2015). However, other descriptions of infants' linguistic input have been broadened to include ID-song (Bergelson, Amatuni, Dailey, Koorathota, & Tor, 2019; Soderstrom & Wittebolle, 2013). The present data contribute to the mounting evidence that songs can indeed be a source for infants' implicit linguistic learning (François et al., 2017; Hahn et al., 2018 (chapter 3); Lebedeva & Kuhl, 2010; Snijders et al., 2020; Suppanen et al., 2019; Thiessen & Saffran, 2009). Consequently, ID-song should be included in descriptions of the linguistically relevant input that infants receive.

In how far could phrase segmentation from ID-song be relevant to infant language acquisition? For one, it could aid infants in identifying smaller linguistic units within the song lyrics, e.g., words occurring at phrase boundaries (see for a similar account for speech e.g., Johnson et al., 2014; Shukla et al., 2011), and help to transfer the song lyrics and melody into working memory by chunking them into manageable units. This, in turn, might help infants to identify the song and its lyrics across different occasions in their daily routines and across different singers, contributing to the formation of context- and singer-independent abstract representations (Kragness, Johnson, & Cirelli, 2021). Phrase segmentation from ID-song might also, indirectly, benefit the processing of (ID)-speech: By attending to melodic phrases in songs infants train to allocate attention to important units in the song input. The same units are also relevant in speech, but presumably less salient and occurring at a much faster time scale. Caregiver singing could thus provide infants with an acoustic playground, a practice field to engage mechanisms that are also at work in the presumably more demanding speech signal.

Future research should investigate the functional relevance of infants' ability to segment songs into phrases. There is ample evidence that prosodic phrase segmentation of speech is a key prerequisite for lexical and morpho-syntactic development (Carvalho et al., 2018). It has even been suggested that impaired recognition of large phrasal boundaries in speech is the key underlying deficit for developmental language disorder (Richards & Goswami, 2019). Consequently, future research should investigate to what extent caregiver singing and other types of rhythmic-melodic input such as rhyming story books (Richards & Goswami, 2019) contribute to infants' perception of phrasal boundaries in speech. Such a relationship between language play and real-life linguistic abilities would speak to recent studies suggesting a link between rhythmic-melodic processing of music and speech on the one hand and grammar development on the other (Gordon, Jacobs, Schuele, & McAuley, 2015; Leong & Goswami, 2015; Politimou et al., 2019). The current study contributes an empirical foundation for such future investigations, by showing that for young infants the major phrasal units in ID-song are at least as accessible as in ID-speech.

Conclusion

Recognizing phrases in continuous speech is a cornerstone of the development of speech perception. This study replicated a previous finding regarding Dutch 6-month-olds' recognition of phrase structure of ID-Speech (Johnson & Seidl, 2008) and extended the results to ID-song. Thus, already within their first half year of life, infants actively process sung input online and memorize well-formed sung phrases. Future research should identify the mechanisms underlying this ability and clarify whether the recognition of the phrasal structure of caregiver singing contributes to linguistic development.

3

Chapter 3

Infants' sensitivity to rhyme in songs

This chapter is a slightly modified version of:
Hahn L. E., Benders T., Snijders T. M., Fikkert P., 2018. Infants' sensitivity to rhyme in songs. *Infant Behavior and Development*. 52:130-139.

Online materials for this chapter are available at: https://osf.io/gnadw/

Abstract

Child songs often contain rhyming words at phrase endings. In this study, we investigated whether infants can already recognize this phonological pattern in songs. Earlier studies using lists of spoken words were equivocal on infants' spontaneous processing of rhymes (Hayes, Slater, & Brown, 2000; Jusczyk, Goodman, & Baumann, 1999). Songs, however, constitute an ecologically valid rhyming stimulus, which could allow for spontaneous processing of this phonological pattern in infants. Novel child songs with rhyming and non-rhyming lyrics using pseudowords were presented to 35 9-month-old Dutch infants using the Headturn Preference Procedure. Infants on average listened longer to the non-rhyming songs, with around half of the infants however exhibiting a preference for the rhyming songs. These results highlight that infants have the processing abilities to benefit from their natural rhyming input for the development of their phonological abilities.

Acknowledgements

We thank Ilona van der Linden for her support in creating the stimuli, Annelies van Wijngaarden for lending us her beautiful voice, Esther Kroese, Renske van der Cruijsen and Francie Manhardt for testing the infants and coding the videos. Furthermore, we would like to thank the First Language Acquisition Group at Radboud University, Nijmegen, NL, Charon Translations, Nijmegen, NL and two anonymous reviewers for their helpful feedback.

Introduction

Infants' daily routines are often accompanied by songs and nursery rhymes. Caregivers recite nursery rhymes during playtime and sing lullabies to soothe their child or to entertain him/her, e.g., while driving or during a diaper change (Trehub et al., 1997). Songs and nursery rhymes clearly serve a social-emotional function. The body of research on phonological processing of rhymes in toddlers and kindergartners is quite extensive, but only rarely makes use of ecologically valid stimuli such as songs and nursery rhymes. The central question in this study is whether 9-month-old infants exploit songs and rhymes to spontaneously process phonological patterns of their native language. A very frequent phonological phenomenon in songs and nursery rhymes is the occurrence of rhymes at phrase endings. A phrase-final rhyme is a sound pattern that is repeated at the end of verses or phrases (Fabb, 1999). In Western child songs, rhymes typically include the last stressed syllable and, if present, the following unstressed syllables. The unit of repetition in phrase-final rhymes typically is a whole word, excluding the word-onset (e.g., Stallworthy, 1996).

Previous research investigating the effect of rhyme on linguistic processing in children mostly focused on kindergartners and has shown clear benefits of rhymes in word learning and phonological processing tasks. For example, 2- to 4-year-olds benefit from rhyming stories when learning new words, as they are better able to learn a novel word if it rhymes with the phrase-final word of the previous verse (Read, 2014). Children in the same age range also benefit from listening to rhyming stories, as opposed to prose stories, before completing a rhyme/alliteration detection task (D. S. Hayes, 2001). These effects might be attributed to the increased predictability of the phonological content of a rhyming story (Cook, 1997; Rubin, 1995). This higher predictability of phonological content might in turn increase phonological sensitivity, as is evident from the results of the rhyme/alliteration detection task (D. S. Hayes, 2001). Similar conclusions can be drawn from a study with 6-year-olds (Sheingold & Foundas, 1978). After having listened to a story in rhyme or prose, children were asked to recall details of the story. Children were equally successful in that task, regardless of whether they had heard a rhyming or non-rhyming story. However, the children were more successful in sorting the events of rhyming stories into ordered sequences than those of non-rhyming stories. The authors hypothesized that the rhyming verse couplets were helpful to reconstruct the order of content appearing in the story, due to the predictability of the phonological form of rhyming words. Thus, toddlers and preschoolers already make use of rhymes to predict upcoming linguistic input and to memorize the content of a story.

A crucial component of rhyme sensitivity is the ability to recognize that rhyming words are different from repetitions of identical words: children have to be able to recognize that e.g., *bear* and *pear* share phonological material, but differ in their onset consonants. This is an important ability, as it has been suggested that the awareness of such subtle phonological differences has important implications for later literacy development (De Cara & Goswami, 2002; Goswami, 2002). Infants as young as 11 and 18 months of age are already sensitive to small changes to onset consonants (Swingley, 2005, 2009). Additionally, 2-year-old children are able to identify the target in a set of rhyming items as soon as they hear the non-overlapping onset consonant (Swingley, Pinto, & Fernald, 1999). Upon acquisition of novel words, 14-month-olds are able to recognize subtle phonological differences between highly similar items, such as *din* and *bin* (Yoshida, Fennell, Swingley, & Werker, 2009). Moreover, infants as young as 16 months are even able to group such rhyming words apart from another set of words that also overlaps in phonological material but do not rhyme (e.g., *dib* and *gib* vs. *deb*) (Floccia, Nazzi, Delle Luche, Poltrock, & Goslin, 2014). At the start of their second year of life infants thus already possess impressive phonological processing abilities that allow them to differentiate between and acquire rhyming single words.

However, there is only sparse evidence for rhyme processing in infants younger than one year of age and none of the studies in question used ecologically valid stimuli that nursery rhymes or songs provide. In one study that used a variant of the Headturn Preference Procedure (Kemler Nelson, Jusczyk, & Mandel, 1995), 9-month-old American-English infants did not show a higher sensitivity for CVC word lists that rhymed as compared to CVC word lists that did not (e.g., *bad, pad, lad (shared –VC) vs. vip, zut, mog (no phonological overlap)*, Jusczyk, Goodman, & Baumann, 1999, experiment 1). However, infants in the same study and using the same procedure were sensitive to the mirror image of rhymes, that is phonological overlap at the beginning of the syllable (shared CV-, experiment 2) and overlapping syllable onsets (shared C-, experiment 3). Another study did find sensitivity to rhyming versus non-rhyming single words in 2-year-old English toddlers, but did not find evidence for this ability in 8.5-month-old infants (Braze, McRoberts, & Mc Donough, 2011). Within this study, both toddlers and infants were tested using the same visual fixation paradigm. While the toddlers looked significantly longer when hearing rhyming versus non-rhyming words, no significant looking time differences were found for infants tested on the same set of words. Infants thus recognize repetition of intra-syllabic units at word onsets in lists of single words, but do not spontaneously display such recognition for rhyming single words.

As a counterpoint to the absence of a spontaneous discrimination between rhyming and non-rhyming stimuli reported in the work by Jusczyk and colleagues (1999) and Braze and colleagues (2011), two other studies with 7-to-13-month-old English infants found that they were able to detect changes from one rhyme pattern to another under specific circumstances (Hayes et al., 2000; Hayes, Slater, & Longmore, 2009). These infants were able to detect a change from one rhyme (e.g., *bad, dad, fad*) to another (*beg, deg, feg*) in a Conditioned Headturn Preference Procedure (Werker, Polka, & Pegg, 1997), suggesting that they are sensitive to the shared rhymes between words. Crucially, infants in both studies by Hayes and colleagues were trained to respond to a change in rhyme rather than showing a spontaneous reaction as assessed in an unconditioned Headturn Preference Procedure. Moreover, due to the elaborate training required for this paradigm, only a subset of the tested infants actually contributed analyzable trials. This gives rise to the possibility that the early rhyme sensitivity observed in this study only holds for rather mature infants that might differ substantially from a set of average infants. The study nevertheless provides a first indication that infants can potentially process rhymes at an early age under certain circumstances.

Taken together, the evidence so far suggests that infants do not have a spontaneous preference for rhyming or non-rhyming word lists (Jusczyk, Goodman, et al., 1999), indicating that they are not able to differentiate between rhyme and non-rhyme. However, they can be trained to react to a change in rhyme in lists of words (Hayes et al., 2000, 2009). Although a specific sensitivity to rhymes is clearly present early in life, these results do not shed light on the question whether infants can and do spontaneously use this sensitivity when they encounter rhymes in their daily lives.

A context that may facilitate infants' spontaneous (instead of trained) processing of rhymes is song. Songs contain rhymes within a solid scaffold of melody and rhythm, which caregivers express through various multi-modal cues (Bergeson & Trehub, 2002; Delavenne et al., 2013; Falk & Kello, 2017; Longhi, 2009; Nakata & Trehub, 2011). These structural cues potentially enable infants to attend longer to infant-directed song compared to infant-directed speech (e.g., Costa-Giomi, 2014). In fact, research has repeatedly identified infant-directed singing as a tool for infant arousal regulation (e.g., Trehub & Nakata, 2002). Prolonged attention to song might subsequently enable infants to extract information from song that is otherwise inaccessible to them. Yet, there are very few studies exploring infants' linguistic processing abilities in song, and none focused on the processing of rhyme. Two studies showed that infants can detect a change in the order of a string of syllables only when the syllable string is sung on a melody, but not when the string is spoken (Lebedeva & Kuhl, 2010; Thiessen & Saffran, 2009). Both studies used the Headturn Preference Procedure, with infants from 6 to 8 months old (Thiessen & Saffran, 2009) and 11 months old (Lebedeva & Kuhl, 2010). These studies provide evidence that infants actively encode the phonological content of songs, an important prerequisite to detect rhymes at phrase endings in songs.

It has recently been shown that the scaffold of melody and rhyme has potential benefits for adult language learners (Bebout & Belke, 2017). In this study, adults acquired the gender system of an artificial language by listening to sentences describing action sequences. Learners exposed to sentences with both rhymes and melody outperformed the learners exposed to spoken prose, only-rhyming or only-sung sentences. These results show that rhyme and melody in a song can form the optimal stimulus for language learning. Perhaps infants are already able to benefit from these structural features in rhyming songs for the acquisition of their first language. For this, however, it is necessary that infants spontaneously process rhymes at phrase endings in songs. Presumably, detecting a rhyme at phrase ends requires more segmentation and working memory skills than detecting rhymes in single words. On the other hand, the rhyming phrases and not the single words are in accordance with the vast amount of rhyming input infants get. In the current study we move from rather artificial single word stimuli to a stimulus that is familiar to infants. In this way we will shed light on infants' processing abilities of natural infant-directed singing, as opposed to simple single words. Spontaneous rhyme processing could also be fostered by the phrase-final position at which rhymes typically occur within Western child songs. In speech, phrase boundaries have been associated with enhanced phonological processing in infants (Johnson et al., 2014). For songs, there is recent evidence that already at 6 months of age, infants are sensitive to the phrase boundaries in song melodies (chapter 2). Consequently, the current study investigates spontaneous processing of rhymes at phrase-endings.

In the current study, we used the Headturn Preference Procedure to investigate whether infants spontaneously process rhymes that are presented at phrase boundaries in songs, an ecologically valid rhyme context that potentially facilitates rhyme processing, without explicit training. With this design we aim to lay the ground for future research that can explore the possibility that rhyming songs are a source of linguistic learning for infants. So far, songs and nursery rhymes have only very rarely been studied as a potential vehicle for language learning in infancy and no study yet has focused on rhyme. To be comparable with previous studies (Braze et al., 2011; Hayes et al., 2000; Jusczyk et al., 1999), we tested 9-month-old Dutch infants. However, novel child songs were used as stimuli, as opposed to the lists of single words used in previous studies. Due to this difference in stimuli, no concrete prediction in terms of the preference direction of infants could be made.

However, given the ubiquity of phrase-final rhymes in song and nursery rhymes (Burling, 1966; Rubin, 1995), a familiarity preference for rhyming songs seemed more likely. Crucially, a preference in either direction would indicate infants' spontaneous processing of rhymes in songs.

Method

Participants

The final sample of participants consisted of 35 infants (21 female, Mage in days = 292; range 275-304). Twelve more infants were tested but excluded because they cried during the experiment (n = 3), because they clapped or danced while listening to the songs and where therefore unable to perform the headturn procedure (n = 6), because they only contributed trials for one experimental condition (n = 1), or because the parent terminated the session (n = 2).

A statistical power analysis using G*Power was conducted to obtain an estimation of the required sample size (Faul et al., 2007). The design and procedure of the current study were comparable to the study by Jusczyk et al. (1999). Their experiment 2 on sensitivity to repeating word-onset clusters was used as a reference for the power analysis, as their experiment 1, on rhyme sensitivity, did not render a statistically significant effect. Based on the means and standard deviations reported for Experiment 2 and a conservatively estimated correlation between groups of 0.2, we calculated Cohen's $d_{\text{Experiment 2}} = 0.506$. The power analysis based on this estimated effect size revealed a minimum required sample size of 33 participants to reach 80% power in a two-sided *t*-test with alpha = 0.05. Consequently, we aimed at a final sample of at least 33 infants, and tested 47 infants, anticipating 30 % drop-out.

All participating infants were reported by the parents to be born full-term, have normal hearing, no familial risk of language or reading problems and come from a monolingual Dutch household. Parents of one child reported that the child is exposed to Berber during occasional family visits on weekends (about six hours per month). On a questionnaire administered before the experiment, the visiting caregivers indicated the amount of exposure to different musical sources for their infant (see online version of this article for an English translation of this questionnaire). The majority of caregivers indicated that their child listens to songs sung live by one of the infants' caregivers on a daily basis (90%, singers are mostly mothers). All children but one were exposed to songs from birth onwards (one child from the age of 4 months). These data were not used for analysis of individual differences, as the frequency of song and rhyme exposure was highly similar across children (93% of participating infants are exposed to Dutch child songs on a daily basis).

All parents were given the choice between a children's book and 10€ as compensation for their participation. The parents gave informed consent prior to participating. Testing adhered to APA ethical standards and was conducted in the Baby and Child Research Center at Radboud University Nijmegen, the Netherlands.

Materials

The stimuli were a set of child songs, consisting of novel melodies with accompanying novel lyrics. Each song occurred in a rhyming and non-rhyming version. These versions only differed in the phrase-final syllable and were otherwise identical.

Melodies. Novel melodies were created to exclude song familiarity as a potential factor influencing the results. All melodies comprised two melodic phrases that had a median length of seven tones each (range 6 – 9 tones). Melodies were set in a major key and consisted of intervals up to a quint. To mimic the two-phrase structure of many Germanic folk songs (Shanahan & Huron, 2011), the second melodic phrase was constructed to sound closed by using a falling melodic contour that ended on the tonic. Melodies were notated using the software MuseScore (musescore.org). An example melody can be seen in Figure 3.1. The sound files and lyrics of the stimuli can be found in the online materials.

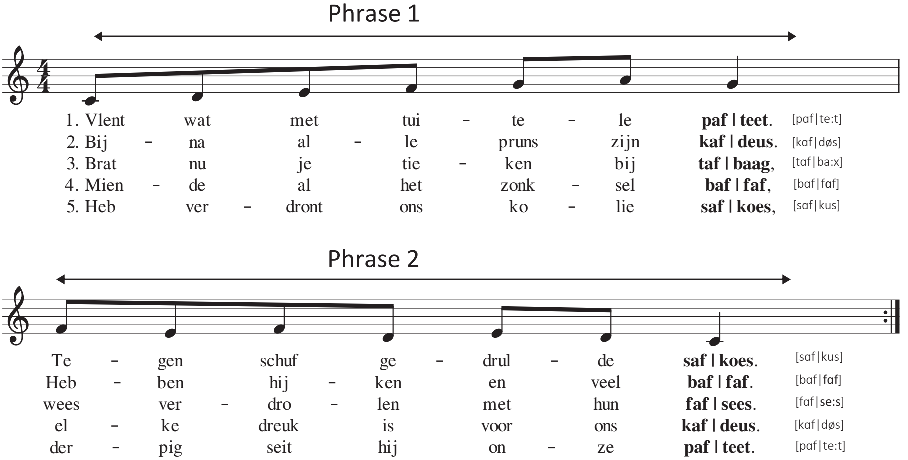


Figure 3.1 Example stimulus. lyrics in Nonsense-Dutch, Phrase-final pseudowords in **bold** and International Phonetic Alphabet (IPA) transcription (International Phonetic Association, 1999); Numbers 1 to 5 = Verse 1 to 5.

Lyrics. For every melody, a rhyming and a non-rhyming version of the lyrics was created, which only differed in the phrase-final pseudowords: monosyllabic CVC combinations of voiceless obstruents [p, t, k, f, s, x] and vowels [i, e, o, ɔ, ɑ, a, u, ɛi, ø]. In the rhyming version, the phrase-final words of the phrases all rhymed, whereas in the non-rhyming version they did not rhyme (see again Figure 3.1, and see online materials for the full set of lyrics).

The lyrics were written using Dutch function words and auxiliaries, and phonotactically legal pseudowords for content words. These Nonsense-Dutch words were used to avoid any word familiarity effects. For every two-phrase melody, five two-phrase verses were written (1 to 5 in Figure 3.1), rendering a 10-phrase song. Each phrase ended with a slot for a monosyllabic phrase-final word. These slots were filled with rhyming pseudowords for the rhyming version (e.g., [pɑf, sɑf, fɑf, kɑf]) and filled with non-rhyming pseudowords for the non-rhyming version (e.g., [tet, kus, døs, bax]). Some phrase-final words were used twice within the same song and version (e.g., [kɑf] and [døs] in Figure 3.1), due to the relatively small number of rhyming mono-syllabic Dutch pseudowords. If a pseudoword had to be repeated within the rhyming version of a song, it was placed as far away as possible from the first occurrence of the word. The corresponding phrase-final word in the non-rhyming version was also repeated at the same position. The text setting of the songs was judged to be appropriate by two Dutch native speakers.

Songs. *Recording procedure.* A female Dutch singer, with more than ten years of singing experience in choirs and solo-lessons, was recorded singing the songs. She was instructed to sing in a friendly, infant-directed manner. A visual metronome was displayed to the singer, set at 61 beats per minute. This tempo was found to be the most convenient tempo for the singer and enabled her to sing both versions of a song at nearly equal tempo. The singer used a tuning fork to start the rhyming and non-rhyming song versions from the same tone. The songs were recorded in a sound-proof recording studio using Adobe Audition and were exported as .wav files with a sampling rate of 44.1 kHz and a resolution of 16 bits.

Acoustic manipulation. Noise and disturbing sounds were deleted from the recording with Audacity (version: 2.0.5). All other acoustic manipulations were performed using Praat (Boersma & Weenink, 2014). All songs were set to 65 dB intensity, which was the median intensity of the original songs. The authors of the study found the original recording of the songs to be rather slow for typical child songs. Therefore, the tempo was increased by shortening the songs' duration by 10%, using a duration tier and resynthesis with overlap-add in Praat. A 500 ms period of silence was inserted between phrases of a song, to ensure that gaps between phrases were of equal duration across the rhyming and non-rhyming versions of each song. All songs faded out at 26 seconds using a two-second fade-out resulting in songs of 28 seconds long (median of 8 phrases per song version, range: 8 to 9 phrases).

Selection of the final song sample. Nine songs were selected from the full set of songs. These nine songs were considered to sound most natural, judged by the authors of the study. Furthermore, their text could easily be transcribed by two Dutch native speakers, indicating that the Nonsense-Dutch lyrics conform to real Dutch phonology. Duration, fundamental frequency and intensity of the songs' phrases and phrase-final words are summarized in Table 3.1. A by-stimulus repeated measures ANOVA confirmed that no significant differences were present between the rhyming and non-rhyming versions of a song in terms of duration, pitch and intensity (see the online materials for more details on this analysis).

Table 3.1 Acoustic characteristics of the song stimuli.

		Rhyming	**Non-rhyming**
Duration (ms)	Phrases	2732 (274)	2728 (255)
	Phrase-final word	648 (99)	626 (92)
Fundamental frequency (Hz)	Phrases	326 (34)	327 (35)
	Phrase-final word	293 (63)	291 (64)
Intensity (dB)	Phrases	65 (2)	64 (2)
	Phrase-final word	53 (7)	55 (7)

Note: mean values with standard deviations in parenthesis. All pitch analyses were done using Praat's standard interpolation method using a pitch range of 75 to 1200 Hz, to accommodate the relatively high pitch in singing.

Procedure

The experiment was run using the Headturn Preference Procedure. Three lights were placed within a three-sided booth at infant eye-level: a blue light in the center and red lights on the right and left walls of the booth. A camera was hidden below the center light to observe infant behavior from outside. Stimuli were presented via loudspeakers below the red lights. The infant and caregiver were seated in the middle of the booth, directly opposite the blue center light, exactly in between the left and right red lights.

Stimulus presentation was controlled from outside the test booth by the experimenter, using the stimulus presentation software Look! (Meints & Woodford, 2008). The experimenter was blind to trial number and trial condition. The experimenter coded the looking behavior of the infant (left, right, center) using assigned keys. The same procedure was used for both practice trials and test trials. The whole session was recorded on video for later reliability coding (see section "reliability coding" below).

Before the experiment started, parents were briefed about the experimental procedure by the experimenter and filled out the questionnaire about their singing habits (see the online materials for an English translation of the questionnaire). Parents informed the experimenter about the current health status of the child (e.g., whether the child had a cold at the day of testing or had pain from teething). Recording this information is standard practice in our lab and serves possible post hoc exclusion of children. However, this did not prove to be necessary within the sample of infants in the current study. At the start of the experiment, infants were seated in the test booth on their caregiver's lap. Both caregiver and experimenter wore headphones throughout the experiment, listening to masking music (samba music played simultaneously with spoken text from various female speakers). Each trial started with the blue light in the center blinking. Once the infant fixated this light, the blue light stopped blinking and either the left or right red light began to blink. The sound stimulus started as soon as the infant turned her head to the blinking light for one second. The stimulus was continuously played until the infant turned her head away from the light for more than two consecutive seconds or until the maximum stimulus duration of 28 seconds was reached. The full testing session lasted about five minutes. Test sessions were aborted earlier if the infant fussed or if the parent indicated to stop the session. After the experiment, parents were debriefed about the research question of the experiment.

The experiment was divided in a training and test phase. During training infants could become acquainted with the lights in the testing booth and the head-turn contingent start and continuation of the stimulus presentation. For the training phase, two songs that were not used during the test phase were played. Both the rhyming and non-rhyming versions of these two songs were used for training, counterbalanced between infants for condition (rhyming / non-rhyming), order (first / second training trial) and presentation side (left / right). After these two training trials, fourteen test trials immediately followed, divided into two blocks of seven songs each. The rhyming and non-rhyming versions of each song appeared in different blocks. Infants thus never heard the same melodies within one block. The order of training songs and test songs was pseudorandomized using Mix (Casteren & Davis, 2006), ensuring that the same side of presentation (left / right) and condition (rhyming / non-rhyming) appeared maximally three times in a row.

Analysis plan

Linear mixed-effect modeling. As opposed to the more traditionally used t-tests or ANOVAs, linear mixed effect models do not require aggregation across trials and include all individual trials of the infants (Quené & van den Bergh, 2004).
The linear mixed effect model was fit onto Box-Cox transformed looking times

(lambda = 0.32), because the residuals of the same model with untransformed data were not normally distributed, as assessed through visual inspection (see also Csibra, Hernik, Mascaro, Tatone, & Lengyel, 2016 for recommendations on why to log-transform infant looking time data). The fixed factors of the model were 1) condition (rhyme, non-rhyme), coded as orthogonal contrasts; 2) test trial number (1 to 14), coded as a linear polynomial; and 3) the interaction of condition and test trial number. The model included random intercepts for subjects (n = 35) and songs (only the ones used during the test phase of the experiment, n = 7) and by-subject and by-song random slopes for the effect of condition. The R-package “lmerTest” was used to run the model and evaluate significance of the effects (Kuznetsova et al., 2016).

t-test analysis. Paired t-tests on the looking time data were run in addition to the linear mixed-effect model to facilitate comparison to previous literature. Looking times per condition were averaged for each participant. These averaged looking times were not transformed, again to stay in accordance with previous literature and since a Shapiro-Wilk test and a Levene’s test was not significant.

Effect sizes. Cohen’s d_z, and Hedge’s g_{av} were calculated for the t-test dataset, according to recommendations given by Lakens (2013) and formulas introduced by Cohen (1988) and Hedges and Olkin (1985). A spreadsheet by Lakens (2013), available under https://osf.io/ixGcd/ was used for the calculation.

Results

Infants on average listened longer to the non-rhyming songs (see Figure 3.2). The linear mixed-effect model (see Table 3.2) however indicated no significant difference in looking times for rhyming and non-rhyming songs (t = 1.78, β = 1.69, p = 0.09; effect of Condition in Table 3.2; see also Figure 3.2). Looking times decreased over the course of the experiment (t = -3.77, β = -13.16, p = 0.0002; effect of Trial in Table 3.2). There was no significant interaction between trial number and condition (p = 0.76).

The paired t-test on the difference in looking time between rhyming and non-rhyming songs was also not significant ($M_{\text{non-rhyming}}$ = 7.22 sec, $SD_{\text{non-rhyming}}$ = 3.71, M_{rhyming} = 6.08 sec, SD_{rhyming} = 3.05, $t(34)$ = 1.661, p = 0.11, 95% CI [-0.25, 2.53], r = 0.30). Eighteen of the 35 participants had a preference for the non-rhyming songs. The effect size was small (Cohen’s d_z = 0.28, Hedges’ g_{av} = 0.33).

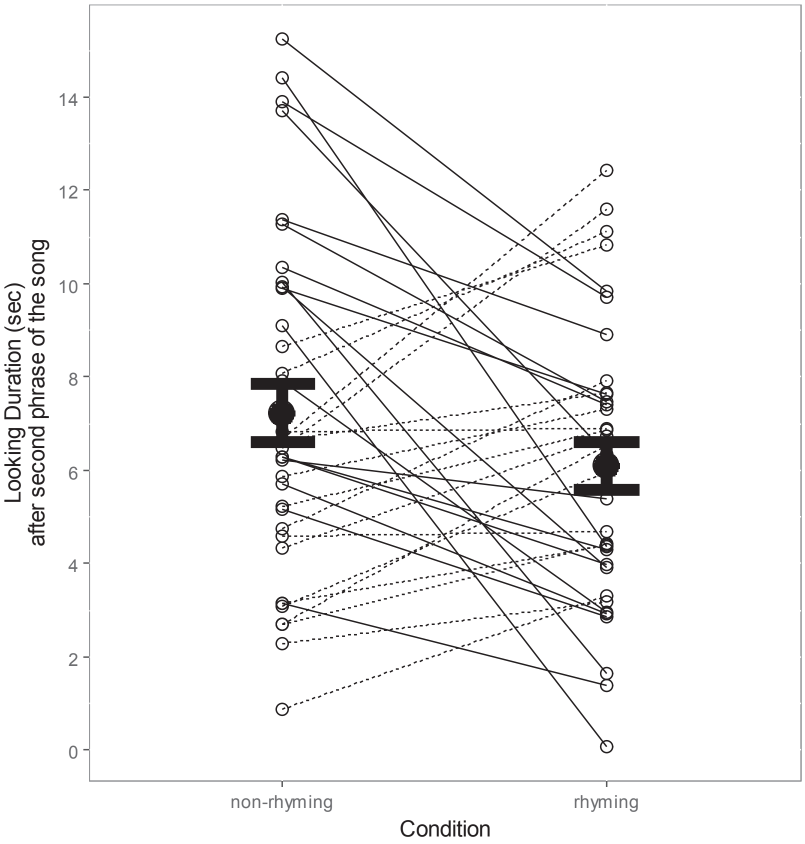


Figure 3.2 Averaged looking times to rhyming and non-rhyming songs; Means and 95% confidence intervals. White dots denote average looking times of individual infants for each condition. Dotted lines indicate infants with a preference for rhyming, solid lines for infants with a preference for non-rhyming songs.

Table 3.2 Parameters of the linear-mixed-effect model, based on looking times after the second phrase of every song

Predictor	*Contrast Coding*	β	*SE*	*t(z)*	*p*
Intercept		43.48			
Condition	non-rhyme (1), rhyme (-1)	1.69	0.95	1.78	0.09
Trial	linear orthogonal	-13.16	3.49	-3.77	0.0002
Condition * Trial		-1.05	3.47	-0.30	0.76

For sake of completeness, we inspected the two training trials of the experiment and found that only 18 of the 35 infants actually contributed a trial for both conditions, that is, they listened long enough to both training trials to reach the end of the second phrase of the song. For this subset of the sample, we found a numeric preference for rhyming over non-rhyming songs ($M_{\text{non-rhyming}} = 11.69$ sec, $SD_{\text{non-rhyming}} = 7.59$, $M_{\text{rhyming}} = 13.05$ sec, $SD_{\text{rhyming}} = 7.17$). We did not confirm this difference statistically, due to the relatively small number of trials, items and participants.

Discussion

The current study investigated whether 9-month-old infants are already able to spontaneously process the phrase-final rhyme in songs. Infants overall listened longer to non-rhyming songs compared to rhyming songs despite the fact that around half of the infants showed a preference for rhyming songs. This suggests that infants are able to process the rhyming pattern at phrase endings in songs. Our results bear the potential that songs in general and rhyming words at phrase ends in songs in particular are an accessible source of phonological structure learning for infants. Early rhyme sensitivity and individual differences there in can thus be studied with relatively simple paradigms and ecologically valid stimuli. Our design can also function as an example for future studies to explore the association between language play and later literacy development. However, since the effect of interest did not reach standard levels of statistical significance, future replications are necessary before firm conclusions can be drawn. Potential directions for methodological improvements will be discussed further below.

The evidence obtained in this study adds to previous research on infants' processing of rhyme in spoken single words (Jusczyk et al., 1999 and Hayes et al., 2000, 2009). So far, there has only been evidence for infants' spontaneous processing of phonological overlap at word onsets, but not offsets (Jusczyk et al., 1999) and for rhyme processing after extensive training (Hayes et al., 2000, 2009). The current study extends these previous results, suggesting that 9-month-old infants can spontaneously process phonological overlap at word offsets and do not require any training to do so. It may be important that the rhyming information is transmitted in a natural and accessible stimulus, which a song provides.

The stimuli in the present study had two crucial features. First, our rhyming stimuli were embedded in prototypical child song melodies. Previous research has shown that language processing in infants as well as adults benefits from the tight coupling of linguistic and musical information in songs (Bebout & Belke, 2017; Lebedeva & Kuhl, 2010; Thiessen & Saffran, 2009). Second, rhymes were placed at phrase boundaries. The phrase boundary position in speech is of particular salience to infants (Johnson et al., 2014) and 6-month-old infants already are sensitive to phrase boundaries in songs (chapter 2). Future studies are needed to disentangle how sung versus spoken material and rhyming phrases versus isolated words each contribute separately to infants' rhyme processing.

A preference for phonological overlap has been reported for adults, when listening to rhyming versus non-rhyming poems (Obermeier et al., 2013), and for 9-month-old infants, who displayed a preference for phonological overlap at word onsets but not at rhyme position (Jusczyk, Goodman, et al., 1999). With respect to the linguistic input that infants are exposed to, rhymes are highly frequent in the lyrics of songs and nursery rhymes (Burling, 1966; Rubin, 1995). The results from the song exposure questionnaire, as filled out by parents of infants participating in this study, suggested that our 9-month-old Dutch participants had vast experience with rhymes in their input, in particular in songs. As we noted earlier, a preference in either direction (rhyming/non-rhyming songs) would indicate infants' spontaneous processing of rhymes in songs. Because listeners tend to prefer phonological overlap, and infants often display a preference for the linguistic structures that surround them (e.g., Höhle, Bijeljac-Babic, Herold, Weissenborn, & Nazzi, 2009; Segal & Kishon-Rabin, 2012; Swingley, 2005), a preference for the rhyming songs could have been expected. Contrary to this expectation, infants in the present experiment overall displayed a preference for the less common non-rhyming songs. As rhymes are particularly frequent in songs, the preference for the non-rhyming pattern may be specific to the sung register, and may not generalize to spoken single words (Jusczyk et al., 1999, experiment 1). This stimulus specific interpretation of the preference direction has been suggested earlier (Bergmann & Cristia, 2016). Yet, given the fact that half of the sample displayed a preference for rhyming songs, infant age and cognitive/linguistic maturation may play a role as well (Hunter & Ames, 1988). These individual differences in early rhyme sensitivity bear great potential for further research, for example exploring the amount and quality of rhyming input and its relation to future linguistic development.

Two approaches were deployed to analyze the data for this study: t-tests, which are traditionally used to analyze infant looking-time data, and mixed effects models, which have been making their way into the analysis of infant behavioral data (Frank et al., 2020). The mixed-effect model may have been more sensitive to the effect of rhyme on infants' looking behavior, as it simultaneously accounts for variance introduced by individual participants as well as individual songs. Moreover, it better fits the data structure of the current study, with varying number of trials per participant and condition. The t-test, in contrast, aggregates over songs and trials. Accounting for nuisance variance in the data may be generally important in infant looking time experiments, as the effect sizes in such experiments tend to be small (Bergmann, Piccinini, Lewis, Frank, & Cristia, 2018; but see again: Csibra et al., 2016). In the present study, the observed effect size ($d_z = 0.28$) was substantially smaller than the anticipated effect size ($d_z = 0.5$). Moreover, it is conceivable that the song stimuli elicited more variance in the responses of the infants than the typically used spoken stimuli. As a result, a statistical method that accounts for all these sources of variance would be the best suited to detect such subtle effects.

Future studies could improve on the methodology employed in this study in several ways: First, a considerable number of trials had to be excluded from the analysis. The first rhyme necessarily only occurred at the end of the second phrase, requiring infants to listen for an average of six seconds before data collection for a given trial could start. On 44% of trials, infants did not listen long enough to reach this time point in the trial. Using shorter melodic phrases could help avoid this problem. Second, the stimuli were rather complex: The Nonsense-Dutch used for the song lyrics could have hampered infants' processing of the phonological content. Future studies could use the infants' natural native language for the verses and Nonsense-Dutch for the rhyme words only. This way, stimulus creation would remain relatively simple while the lyrics in general would possibly be more accessible for the infants. Finally, a relatively large set of unknown melodies was used for the stimuli, possibly tiring out the infants. Taken together, a simpler approach using a smaller number of songs that consist of shorter melodic phrases and carry natural native language lyrics could be a promising road for future investigations. Other studies using the Headturn Preference Procedure have successfully shown that infants process songs, using song stimuli that have shorter phrases, and with melodies and lyrics that are repeated in several trials of the experiment (Corbeil et al., 2013; Lebedeva & Kuhl, 2010; Thiessen & Saffran, 2009).

Child songs and nursery rhymes have a clear and repetitive structure. Young infants might benefit from this structure when detecting complex phonological patterns. In this study, infants detected rhymes at phrase endings, a phonological pattern that they do not seem to detect when presented within word lists (Jusczyk, Goodman, et al., 1999) unless they receive explicit training (Hayes et al., 2000, 2009). Detecting phonological patterns such as rhymes is of great importance to children over the course of linguistic development. Being aware of the rhyme relationship between words has been associated with various linguistic abilities: rhythmic awareness (Wood, 2006), articulation (Mann & Foy, 2007), phonological perception (Foy & Mann, 2001) and acquisition of literacy in general (Bryant, Maclean, & Bradley, 1990; Goswami, 2002; Wood & Terrell, 1998). Child songs and nursery rhymes might be the earliest contexts in which infants encounter rhymes in their linguistic input. A small number of studies has even shown that early knowledge of nursery rhymes is directly linked to later success in reading and spelling (Bryant et al., 1989; MacLean et al., 1987).

Here we explored infants' ability to extract phonological patterns from songs. Infants are exposed to child songs on a daily basis. Potentially, the songs' inherent coupling of melody and lyrics makes them not only an attractive social stimulus to infants, but also a tool for language learning (Thiessen & Saffran, 2009). In this study, infants overall listened longer to the songs that did not contain a rhyme, suggesting that infants are sensitive to phonological overlap at phrase boundary position in sung input. These results warrant more research to further our understanding of how rhyming songs and other forms of language play support spontaneous language processing abilities in infants.

4

Chapter 4

Infants’ implicit rhyme perception in child songs and its relationship with vocabulary

This chapter is a slightly modified version of:
Hahn, L. E., Benders, T., Fikkert, P., Snijders, T. M., (2021). Infants’ implicit rhyme perception in child songs and its relationship with vocabulary. *Frontiers in Psychology, 1-8*.

Stimuli for this chapter are available at: https://osf.io/gnadw/

Abstract

Rhyme perception is an important predictor for future literacy. Assessing rhyme abilities, however, commonly requires children to make explicit rhyme judgements on single words. Here we explored whether infants already implicitly process rhymes in natural rhyming contexts (child songs) and whether this response correlates with later vocabulary size. In a passive listening ERP study, 10.5-month-old Dutch infants were exposed to rhyming and non-rhyming child songs. Two types of rhyme effects were analysed: (1) ERPs elicited by the first rhyme occurring in each song (rhyme sensitivity) and (2) ERPs elicited by rhymes repeating after the first rhyme in each song (rhyme repetition). Only for the latter a tentative negativity for rhymes from 0 to 200 ms after the onset of the rhyme word was found. This rhyme repetition effect correlated with productive vocabulary at 18 months-old, but not with any other vocabulary measure (perception at 10.5 or 18 months-old). While awaiting future replication, the study indicates precursors of phonological awareness already during infancy and with ecologically valid linguistic stimuli.

Acknowledgements

We would like to thank Katharina Menn and Iris Schmits for their exceptionally professional and motivating support during study preparation, scheduling and data collection. Clementine Pizzey-Grey and Anna Bolhuis are gratefully acknowledged for their help during data collection and Jessica Askamp and Angela Khadar for their support with general lab procedures and scheduling.

Introduction

Being able to judge whether words rhyme is one of the earliest forms of phonological awareness children develop (Ziegler & Goswami, 2005). Rhyme awareness reflects children's metalinguistic ability to separate the initial phoneme of a syllable from the rest and is usually assessed with a rhyme judgment task, in which children decide whether a set of words rhymes (e.g., *king/ring*) or not (e.g., *king/pear*). Measured from around four years of age (Bradley & Bryant, 1983; Bryant & Bradley, 1978), rhyme awareness, among other phonological awareness measures, serves as a standard predictor for future literacy (Melby-Lervåg, Lyster, & Hulme, 2012; Wood & Terrell, 1998).

Rhymes already play a role in the developing lexicons of much younger children. Many frequent words in children's input rhyme (e.g., *hit, bit, pit*) (De Cara & Goswami, 2002), presumably resulting in the first words in infants lexicons to also rhyme, rather than to share phonemes at the word onset (e.g., *pin/bin* vs. *pin/pit*, respectively) (Zamuner, 2009). It has been suggested that the necessity to differentiate and recognize such similar-sounding words causes children's initially holistic lexical representations to be re-structured into segmental-level specifications (Fikkert & Levelt, 2008; Metsala & Walley, 1998). Consequently, a growing lexicon might lead to an increase in phonological awareness (Carroll, Snowling, Stevenson, & Hulme, 2003; Metsala & Walley, 1998) and this relationship is probably reciprocal: infants' developing phonological sensitivity also enables a further growth in the number of lexical representations (Curtin & Zamuner, 2014). In fact, there is a well-established association between pre-school phonological awareness and vocabulary size (e.g. Lonigan et al., 2000; Stadler et al., 2007), where phonological awareness is mainly assessed with explicit rhyme judgment tasks and has been attributed to a need to differentiate words with offset-overlap in growing lexicons (De Cara & Goswami, 2002, 2003; Metsala & Walley, 1998).

Despite the evidence regarding the role of rhyming words in infants' lexicons, previous studies have been equivocal concerning infants' ability to recognize rhymes. This is mainly due to differences in test procedures and stimuli: while infants respond to a change from one single-word rhyme pattern to another in the Conditioned-Headturn-Procedure (R. A. Hayes et al., 2000, 2009), 9-month-olds may only display spontaneous differentiation between rhyming and non-rhyming child songs (chapter 3), but not between rhyming and non-rhyming word lists (Jusczyk, Goodman, et al., 1999). The tentative evidence in favour of infants' rhyme detection in Hayes et al. (2000, 2009) and chapter 3 suggests that infants' rhyme detection might be subtle. An implicit paradigm, like passive-listening EEG, might therefore be more sensitive to infants' emerging rhyme abilities (Kooijman et al., 2008).

The results in chapter 3 suggest that infants are able to recognize recurring rhymes in their natural linguistic context. Caregivers rhyme and sing for their infants on a daily basis (Ilari, 2005) and it has been suggested that at-home musical and language play contributes to vocabulary growth (Franco et al., 2021) and emergent literacy (Krijnen, van Steensel, Meeuwisse, Jongerling, & Severiens, 2020; Politimou et al., 2019). Specifically, toddlers' experience with nursery rhymes was associated with several phonological abilities in a number of studies (Bradley & Bryant, 1983; Bryant & Bradley, 1978; Dunst et al., 2011) and caregiver singing during infancy positively influences later vocabulary size (Franco et al., 2021). Potentially, the acoustic shape of songs creates a perceptual boost for infants, due to rhymes in songs being placed at a salient phrase-final position and in a predictable rhythmic context (Kotz & Schwartze, 2010). In the current study, we employ a passive-listening EEG paradigm to answer the following research questions: 1) Do infants process rhyming songs differently from non-rhyming songs? and 2) Is the ability to detect recurring rhymes in songs related to later vocabulary?

The ERP literature on rhyme processing contains a range of effects, including the classic N450 rhyme effect, a negativity for non-rhymes at posterior electrodes elicited during explicit rhyme judgement tasks (Rugg, 1984a,b). There is no consensus whether pre-literate children already show this effect (Andersson, Sanders, Coch, Karns, & Neville, 2018; Wagensveld, van Alphen, Segers, Hagoort, & Verhoeven, 2013), but one study observing the N450 at this young age found that it correlated with phonological awareness (Andersson et al., 2018). An anterior negativity for rhyming pseudo-words has been reported for 4-year-olds in the absence of a rhyme task (Andersson et al., 2018), suggesting that the anterior negativity reflects implicit automatic rhyme processing. Note, however, that preschoolers' executing rhyme judgements also displayed an early anterior negativity, which reduced in amplitude with increased letter knowledge (Wagensveld et al., 2013). In the present study, we expect an early negativity for rhymes, most likely at anterior electrode sites, as the infants will not be executing a task and still have limited phonological awareness.

The infant (EEG) research tradition has not yet assessed infants' rhyme abilities, but laid the foundation by providing evidence for infants' ability to detect repeated phonemes and words in speech and associating this detection with vocabulary size (see Cristia et al., 2014 for a review). Specifically, the ERP word familiarity effect usually occurs between 200-500 ms after word onset as a left anterior negativity for the familiar word, which becomes more negative with each repetition. The effect occurs in response to several repetitions (e.g. Kooijman et al., 2005), but also after a single repetition of the same word (e.g. E. Kidd, Junge, Spokes, Morrison, &

Cutler, 2018) in continuous speech, and is influenced by stimulus features and task difficulty (Junge et al., 2012; Mills et al., 2004; Snijders et al., 2020). For example, Snijders, Benders and Fikkert (2020) observed a positive word familiarity effect for words occurring in child songs. The word familiarity effect has been established as a reflection of infants' ability to recognize repeating words in speech (see e.g. Teixidó, François, Bosch, & Männel, 2018 for review). Individual differences in the polarity of the ERP word familiarity effect are associated with vocabulary size (Junge & Cutler, 2014; E. Kidd et al., 2018; Kooijman et al., 2013): Infants with more negative word familiarity effects tend to have larger vocabularies.

The current study builds on the discussed EEG research on rhyme processing with adults and children in combination with the infant word familiarity effect to ask whether infants detect rhymes in songs and whether individual differences in this ability are related to infant vocabularies. We specifically aim to extend the word familiarity effect to another phonological unit: rhymes. We presented 10.5-month-old Dutch infants with child songs of 10 phrases long from chapter 3 in a rhyming and non-rhyming version, which only differed in the final pseudo-word at the end of every phrase (e.g.: *paf, taf, kaf* vs. *teet, deus, bag*).

Two types of effect will be investigated*: rhyme sensitivity* and *rhyme repetition*. Rhyme sensitivity will be measured at the first point of diversion between rhyming and non-rhyming songs, i.e., at the end of the second phrase where the rhyme is repeated for the first time in rhyming songs but not repeated in non-rhyming songs. Measuring rhyme sensitivity corresponds to earlier studies where critical words (Junge et al., 2012; E. Kidd et al., 2018) or rhymes (Andersson et al., 2018) were repeated only once. The rhyme repetition effect will be measured as the averaged response to rhymes occurring at the end of the 3rd through 10th phrase of the songs. Measuring responses to repeated rhymes is comparable to the ERP word familiarity response to words repeated across successive sentences (Junge et al., 2012; Kooijman et al., 2005). For both effects, we expect a left anterior negativity for rhymes, which might be more pronounced in the rhyme repetition effect due to repetition enhancement (Nordt, Hoehl, & Weigelt, 2016). The onset of the ERP effects is probably slightly later than the 200 ms reported in earlier word familiarity studies (Teixidó et al., 2018), due to the beginning of the phonological overlap being shifted to after the onset phoneme. Finally, both ERP effects will be correlated with Dutch CDI scores (Zink & Lejaegere, 2002) for productive and receptive vocabulary at 10.5 and 18 months, to investigate a possible link between rhyme perception and vocabulary size.

Method

This study was approved by Ethical Board of the Faculty of Social Sciences, Radboud University Nijmegen, CMO 2012/012 and parents of participating infants gave informed consent prior to data collection in accordance with the Declaration of Helsinki.

Participants

In total 40 10.5-month-old infants from monolingual Dutch households were tested, from which 12 infants were excluded from data analysis due to not contributing at least 30 trials each for rhyming and non-rhyming songs (5 infants) or due to having more than two neighbouring noisy channels (7 infants). Twenty-eight datasets were used for the rhyme repetition effect (see analysis section below): mean age: 320 days, range: 304-338, 13 girls. From this subsample, 18 infants contributed enough trials to be analysed for the rhyme sensitivity effect (see analysis section below). Sample size was determined prior to data collection based on similar studies by Kooijman and colleagues (2005, 2013) and Junge and colleagues (2014, 2012).

Stimuli

Song stimuli were taken from Hahn and colleagues (2018) and comprised of nine novel songs that were unknown to the infants. An example stimulus is depicted in Figure 4.1. See the original publication for more detail on creation and acoustic characteristics of the songs and a link to the song stimuli.

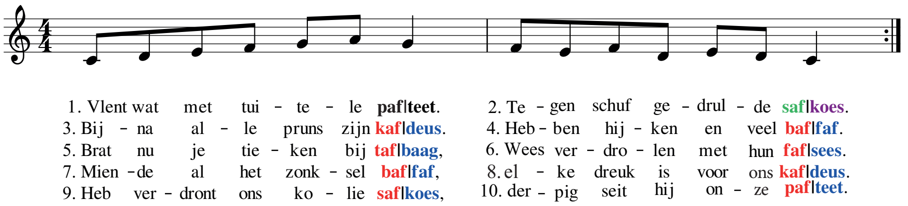


Figure 4.1 Example Stimulus. Song with lyrics in non-sense Dutch. Critical pseudo-words used to assess the rhyme sensitivity effect at the end of phrase 2 of the songs highlighted in green (rhyme) and purple (non-rhyme); Critical pseudo-words to assess the rhyme repetition effect from phrase 3 to 10 of the songs are highlighted in red (rhyme) and blue (non-rhyme). Stimuli were taken from chapter 3. For the analysis, the first phrase-final pseudo-word of every song (in bold) was disregarded, as this cannot rhyme.

Each song occurred in two versions (rhyming and non-rhyming) and consisted of ten phrases (five verses consisting of two phrases each). Song lyrics were in non-sense Dutch to control for word familiarity: content words were replaced with legal

Dutch pseudo-words, while function words were regular Dutch function words. Critical words at the end of every phrase comprised of CVC pseudo-words. Rimes of CVC pseudo-words were of medium frequency and all occurred as word ends in the Dutch Celex (Baayen et al., 1995), while the full CVC pseudo-words (including the initial consonant) did not appear in Celex or had a very low frequency (<100 raw). The full songs were on average 33 sec long (range rhyming: 28-36 sec, range non-rhyming: 28-35 sec), including 500 ms between consecutive phrases of every song. The ten single phrases of each song were on average 2.7 sec long (range rhyming: 2.2-3.3 sec, range non-rhyming: 2.2-3.2 sec). Critical pseudo-words were on average 690 ms long (ranges rhyming: 511-1022 ms, range non-rhyming: 489-908 ms). The rhyming and non-rhyming version of a song were identical apart from the final pseudo-word of every phrase (critical pseudo-word). These either rhymed (e.g.: *paf, taf, kaf*) for the rhyming version of a song, or did not rhyme (e.g.: *teet, deus, bag*) for the non-rhyming version.

Procedure

Each test session was run by two experimenters. One experimenter briefed the parents and ran the measurement, while the other entertained the infant during placement of the electrode cap and data collection. The entire testing procedure from arrival to farewell lasted approximately 1 hour. During the test session, the infant sat on a caregiver's lap in an electrically shielded room. Silent baby-friendly movie clips were played at a PC screen in front of the infant. One experimenter sat next to the caregiver to silently entertain the infant during the measurement if necessary. Both caregiver and experimenter listened to masking music over headphones throughout data collection.

Stimulus presentation was controlled by Presentation software (Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com). The nine song stimuli (ten phrases each) in their two versions (rhyming / non-rhyming) were randomized across two experimental blocks. Songs of the same condition (rhyming / non-rhyming) never occurred more than twice in a row. Data collection lasted around 11 minutes.

Parents filled in questionnaires on their musical and demographic background and the vocabulary of their child (N-CDI 1 at 10.5 and N-CDI 2B at 18 months of age). So far, only the vocabulary questionnaires have been analysed.

EEG activity was collected from 32 Ag/AgCl electrodes (ActiCAP) using BrainAmp DC and BrainVision Recorder Software (Brain Products GmbH, Germany). Electrode locations were in accordance with an extended 10/20 system: F7/3/4/8, FC5/1/2/6, C3/4, CP5/1/2/6, P7/3/4/8, Fz, FCz, Cz, CPz, Pz, POz for collection of EEG activity.

Electro-oculogram (EOG) was recorded using an electrode on the left or right cheek and above the eye (Fp1/2) for vertical EOG, and left and right of the eyes (FT9/10) for horizontal EOG. AFz served as Ground, FCz as online reference. Impedance were typically kept below 25 kΩ. Data was collected with a sampling rate of 500 Hz using an online low-cut off filter of 10 sec and high-cut off of 1000 Hz.

Data preprocessing

EEG data were analysed in MATLAB (The MathWorks, Natick, MA, USA) using the Fieldtrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). Data was filtered offline at 0.1-30 Hz. For one infant, a 0.5 Hz high-pass filter was used, due to slow drifts during the measurement. Bad channels were manually removed, as were data segments with flat channels or large artefacts (>150 µV for EEG channels, >250 µV for EOG channels). Eye- and single-electrode noise components were identified using Independent Component Analysis (Makeig, Bell, Jung, & Sejnowski, 1996) as implemented in the EEGlab toolbox (Delorme & Makeig, 2004) with infomax ICA (Bell & Sejnowski, 1995) on 1-sec data snippets.

Critical pseudo-words were the final pseudo-words from phrase 2 to 10 from every song, resulting in 162 possible trials in total (see Figure 4.1 for an example, the rhyming (green/red) and non-rhyming (purple/blue) pseudo-words are the critical pseudo-words). Raw data was epoched from -200 to 900 ms around critical pseudo-word onset, using 0.1-30 Hz filters (0.5 Hz high-pass filter for one infant), and the eye- and noise-components identified with ICA were removed from the data (average of 2 eye and 3 noise components per infant). Critical pseudo-word-epochs were re-referenced to linked-mastoids. For three infants a single mastoid electrode was used as a reference, due to the other reference electrode being noisy. Time-locked data was baseline corrected by normalizing waveforms relative to the 200 ms epoch preceding the onset of the critical pseudo-word. Trials containing activity exceeding +/-150 µV were removed, leading to exclusion of five infants who did not contribute a minimum of 30 of the 81 possible trials for rhyming and non-rhyming songs. Nine channels were not analysed due to being noisy in too many datasets: F7/F8, F3/F2, Fz, T7/8, P7/8. Further noisy/missing channels were repaired using spline interpolation and a custom neighbourhood structure (a total of 17 channels repaired in 16 infants). Seven infants were excluded from further data analysis due to having more than two neighbouring noisy channels, making channel repair unreliable. Event-related potentials were computed for the remaining 13 channels (FC5/6, FCz, C3/4, Cz, CP5/6, CP1/2, P3/4, Pz) by averaging the rhyming and non-rhyming trials.

Statistical analysis

Two ERP effects were investigated: 1) the *rhyme sensitivity effect*, only on ERPs from phrase 2 of the songs, the moment where rhyming and non-rhyming songs first differed (green/purple in example Figure 4.1) (minimum of 5 trials per condition for this analysis, *Mean (SD)* number of trials rhyming: 7 *(1.09)*, and non-rhyming trials: 7 *(1.47)*, n = 18 infants) and 2) the *rhyme repetition effect*, averaged over ERPs from phrase 3 to 10 of every song (red/blue in example Figure 4.1, n = 28 infants, *Mean (SD)* number of trials rhyming: 48 *(11.43)*, and non-rhyming trials: 48 *(10.39)*). Non-parametric cluster-based permutation tests (Maris & Oostenveld, 2007) were used to evaluate differences in the ERPs between the rhyming and non-rhyming conditions. For these tests, first dependent-samples t-tests are calculated to compare rhyming and non-rhyming conditions (for all 13 remaining electrodes and all time-points between 0 to 900 ms after onset of the critical pseudo-word). Then, clusters are made of neighbouring electrodes and time-points that exceed a threshold alpha of 0.05 (uncorrected). A cluster-level statistic (sum of t-statistics in the cluster) is then computed and, using Monte-Carlo resampling (1000 permutations), a reference distribution is made for random data, to which the observed cluster-statistic is compared to get a Monte Carlo p-value. This effectively controls for multiple comparisons while taking the electrophysiological properties of EEG into account (Luck & Gaspelin, 2017; Maris & Oostenveld, 2007; Sassenhagen & Draschkow, 2019)

From the N-CDI questionnaire for each infant the following scores were derived: comprehension at 10.5 months of age and production and comprehension at 18 months of age. To adhere to previous research (e.g., Kidd et al., 2018), non-parametric Kendall's Tau rank correlations were calculated to investigate the relationship between ERPs (mean of cluster electrodes and cluster time-points of identified clusters in cluster-based permutation test comparing rhyming and non-rhyming conditions) and the vocabulary scores (not normally distributed).

Results

Rhyme sensitivity effect

In response to the first occurrence of the rhyme/non-rhyme (phrase 2 of each song), rhyming pseudo-words induced a more positive ERP waveform compared to non-rhyming pseudo-words (Figure 4.2A). This difference was not significant (lowest cluster *p* = .6).

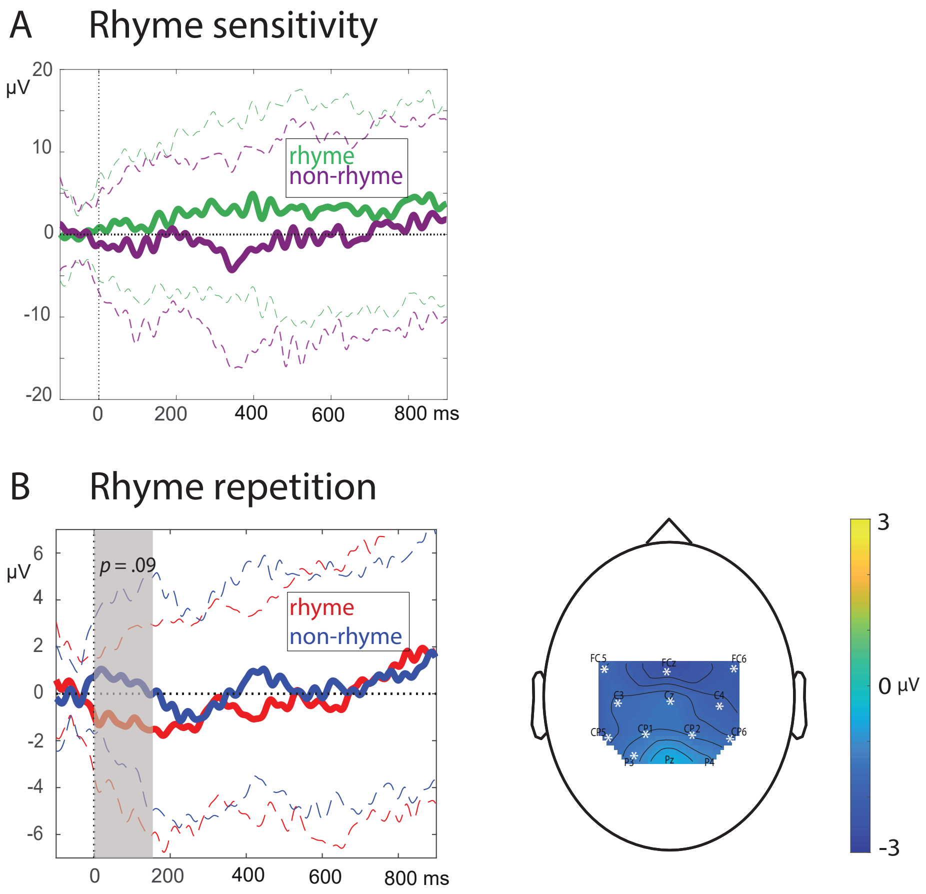


Figure 4.2 Rhyme sensitivity effect (based on the final pseudo-word of phrase 2 of every song, n= 18 infants) and Rhyme repetition effect (based on the final pseudo-words of phrase 3 to 10 of the songs, n =28 infants). Solid lines = Means; dotted lines = +/- 1 SD; A: ERPs, averaged over all 13 electrodes; B left: ERPs averaged over the electrodes in the largest identified cluster (0-178 ms, p = .09), right: topographic isovoltage maps of the difference between rhyme and non-rhyme within the rhyme repetition effect cluster time window (0–178 ms) for all tested electrodes, cluster electrodes are marked with white star.

Rhyme repetition effect

ERPs for rhyming pseudo-words occurring at the end of phrase 3 to 10 of the songs were more negative than ERPs for non-rhyming pseudo-words within the first 200 ms after pseudo-word onset (Figure 4.2B). None of the identified clusters in the cluster-randomization test of the 0-900 ms time-window survived multiple comparisons correction (lowest cluster *p* = .09). The cluster with the lowest p-value ranged from 0-178 ms and contained all electrodes except for Pz and P4: *Mean (SD)* rhyme: -1.27 (2.55), *Mean (SD)* non-rhyme = 0.48 (2.57).

Individual differences in the *rhyme repetition effect* within the largest identified cluster (mean of all cluster electrodes within the 0-178 ms time window) were significantly correlated with productive vocabulary at 18 months (τ = -0.3, *p* = 0.03). Infants with a larger negative ERP difference (rhyme more negative than non-rhyme) produced more words at 18 months old (Figure 4.3). There were no correlations with comprehension at 10.5 or 18 months.

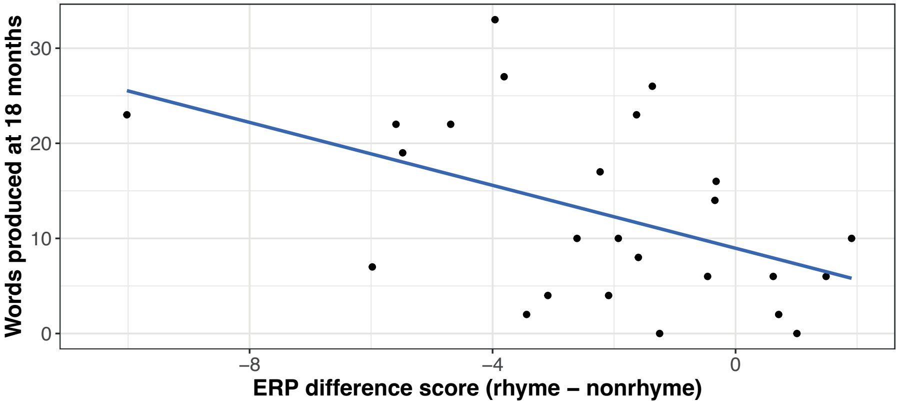


Figure 4.3 Correlation between words produced at 18 months (y-axis) and ERP difference scores (rhyme – non-rhyme, averaged for the cluster electrodes and time window) on the x-axis: τ = -0.3, p = 0.03. The correlation remains marginally significant when excluding the outlier at (-10, 23): τ = -0.3, p = 0.07.

Discussion

Explicit phonological awareness during preschool years is an important predictor for literacy (Melby-Lervåg et al., 2012; Wood & Terrell, 1998). Potentially, infants' perception of rhymes during informal language play contributes to emerging phonological awareness skills and vocabulary (Politimou et al., 2019; Krijnen et al., 2020; Franco et al., 2021). Previous studies were equivocal concerning infants' rhyme abilities (Hayes et al., 2000, 2006; Hahn et al., 2018; Jusczyk et al., 1999), presumably due to behavioural paradigms concealing infants' subtle processing abilities (Kooijman et al., 2008). The current study employed a passive listening EEG paradigm with 10.5-month-old Dutch infants to explore whether infants differentiate rhyming from non-rhyming pseudo-words in child songs in the absence of an explicit behavioural task. Infants' response to the first rhymes occurring in each song (at the end of the second phrase, *rhyme sensitivity effect*) did not render significant results. For rhymes repeated throughout the song (at the end of phrase 3 to 10, *rhyme repetition effect*) an early negative effect for rhymes was found that approached conventional significance levels and correlated with productive vocabularies at 18 months.

For the rhyme sensitivity effect there was a numeric indication for rhyming pseudo-words to elicit a sustained positivity (Figure 4.2A), but the effect did not survive multiple comparison correction. The inconclusive results might be attributable to a lack of statistical power, as data from only 18 infants was available for this analysis, with each infant providing on average only seven trials per condition.

Repeated rhymes occurring at the end of phrase three to ten of the songs elicited a central negativity within the first 200 ms after pseudo-word onset, which was marginally significant when corrected for multiple comparisons. Consequently, this effect delivers tentative evidence for an implicit neural rhyme response in infants. The early negative effect for rhyming pseudo-words is similar to the early negative rhyme effect that was identified by Wagensveld et al. (2013) for single words in pre-literate 5-year-olds, but not 7-year-olds. Later negative effects for rhyming words have also been identified in 3–5-year-olds (Anderson et al., 2018). For complete words repeated in continuous speech, both negative and positive word familiarity effects have been found (see Teixidó et al., 2018). A previous study that used child songs, reported a positive word familiarity effect (Snijders et al., 2020). The opposite polarity in the current study might be attributable to rhymes occurring consistently at the end of the song phrases. The fixed rhyme position might have lessened working memory load in comparison to Snijders et al. (2020), where critical words were occurring at various phrase positions (see Snijders et al., 2020 for more background on polarity differences in the word familiarity effect). In terms of polarity, our negative rhyme effect rather adheres to other studies with recurring spoken single words and rhymes (Teixidó et al., 2018, Andersson et al., 2018; Wagensveld et al., 2013). The rhyme repetition effect observed here occurred immediately after pseudo-word onset, which is different from the word familiarity effect which usually occurs between 200 and 500 ms after word onset (Teixidó et al., 2018), but again similar to the early negative rhyme effect in 5-year-olds identified by Wagensveld et al. (2013). The latency and polarity of our effect have to be interpreted with caution, due to the effect being the result of an average over the 3rd through 10th phrase of the songs. Future studies, with more trials available per phrase of the songs, should investigate whether the effect changes (gradually) in latency and polarity upon every rhyme repetition (Kooijman et al., 2013; Nordt et al., 2016).

Infants with more negative rhyme repetition effects at 10.5 months had larger productive vocabularies at 18 months. This is the first study to report a relationship between infants' vocabularies and their perception of repeating rhymes in songs.

This finding extends previous studies, which established such relationship based on the detection of phonemes and words in fluent speech (see Cristia et al., 2014 for a review).

The functional relevance of rhyme sensitivity for infant development requires further research. Infants might experience no communicative pressure to utilize their implicit knowledge about the syllabic units of onsets and rhymes, due to their small lexicons not yet containing many rhyming words (Johnson, 2016). Rhymes in songs, however, are placed within a particularly intriguing stimulus that is highly ritualized, repetitive, rich in structural cues and progressing at a rather slow pace (Falk & Kello, 2017; Longhi, 2009; Trainor, 1996; Trehub et al., 1997; Trehub, Unyk, & Trainor, 1993b). The acoustic context of language play might provide infants with a chance to recognize the syllabic structure of rhyming words, while this might be much more difficult in ordinary speech. So far, there is mounting evidence for a relationship between processing and production of spoken nursery rhymes and literacy and phonological awareness skills in pre-schoolers (see Dunst et al., 2011 for review). Based on the current study, songs and nursery rhymes might have an impact on phonological processing and vocabulary already during infancy (see also Franco et al., 2021).

The tentative relationship between implicit rhyme processing and vocabulary observed in the current study requires future replication. Only productive vocabulary, and not word comprehension, was related to early rhyme abilities. This might be due to a more reliable parental estimate of productive vocabulary. Another tentative explanation would be the use of prediction in production (van Alphen, Brouwer, Davids, Dijkstra, & Fikkert, 2021), which might also have impacted processing of our predictable rhyming stimuli.

The limited number of trials, electrodes and infants in our sample makes future replication and extension of the study necessary. Future research should also settle to what extent the rhyme effect reported here differs from the ERP word familiarity effect. Both effects are elicited by repeating phonological material in infants' input and could thus stem from the same underlying auditory processing mechanism. One possible interpretation is that the rhyme effect we identify in the current study is just a word familiarity effect that appears early due to the predictability of the appearance of the rhymes at phrase ends. Solving this issue is impeded by the exclusion of anterior electrodes in the current study, the standard location of measuring the word familiarity effect. An alternative interpretation would be that the effects differ, and possibly depend on the perceived lexicality of the repeated stimuli. The ERP word familiarity effect has mainly been reported for existing words, while the rhyme negativity in pre-schoolers can be elicited by existing as well as pseudo-words (Andersson et al., 2018). Additionally, the ERP word familiarity effect arises from repetition of full words, while the rhyme negativity is based on repetition of syllable nucleus and coda only, and no other kinds of phonological overlap (Wagensveld et al., 2013). Whether infants in the current study recognize the change in syllable onsets between successive rhyming pseudo-words (eg., *paf, taf, kaf*) or rather consider them repetitions of the same pseudo-word and ignore onset differences (e.g., *paf, paf, paf*) is another question that remains for future research (see also Ngon et al. (2013)).

The current study complements and extends previous behavioural results (chapter 3) with an ERP response for 10.5-month-old infants' early implicit rhyme detection in a natural rhyming stimulus and a relationship of this ERP response with productive vocabulary at 18 months of age. Implicit rhyme detection in language play might contribute to the development of explicit phonological awareness abilities and vocabulary (Krijnen et al., 2020; Franco et al., 2021). The current early evidence of phonological awareness in infancy might thus reflect the origin of a key predictor of reading achievement (Wood & Terrel, 1998; Ziegler & Goswami, 2005; Melby-Lervåg et al., 2012).

5

Chapter 5

Rhyme perception in poetry – electrophysiological effects of rhyme in varying metrical contexts

This chapter is a modified version of Stroo, M., 2020. Investigating auditory processing of nursery rhymes: An EEG study of rhyme and meter and their relationship to musical aptitude. *Master thesis submitted to the study programme Cognitive Neuroscience at Radboud University*, Nijmegen, the Netherlands.

Stimuli for this chapter are available at: https://osf.io/87b6y/

Abstract

In poetic language, line-final rhymes typically occur in a metrical context. Here we explored whether rhyme perception by adult listeners is altered by surrounding metered versus non-metered verse-lines. Participants passively listened to Dutch nursery rhymes of four lines each, with rhyming or non-rhyming pseudowords at the end of each line. Meter was manipulated by changing the pattern of stressed and unstressed syllables per line. Rhyming nursery rhymes elicited an anterior negativity around 250 ms at the end of line 2 and a posterior positivity from around 400 ms onwards at the end of lines 2 and 4. The posterior rhyme effect represents a delayed phonological N400 that is robust to lexicality and task effects. The occurrence at both line 2 and line 4 suggests the effect to be a strong index of subconscious processing of phonological repetition. The anterior rhyme effect at the end of line 2 only was interpreted as a reflection of the working memory operation of associating rhyming verse lines with each other. Regular meter resulted in sustained positive ERP deflections at the end of both line 2 and line 4, but had only a subtle influence on rhyme perception. The anterior rhyme effect set in slightly earlier in non-metered nursery rhymes, possibly reflecting more effortful processing of non-metered nursery rhymes. The study contributes ERP correlates of rhyme processing in a task-free poetic context, furthering our understanding of the role of poetic devices like rhyme and meter in natural language processing.

Acknowledgements

We thank Margret van Beuningen and Bob Rosbach for their administrative and practical support during testing.

Introduction

The role of rhyme and meter in language processing

Rhymes in poetry typically occur in a rhythmic context, as exemplified in this beginning verse of the famous children's book "The Cat in the Hat" (Seuss, 1957): (stressed syllables in CAPitals) "The SUN did not SHINE, it was TOO wet to PLAY. So we SAT in the HOUSE, all that COLD, cold, wet DAY." The final words of the second and fourth line here rhyme (*play - day*), i.e. they overlap in syllable nucleus and coda. Moreover, stress-feed here are mainly organized into anapestic tetra-meter, i.e., two unstressed syllables precede a stressed syllable. The poetic devices rhyme and meter originate from oral traditions from thousands of years ago and until today the combination of rhyme and meter in poetry elicit a sense of beauty and pleasure in the hearer (Fabb, 2014; Menninghaus, Bohrn, Altmann, Lubrich, & Jacobs, 2014; Obermeier et al., 2016, 2013; Reber, Schwarz, & Winkielman, 2004; Schön et al., 2010; Wassiliwizky, Koelsch, Wagner, Jacobsen, & Menninghaus, 2017) and are perceived as emotionally rewarding (Wassiliwizky et al., 2017).

The aesthetic and emotional appeal of rhyme and meter might be grounded in their facilitating effect on language processing and working memory (e.g., Fabb, 2014; Obermeier et al., 2013). For example, rhyming texts are easier to remember in the long term than prosaic texts (Lea, Rapp, Elfenbein, Mitchel, & Romine, 2008; Rubin, 1995; Tillmann & Dowling, 2007), rhyme enhances working memory during language-related tasks (Chow, Macnamara, & Conway, 2016; Gupta, Lipinski, & Aktunc, 2005; Lindstromberg & Boers, 2008) and verse lines in poetry are considered salient chunks for working memory (Fabb, 2014). Target words in priming studies are processed faster and more easily once preceded by a rhyming prime (e.g., Coch, Hart, & Mitra, 2008; Davids, van den Brink, van Turennout, & Verhoeven, 2011; Rugg, 1984), due to engagement of working memory (Baddeley, Lewis, & Vallar, 1984). Regular meter, like rhyme, also enhances text memorability (Menninghaus et al., 2014), and facilitates lexico-semantic integration (Rothermich & Kotz, 2013; Rothermich, Schmidt-Kassow, & Kotz, 2012) and syntactic processing (Schmidt-Kassow & Kotz, 2009). Moreover, listeners process regularly metered speech faster and with greater ease than non-metered speech (Beier & Ferreira, 2018).

Rhyme and meter are usually assumed to enhance speech processing, though some studies offer opposing views. For example, although rhyme and meter in speech are generally assumed to facilitate prosodic processing and enhance memorability, they may simultaneously increase demands on semantic processing (Menninghaus et al., 2015b; Wallot & Menninghaus, 2018). In particular, processing costs are increased by constraints posed by rhyme and metrical regularity on word choice and word order, sometimes rendering rhyming and metered sentences less natural.

The neural correlates of poetic rhyme perception are so far poorly understood. Especially the complex interplay of task effects, differences in rhyme paradigms, and the use of pseudo as well as lexical rhyme words make it difficult to derive strong conclusions from previous studies. The current study aims at improving our understanding of poetic rhyme processing by providing adult listeners with more natural auditory rhymes in less demanding listening conditions.

Electrophysiology of rhyme perception

Current knowledge about the electrophysiology of rhyme perception mainly stems from single word priming studies, where participants judge a prime word and following target to be either rhyming or not. Such paradigms leave open how rhyme is processed in the natural rhythmic and rhyming context of poetry. The lack of more natural rhyme paradigms might be attributable to the primarily clinical interest of ERP rhyme research (Coch, Grossi, Coffey-Corina, Holcomb, & Neville, 2002; Grossi, Coch, Coffey-Corina, Holcomb, & Neville, 2001; Wagensveld, Segers, Alphen, Hagoort, & Verhoeven, 2012): rhyme awareness is a reliable predictor for literacy (Wood & Terrell, 1998) and is standardly assessed by asking explicit judgments about single words (e.g., *king – ring*). ERP rhyme effects in single word priming studies are also grounded in research on the neural correlates of phonological working memory (Rugg, 1984a, b), as both prime and target need to be segmented into their phonological constituents, retrieved from memory and matched against each other (Grossi et al., 2001). Phonological working memory is another important prerequisite for the development of reading and writing (Gathercole, 1999). The current study will move beyond single word rhyme paradigms in order to gain a better understanding of the interplay of rhyme and meter in implicit auditory rhyme perception in poetry. In this domain of spoken language, phonological working memory plays a role as well, yet the underlying ERP effects might be different due to the rhyme-licensing context of poetry (Fabb, 2014).

The N400 rhyme effect (sometimes also referred to as N450) originates from single word priming studies (Coch et al., 2008; Grossi et al., 2001; Rugg, 1984a, 1984b; Rugg & Barrett, 1987) and denotes a negativity for non-rhyming compared to rhyming target words, most pronounced bilaterally at posterior electrodes. The effect is modified as a response to auditory versus visual rhymes (Coch et al., 2002; Davids et al., 2011; Praamstra, Meyer, & Levelt, 1994) and words versus pseudowords (Coch, Grossi, Skendzel, & Neville, 2005; Davids et al., 2011; Rugg, 1984a). Similar ERP effects have been obtained for other types of phonological overlap (Dumay et al., 2001; Praamstra et al., 1994; Radeau, Besson, Fonteneau, & Castro, 1998; Wagensveld et al., 2012), but might be modulated by lexicality of target items and task demands (see discussions in Wagensveld et al., 2012; Wagensveld, van Alphen, et al., 2013). The N400 rhyme response develops during pre-school years (Andersson et al., 2018; Coch et al., 2005) and its specific association with reading and phonological awareness is still a matter of debate (Coch et al., 2005, 2008; Desroches, Newman, Robertson, & Joanisse, 2013; Noordenbos, Segers, Wagensveld, & Verhoeven, 2013).

Several studies also report anterior rhyme effects with reversed polarity but similar time course as the posterior N400 auditory rhyme effect (e.g., Coch et al., 2002; Cross & Fujioka, 2019; Davids et al., 2011; Grossi et al., 2001), i.e. rhymes giving a more negative ERP response over frontal electrodes than non-rhymes. The overlap in timing between the anterior and posterior rhyme effects might merely indicate a switch in polarity from posterior to anterior scalp regions. Yet, given that the amplitudes of the posterior and anterior rhyme effects are not correlated (Coch et al., 2002) and that there are more individual differences in the anterior than in the posterior rhyme effect (Cross & Fujioka, 2019; Davids et al., 2011; Mohan & Weber, 2015), the anterior rhyme effect might as well be a functionally different ERP component (Coch et al., 2002; Cross & Fujioka, 2019; Davids et al., 2011). Source localization indeed suggests different neural generators underlying the anterior and posterior rhyme effects (Khateb et al., 2007), but their specific functional role still remains to be clarified (Davids et al., 2011).

Both the anterior and the posterior ERP rhyme effects are modulated by task effects. While the posterior N400 rhyme effect is commonly found when participants perform rhyme judgements, it disappears during semantic or lexical decision tasks (Perrin & García-Larrea, 2003; Praamstra et al., 1994; Praamstra & Stegeman, 1993), during melodic tasks (Yoncheva, Maurer, Zevin, & McCandliss, 2013), and during passive listening (Davids et al., 2011; Perrin & García-Larrea, 2003). Lexical status of the stimuli might also be of relevance, as delayed N400 rhyme effects were reported for rhyme judgments on pseudoword pairs (Coch et al., 2005; Davids et al., 2011; Perrin & García-Larrea, 2003) and illegal non-words (Praamstra 1993). In the study by Praamstra and Steegeman (1993), the N400 rhyme effect was observed for both words and phonotactically illegal non-words during a rhyme judgement task, but only for words during a lexical decision task. The anterior rhyme effect was elicited in adults and pre-schoolers during passive listening (Davids et al., 2011; Andersson et al., 2018), with the pre-schoolers also showing the posterior N400 rhyme effect (Andersson et al., 2018).

The combination of these findings leaves open whether the anterior and/or the posterior ERP rhyme effects reflect implicit rhyme perception and whether they generalize to other rhyming contexts such as poetry. Moreover, the role of active versus passive rhyme processing and the influence of lexical status thereon remain unclear.

Electrophysiology of meter perception

As noted above, a regular meter as found in poetry influences speech perception. The regular alternation of stressed and unstressed syllables creates a scaffold for prosodic expectations and directs listeners' attention to relevant moments in the speech stream (Port, 2003; Quené & Port, 2005). Consequently, metrical speech is processed with greater ease and listeners prefer metered over non-metered poetry (Obermeier et al., 2016, 2013). Regular meter also influences semantic processing, as lexical activation and word comprehension are facilitated in rhythmic contexts (Magne et al., 2007; Marie, Magne, & Besson, 2011; Rothermich & Kotz, 2013; Rothermich et al., 2012).

ERPs to words in metered speech are well studied and are usually more positive than ERPs in non-metered speech around 400 ms after critical word onset. Yet, the associated ERP components for meter processing are more heterogenous than the prototypical N400 rhyme effects (see Breen & Fitzroy (2019) and Magne et al. (2016) for comprehensive summaries). In short, more pronounced N400 and P600 components were observed for target words presented in metrically irregular sentences or words that violated the metrical context (e.g., Bohn et al., 2013; Henrich et al., 2014; Magne et al., 2007; Marie et al., 2011; Roncaglia-Denissen et al., 2013; Rothermich et al., 2012; Schmidt-Kassow & Kotz, 2009) with some studies reporting an N400 followed by a P600 (e.g., Y. Luo & Zhou, 2010; Marie et al., 2011; Rothermich et al., 2012). Moreover, P200 effects (Böcker, Bastiaansen, Vroomen, Brunia, & De Gelder, 1999; Henrich et al., 2014; Marie et al., 2011), LANs/ELANs (Bohn et al., 2013; Kriukova & Mani, 2016; Rothermich, Schmidt-Kassow, Schwartze, & Kotz, 2010) and CNVs (Contingent Negative Variation) were associated with violations of metrical expectancies (Domahs, Wiese, Bornkessel-Schlesewsky, & Schlesewsky, 2008; McCauley, Hestvik, & Vogel, 2013). While most of these studies used active listening paradigms, which required participants to make judgments about metric, semantic or lexical stimulus features, similar effects have been obtained in a passive listening context with word list stimuli (Böcker et al., 1999). Based on these previous studies we expect the metrical context as implemented in poetic speech to influence processing of line-final rhymes in poetry.

Rhyme perception is influenced by poetic meter

So far, only very few ERP studies have assessed rhyme perception in a natural rhyming context such as poetry. Vaughan-Evans et al., (2016), Obermeier et al., (2016) and Chen et al., (2016) observed typical posterior negativities for non-rhymes at the end of (Welsh, German, Chinese) verse lines, which largely resembled the N400 rhyme effect. Obermeier et al. (2016) directly assessed whether rhyme perception is affected by the occurrence of rhymes in a metered versus non-metered verse line. Participants in that study were listening to spoken folk-song stanzas of four lines each, which carried either a rhyming or non-rhyming word at line endings. Additionally, the stanzas were either regularly metered or non-metered. The paradigm required active listening, as participants were judging each stanza on rhythmicity (very irregular – very regular) and liking (very bad – very good). A typical N400 rhyme effect was observed, with ERPs for non-rhyming line-final words being more negative than rhymes. Around 800 ms after word onset, non-rhymes elicited more positive responses than rhymes, termed a P600 component. Crucially, both effects were larger in the metered stanzas. A recent fMRI study with poetic stimuli also reports a rhyme-meter interaction, namely at the superior temporal sulcus and the putamen of the right hemisphere (Hurschler, Liem, Jäncke, & Meyer, 2013; Hurschler, Liem, Oechslin, Stümpfli, & Meyer, 2015). Taken together, these findings indicate that the rhythmic-poetic verse context facilitates rhyme perception.

While these previous studies further our understanding of rhyme perception in naturalistic rhyming contexts by using realistic poetic stimuli, it is currently unclear to what extent these poetic rhyme effects are influenced by semantics and lexicality. During reception of poetry, lexicality together with rhyme and meter contribute to aesthetic liking and emotional appeal (Obermeier et al., 2013). Specifically, the positive emotional effect of rhymes was found to be even larger for pseudoword stanzas compared to real word stanzas. Moreover, regular meter and predictable word stress facilitate lexico-semantic processing (Magne et al., 2007; Marie et al., 2011; Rothermich & Kotz, 2013; Rothermich et al., 2012). Given these complex interaction effects between rhyme, meter and lexicality, it is worthwhile to study purely phonological rhyme perception in metrical poetry in the absence of lexico-semantic confounds (Hurschler et al., 2013; 2015).

The current study

The ERP studies on poetic rhyme perception mentioned above all relied on active paradigms, where participants were executing a judgment task of some kind (rhyme, semantic, rhythmicity or poetic appeal). This raises the question whether rhyme and meter perception in poetry arise from automatic implicit processes or rather depend on attention. This is important to assess, as we normally listen to poetry outside of the lab, without executing a behavioural task. Moreover, investigating ERP rhyme effects in the absence of a task might be of interest for potential clinical and developmental future studies (Davids et al., 2011; Andersson et al., 2018; chapters 4 and 6). In addition, previous studies have provided rhyme and meter effects that were potentially confounded with semantic processing.

Consequently, the current study set out to investigate adults' neural correlates of implicit rhyme processing in a natural rhyming stimulus, in the absence of a task, and using rhyming pseudowords. Specifically, we ask whether the implicit processing of rhyming poetry is reflected in ERPs and whether these are influenced by metrical context. Stimuli will be recordings of prototypical Dutch spoken nursery rhymes of four lines each with a simple AAAA rhyme pattern, specifically designed for an infant version of this study (chapter 6). Nursery rhyme stimuli will be manipulated for the factor rhyme (rhyming / non-rhyming) and meter (metered / non-metered) (Obermeier et al., 2016; Hurschler et al., 2015). Pseudowords will be used as line-final target words, while real words will be used for the rest of each verse line. The use of pseudowords facilitates matching of phonological features across conditions. But more importantly, this unique feature of our paradigm allows to isolate rhyme and meter effects from the influence of lexical processing, while maintaining the rhyme licensing context of relatively natural children's-poetry.

Two types of effects will be investigated (in analogy with chapters 4 and 6): *rhyme sensitivity* and *rhyme repetition*. *Rhyme sensitivity* will be measured at the first point of diversion between rhyming and non-rhyming nursery rhymes, i.e., at the end of the second line where the rhyme is repeated for the first time in rhyming nursery rhymes but not repeated in non-rhyming nursery rhymes. Measuring rhyme sensitivity corresponds to earlier studies that presented rhymes only once in single word priming or poetry. Consequently, we expect a posterior positivity for rhymes, presumably arising later than 400 ms post word onset due to the use of pseudoword stimuli.

The *rhyme repetition* effect on the other hand will be measured as the response to rhymes occurring at the end of line 4 of each nursery rhyme, i.e., the third time that the rhyme occurs in rhyming nursery rhymes. The analysis of responses to repeated rhymes is more of an exploratory nature, due to the absence of previous studies that report results for repeating rhymes. Based on the ERP literature on (pseudo)word repetition (Rugg, Doyle, & Wells, 1995; Snijders, Kooijman, Cutler, & Hagoort, 2007) we expect responses to rhymes to be more positive than responses to non-rhymes. However, we have less clear expectations about the onset and topography of the effect. In addition, the effect of a metrical versus non-metrical verse context on rhyme perception will be investigated. Given the swift build-up of prosodic and phonological expectations during perception of metrical speech (Port, 2003; Quené & Port, 2005), we expect that the metered nursery rhymes will elicit a more pronounced rhyme sensitivity effect (line 2; Obermeier et al., 2016). Potentially, meter effects on rhyme processing are carried through until later in the verse, until the rhyme repetition effect at the end of line 4.

Methods

Participants

Twenty-eight native Dutch speakers (7 male, 18-33 years of age, mean age 23.0 years) took part in this study after signing an informed consent form. Sample size was based on a power analysis of the rhyme-meter interaction reported in Obermeier and colleagues (2016) using GPower (version 3.1) (Faul, Erdfelder, Buchner, & Lang, 2009). Data from 23 participants were included in the final sample. See below for details on exclusion of trials and participants. All participants were right-handed (mean laterality coefficient = 89,7; Oldfield, 1971). Participants had normal or corrected-to-normal vision, no history of neurological disorders or self-reported hearing deficits, and no reading or other language-related disorders. Participants did not have professional experience with poetry or nursery rhymes, but had all been exposed to some poetry or nursery rhymes in their youth.

Materials

Nursery rhyme stimuli. Stimuli were the same as in the infant version of the experiment (chapter 6), 52 nursery rhymes, reminiscent of typical Dutch nursery rhymes in terms of structure and linguistic content (see Table 5.1 for an example stimulus). The entire stimulus set is provided here: https://osf.io/87b6y/. Each nursery rhyme consisted of four lines, each ending in a monosyllabic phonologically legal CVC pseudoword. Nursery rhymes were manipulated for the factors rhyme (rhyming vs. non-rhyming) and meter (metered vs. non-metered), resulting in four different stimulus conditions (metered_rhyming; non-metered_rhyming; metered_non-rhyming; non-metered_non-rhyming) and a stimulus set of 208 unique nursery rhymes (52 original nursery rhymes * 4 experimental conditions).

Table 5.1 Example Stimulus in each of the four conditions.

Metered_Non-rhyming	**Metered_Rhyming**
GiRAFfen Eten VERse FIP.	GiRAFfen Eten VERse BIJP.
Ze KNAgen TELkens GROEne SOS.	Ze KNAgen TELkens GROEne FIJP.
BeHOORlijk KAAL is DUS die LIJG,	BeHOORlijk KAAL is DUS die DIJP,
Na EEN beZOEK van VIJFtien FAAF.	Na EEN beZOEK van VIJFtien KIJP.
Non-metered_Non-rhyming	**Non-metered_Rhyming**
GiRAFfen Eten graag VERse FIP.	GiRAFfen Eten graag VERse BIJP.
Ze KNAgen ook TELkens GROEne SOS.	Ze KNAgen ook TELkens GROEne FIJP.
BeHOORlijk KAAL is DUS nu die LIJG,	BeHOORlijk KAAL is DUS nu die DIJP,
Na TWEE beZOEKen van VIJFtien FAAF.	Na twee beZOEKen van VIJFtien KIJP.
Literal translation (metered)	***Literal translation (non-metered)***
Giraffes eat fresh fip/bijp.	*Giraffes like to eat fresh fip/bijp.*
Again and again they gnaw green sos/fijp.	*Again and again they also gnaw green sos/fijp.*
Pretty bald is thus the lijg/dijp.	*Pretty bald is thus now the lijg/dijp.*
After one visit of fifteen faaf/kijp.	*After two visits of fifteen faaf/kijp.*

Note: Lyrics consist of real Dutch words, line-final words are legal pseudowords. Stressed syllables in CAPitals. Non-metered version created by adding/removing a syllable (counterbalanced). Critical words used to assess the Rhyme Sensitivity effect at the end of line 2 and the Rhyme Repetition effect from line 4 underscored. For the analysis, the line-final word of line 1 was disregarded, as this cannot rhyme.

Rhyme was manipulated by creating a rhyming and a non-rhyming version of each nursery rhyme: placing either rhyming or non-rhyming CVC pseudowords (the *critical word*) at the end of each line of the nursery rhymes (rhyming: *bijp, fijp, dijp, kijp* vs. non-rhyming: *fip, sos, lijk, faaf* in the example stimulus in Table 5.1). The same 208 pseudowords were used across conditions, only in a different order to create rhyming versus non-rhyming versions. Meter was manipulated by varying the number of syllables per nursery rhyme line across conditions: metered nursery rhymes had 8-syllable lines in iambic meter, thus starting each line with an unstressed syllable (Table 6.1). This iambic meter was disrupted in the non-metered versions of each nursery rhyme by adding or removing a single syllable from each line. Thus, the corresponding non-metered version of a nursery rhyme either had 7-syllable lines or 9-syllable lines, counterbalanced across the set of 52 original nursery rhymes. Changing the number of syllables was achieved by adding or removing particles and function words or by replacing content words with ones of similar meaning and leaving the rest of the nursery rhyme intact (the stimulus example in Table 5.1 shows a 9-syllable non-metered version). Within the entire stimulus set, the pre-final word of each line of each nursery rhyme was the same across conditions.

Stimuli were recorded from a female Dutch native speaker (the last author of the study), reading the stimuli in a child-directed manner. Recitation of the nursery rhymes was paced to a visual metronome set at 90 bpm to achieve a regular 1.5 Hz stress rate and 3 Hz syllable rate in the stimulus amplitude envelop (Figure 6.1, chapter 6). Silences were inserted at the beginning (50 ms) and end (30 ms) of every nursery rhyme and silences of 666 ms were inserted between the consecutive lines using Praat (Boersma & Weenink, 2014), resulting in nursery rhymes between 10 and 15 seconds long (mean = 12.7 sec, SD = 987 ms) and critical words of approximately 500 ms long (see Table 6.2 (chapter 6)).

Questionnaires. Prior to data collection, participants filled out a custom made questionnaire about demographic information (age, nationality, native language), handedness (Oldfield, 1971), and familiarity with and appreciation of poetry, nursery rhymes, literature and theatre (not reported here, see Stroo, 2020 for details).

Musical aptitude measures. Before the start of the experiment and during data collection, participants provided several measures of their musical aptitude. All these measures were disregarded here, but see Stroo (2020).

Procedure

Participants were seated in front of a computer screen in a dimly lit and sound-attenuated room and were instructed to listen attentively to the stimuli. The nursery rhymes were presented via headphones with 1300 ms pauses between consecutive nursery rhymes. During stimulus presentation, participants watched a silent nature documentary, to ensure they remained awake. Stimuli were presented in 4 blocks of 52 nursery rhymes each (208 nursery rhymes in total). The order of stimulus presentation was pseudorandom, using 4 different stimulus lists (counterbalanced across participants). Participants would hear two nursery rhymes of the same condition no more than twice successively. After each block of nursery rhymes, participants were tested for their musical aptitude (see Stroo, 2020). An entire experimental session lasted approximately 75 minutes. Participants received a €22.50 gift voucher as a reward for participation and were debriefed about the aim of the experiment upon finishing the session.

EEG data was collected using 32 Ag-AgCl electrodes mounted in a standard electrode cap according to the modified 10-20 system with a left mastoid on-line reference. Vertical/horizontal EOG was measured for eye artefact rejection. Data was sampled at 500 Hz and filtered on-line with a time constant of 10 seconds and a high cut-off at 1000 Hz. Impedances were kept below 10 kΩ.

Analysis

Data preprocessing. EEG data were analysed in MATLAB (version R2018a and upwards, The MathWorks, Natick, MA, USA, 2018) using the Fieldtrip toolbox (Oostenveld et al., 2011). Data from each infant were filtered offline at 0.1 - 30 Hz and were cut in 1-sec data snippets. Those snippets containing flat channels or large artefacts (>150 µV for EEG channels, >250 µV for EOG channels) were removed. Data from three participants were excluded from analysis altogether, due to technical errors during recording. From the remaining 25 participants, eye/noise artefact components were identified in the 1-sec data snippets (average of 3 eye and 3 noise components per participant; range 0 to 8 components) using Independent Component Analysis (Makeig et al., 1996) as implemented in the EEGlab toolbox (Delorme & Makeig, 2004) with infomax ICA (Bell & Sejnowski, 1995). Raw EEG data was subsequently epoched into the critical words (-200 to 900 ms around critical word onset, see Note Table 5.1) using 0.1 - 30 Hz filters. The eye- and noise-components identified with ICA were removed from the critical word-epochs, and the trials were re-referenced to linked mastoids. Baseline correction was applied to the time-locked data with waveforms normalized relative to the 200 ms epoch preceding the onset of the critical word. Remaining artefacts were removed using automatic artefact rejection (cut-off +/-75mV), resulting in exclusion of another two participants who contributed less than 80% of trials in each condition. In the data from the remaining 23 participants, further noisy/missing channels were repaired using spline interpolation and a custom neighbourhood structure (3 channels in 2 participants). Event-related potentials were computed for the remaining 27 channels (Fp1/2, F3/4, F8/7, FC5/6, FC1/2, C3/4, T7/8, CP1/2, CP5/6, P3/4, P7/8, O1/2, Fz, Cz, Pz) by averaging the amplitude time-locked to the onset of the critical words for all four conditions for line 2 and line 4 (maximum of 52 trials per condition and line). Participants on average contributed 204 trials for line 2 (SD = 3.99 trials) and 204 trials for line 4 (SD = 3.48 trials).

ERP data analysis. The ERP analysis first evaluated the effects of Rhyme and Meter, by assessing the difference between rhyme/non-rhyme and meter/non-meter at the critical words (the final words). This was done separately with ERP data from line 2 and 4 of the nursery rhymes. ERP rhyme effects will be termed in accordance with the infant version of this study (chapter 6): the Rhyme Sensitivity effect, concerning ERPs for rhyme/non-rhyme from line 2 (Table 5.1), and the Rhyme Repetition effect, concerning ERPs for rhyme/non-rhyme from line 4 (Table 5.1). To assess whether the effect of rhyme was different within the non-metered and metered condition (interaction effects), the difference score between rhyme and non-rhyme was compared between the two Meter conditions. Non-parametric cluster-based permutation tests (Maris & Oostenveld, 2007) were used to evaluate differences between conditions in the ERP waveform across 27 electrodes from 100 to 900 ms post critical word onset (Monte-Carlo permutation distribution, 1000 permutations as standard, 10000 permutations where p-values were around 0.05). Cluster tests assessed all electrodes and time-points in one test and were not limited to certain electrodes or time-ranges.

Results

Rhyme sensitivity- Line 2

For critical words occurring at the end of line 2 of the nursery rhymes, inspection of grand average waveforms indicated an early interaction between rhyme and meter, followed by a sustained positivity for both rhyme and meter, most notably over posterior and midline electrodes (Figure 5.1).

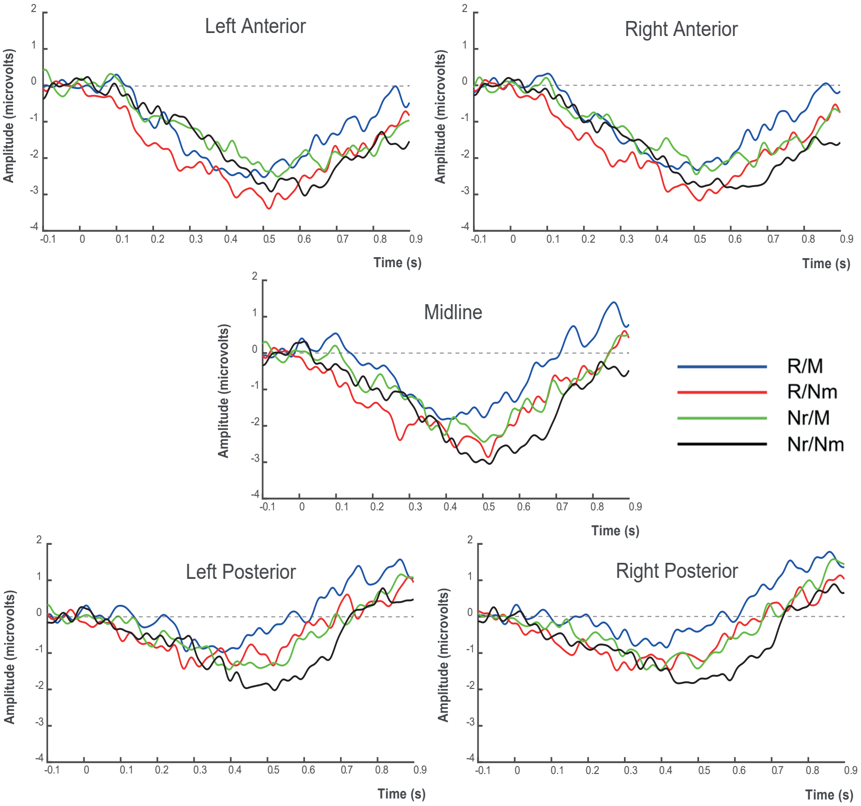


Figure 5.1 Grand averages for critical pseudowords at the end of line 2. R/M = rhyming_metered; R/NM = rhyming_non-metered; Nr/M = non-rhyming_metered; NR/NM = non-rhyming_non-metered. Left anterior = Fp1, F7, F3, FC5, FC1; right anterior = Fp2, F4, F8, FC6, FC2; midline = Fz, Cz and Pz; left posterior = CP5, CP1, P7, P3, O1; right posterior = CP2, CP6, P4, P8, O2.

The cluster test comparing rhyming with non-rhyming nursery rhymes resulted in a cluster from 242 to 452 ms over anterior electrodes (cluster *p* = .042, Figure 5.2A) where rhyme was more negative than non-rhyme and a second cluster from

426 to 900 ms (cluster *p* = .002, Figure 5.2B), where rhyme was more positive than non-rhyme. The second cluster involved all electrodes but was most pronounced at centro-posterior electrodes.

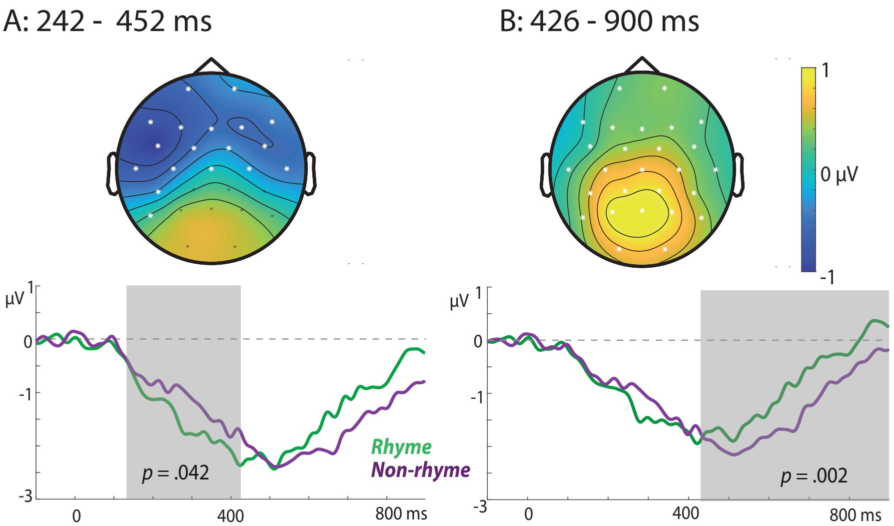


Figure 5.2 Rhyme sensitivity effects, collapsed across metered/non-metered nursery rhymes. A: negative anterior rhyme sensitivity effect (242-452 ms); B: positive centro-posterior rhyme sensitivity effect (426-900 ms); Top: Topographic isovoltage maps of the difference between rhyming and non-rhyming nursery rhymes within the cluster time window; cluster electrodes highlighted with white asterisk. Bottom: ERPs for rhyme and non-rhyme averaged for the cluster electrodes. Shaded areas indicate cluster time windows. All ERPs collapsed across metered/non-metered nursery rhymes.

The cluster test for metered versus non-metered nursery rhymes resulted in a positive cluster ranging from 408 to 766 ms (cluster *p* = .002) involving all electrodes and a second positive cluster, ranging from 784 to 900 ms, which was marginally significant (cluster *p* = .058) for all electrodes except for F8, F7 and T7, indicating a sustained positivity for metered compared to non-metered nursery rhymes (see Figure 5.1 for all meter effects).

The interaction between rhyme and meter was assessed by comparing the rhyme effect within the metered condition to the rhyme effect within the non-metered condition (i.e., comparing the difference score of rhyme/meter minus non-rhyme/meter with the difference score of rhyme/non-meter minus non-rhyme/non-meter) from 100 to 900 ms across all channels. This comparison resulted in a marginally significant positive cluster, ranging from 128 to 234 ms at fronto-central electrodes (*p* = .068) involving all electrodes except for P7/8, O1/2, P4 and

CP6, Figure 5.3. To further inspect this interaction effect, the rhyme effect within the metered condition and the rhyme effect within the non-metered condition were analysed separately, averaging over time between 128 and 234 ms using all channels. In the non-metered condition, a negativity for rhyme emerged at fronto-central electrodes (cluster *p* = .03; cluster channels: Fp1/2, F7, F3/4, Fz, FC5/6, FC1/2, T7, C3). In the metered nursery rhymes, there was no rhyme effect in the 128-234 ms time-window (no clusters identified).

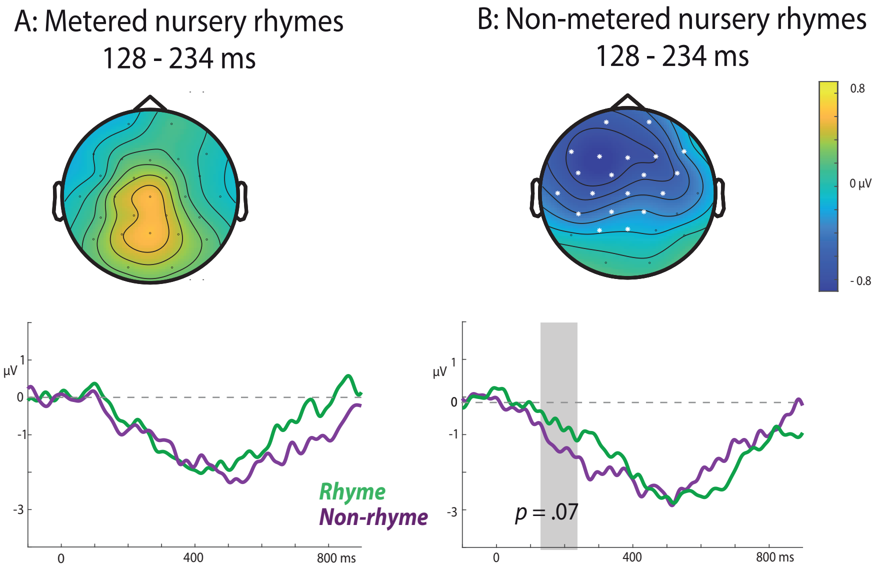


Figure 5.3 Early rhyme sensitivity effect within metered and non-metered nursery rhymes. Rhyme sensitivity effects within (A) metered and (B) non-metered nursery rhymes averaged for the interaction cluster electrodes. The p-value denotes the interaction cluster p-value. Top: topographic isovoltage map of the difference between rhyming and non-rhyming nursery rhymes within metered/non-metered nursery rhymes and the interaction cluster time window (128-234 ms); cluster electrodes highlighted with white asterisk. Bottom: ERPs for rhyme and non-rhyme averaged for the interaction cluster electrodes. Shaded area indicates cluster time window of the interaction effect.

Rhyme repetition - Line 4

Critical words at the end of line 4 of the nursery rhymes elicited differences between the four conditions from around 400 ms post word onset, with a sustained positivity for rhyme, especially pronounced over posterior electrodes, and potentially a later interaction between rhyme and meter over anterior electrodes (Figure 5.4).

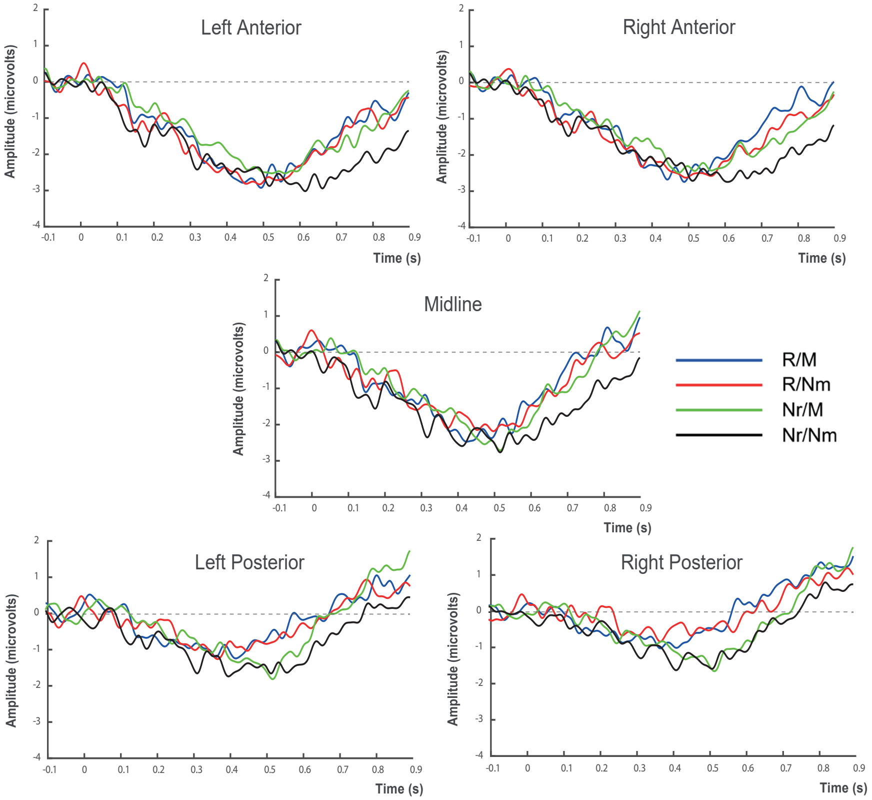


Figure 5.4 Grand average waveforms for the critical pseudowords at the end of line 4. R/M = rhyming_metered; R/NM = rhyming_non-metered; Nr/M = non-rhyming_metered; NR/NM = non-rhyming_non-metered. Left anterior = Fp1, F7, F3, FC5, FC1; right anterior = Fp2, F4, F8, FC6, FC2; midline = Fz, Cz and Pz; left posterior = CP5, CP1, P7, P3, O1; right posterior = CP2, CP6, P4, P8, O2.

The cluster test comparing rhyming with non-rhyming nursery rhymes from 100 to 900 ms across all channels resulted in a significant positive cluster ranging from 386 to 900 ms and involving all channels except for T7 (cluster p = .002, Figure 5.6), indicating a positivity for rhyme that is reminiscent of the late positivity observed for line 2.

Comparing metered and non-metered nursery rhymes from 100 to 900 ms across all channels resulted in a marginally significant positive cluster at all electrodes except for F4, ranging from 790 to 858 ms (cluster p = .064).

Testing for the interaction between rhyme and meter resulted in no significant clusters (lowest cluster p = .1).

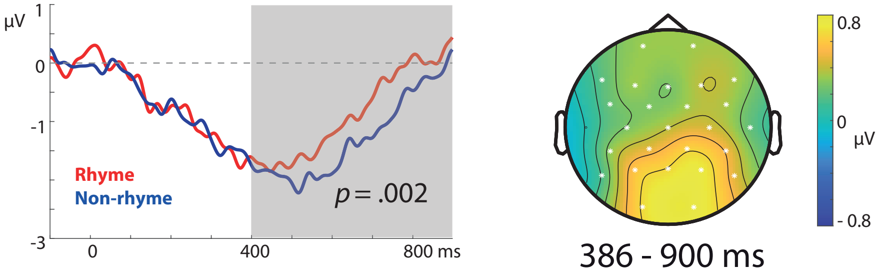


Figure 5.5 Rhyme repetition effect, collapsed across metered/non-metered nursery rhymes. Left: ERP waveforms of rhyming and non-rhyming nursery rhymes within the positive rhyme cluster, collapsed across metered/non-metered nursery rhymes; right: Topographic isovoltage map of the difference between rhyming and non-rhyming nursery rhymes between 386 to 900 ms, significant cluster channels highlighted with white asterisks.

Discussion

Summary of current findings

The current study investigated poetic rhyme perception and the influence of regular meter thereon in adult listeners in the absence of an explicit task. Pseudoword rhymes were used in nursery rhymes that otherwise contained real words, in order to isolate rhyme and meter effects from potential semantic and lexical confounds. Two types of rhyme effects were investigated, *rhyme sensitivity,* for the first occurrence of rhyme at the end of line 2 of each nursery rhyme stimulus, and *rhyme repetition* at the end of line 4. The contribution of a regular metrical context on both rhyme effects was explored. Cluster-based permutation tests were used to test for differences between conditions. Note that these tests do not statistically assess the extent of the effect in time and space, but only whether there is a difference between conditions (Sassenhagen & Draschkow, 2019). Consequently, precise topography and latency of the reported clusters have to be interpreted with caution and warrant future replication.

For the first occurrence of rhyme at the end of line 2, rhymes resulted in a larger positivity than non-rhymes, which was strongest at posterior electrodes and lasted from around 450 to 900 ms (Figure 5.2B). This effect resembles a (presumably delayed) version of the prototypical N400 rhyme effect, reported earlier for poetic stimuli and single word priming studies. In addition, an early negativity for rhymes was observed at anterior electrodes between 250-450 ms post target word onset (Figure 5.2A). The anterior negativity for rhymes set in earlier in non-metered nursery rhymes at around 130 ms (Figure 5.3), as is evident from a marginally significant interaction between rhyme and meter in this early time window. The effect of meter alone resulted in more positive waveforms in metered nursery rhymes across the entire scalp, within a single elongated cluster starting at around 400 ms. The positivity for meter is in line with previous reports of positive ERP deflections in response to metered speech.

For the rhyme repetition effect at the end of line 4, only the late positivity for rhymes was observed: repeating rhymes were again more positive than non-rhymes from around 400 ms onwards and this difference was most pronounced at posterior electrodes (Figure 5.5). Based on visual inspection, it appears that these late posterior effects for line 2 and line 4 overlap in latency, topography and polarity, rendering it highly likely they reflect the same underlying cognitive process. There was no interaction between rhyme and meter at the target word on line 4. Metered nursery rhymes again elicited more positive waveforms across the entire scalp, however only within a marginally significant cluster and at a short and late time window (around 800 ms). Based on visual inspection, the effect of meter thus seems less pronounced at the end of line 4 than of line 2. We will now discuss differences in results between the current and a comparable previous study and then provide an interpretation of the posterior and the anterior rhyme effects.

Comparison with the comparable previous study

Obermeier et al., (2016) studied how ERP rhyme and meter effects correlate with aesthetic liking of German folk song stanzas. ERP effects for both rhyme and meter, were more positive in the N400-time window and more negative in the (late) P600 time window. These effects were most pronounced in the metered condition, which was attributed to increased processing ease for prototypical rhyming metered stanzas compared to the other conditions. In the current study, no influence of meter was observable at the delayed posterior (N400) rhyme effect, neither at line 2 nor at line 4. Instead, a marginally significant interaction between meter and rhyme became apparent in a much earlier time window, indicating the anterior negativity for rhymes to set in earlier in non-metered stimuli.

The differences in effects are probably attributable to several differences in experiment design, especially concerning lexicality and task effects. Firstly, while Obermeier et al. used real words at line endings, here we used pseudowords that adhered to Dutch phonotactics, but have no meaning. The exact phonological form of rhyming words was thus more predictable in the study by Obermeier and colleagues, due to the semantics of the surrounding verse context constraining the number of possible lexical items occurring as rhyme words (see Hurschler et al., 2015 and Magne et al., 2007 for similar suggestions). The rhyme effects in the current study can rather be considered as purely phonological effects, unbiased by lexical activation resulting from the verse context. Secondly, the effect of meter in the previous study might have been enhanced by the rhythmicity judgment required for each stimulus. This might have enhanced the processing of meter and might in turn have affected the integration of line-final rhyme words into the verse context. The interaction between rhyme and meter in the study by Obermeier et al. was thus at least in part mediated by a third variable, namely lexical activation, which was facilitated in metered compared to non-metered and probably strengthened by the explicit judgment required about rhythmicity of each verse.

In the current study, listeners were allowed more shallow and implicit rhyme and meter processing, as no explicit judgement was required. Consequently, the interaction effects between rhyme and meter in the current study were only observable to a lesser extent and were limited to the early anterior rhyme effect, suggesting meter effects on implicit phonological processing to be subtle. Future passive listening studies could rely on block designs, wherein metered and non-metered stimuli are not presented directly back-to-back. This might allow participants to become more immersed in the metrical context, presumably a necessity for stronger rhyme-meter interaction effects to unfold with pseudoword rhymes during passive listening.

On a final note, the manipulation of meter was more strictly controlled in the current study: non-metered nursery rhymes were always one syllable longer or shorter than their metered counterparts and the number of syllables per verse line was counterbalanced across the whole stimulus set, see Materials section above. In the previous study, non-metered stimuli were always longer than their metered counterparts (one or two syllables, no info about counterbalancing). Non-metered stimuli being consistently longer than their metered counterparts might have biased the effects reported in the previous study, as the shorter metered stimuli might have been perceived as more succinct than longer non-metered stimuli. Our meter manipulation might also have rendered the acoustic difference between metered and non-metered stimulus versions smaller than in the previous study, as a two-syllable-difference might result in a stronger perceptual contrast between metered and non-metered stimuli than the one-syllable-difference used here.

Functional specification of posterior rhyme effects

The posterior positivity for line-final pseudoword rhymes found here corresponds well with previous reports of phonological repetition effects in the N400-range. These were previously reported during rhyme judgements of words and pseudowords, but only inconsistently during lexical decision tasks (Perrin & García-Larrea, 2003, Praamstra et al., 1993) and never in passive listening with pseudoword items (Davids et al., 2011). It was therefore thought that the N400 rhyme effect is not the result of an automatic rhyme response, but rather reflects explicit rhyme judgements, requiring attention to phonological stimulus features (Yoncheva et al., 2013; Perrin & García-Larrea, 2003; Davids et al., 2011; Mohan & Weber, 2015).

The current results contradict this notion, given that phonological N400 effects were found during passive listening. We suggest that phonological repetition effects in the N400 range during passive listening depend on stimuli which strongly license the occurrence of phonological repetition, such as rhymes in poetry. The absence of N400 rhyme effects in the previous passive listening studies (Davids et al., 2011; Perrin & García-Larrea, 2003) might be explained by these studies using (pseudo)word pair stimuli, which were not embedded in verse lines or sentences. Note again that the posterior rhyme effects reported here only started around 400 ms after critical word onset (Figure 5.1 and 5.4), whereas the typical N400 rhyme effect usually already peaks around 400/450 ms. Delayed N400 rhyme responses were reported earlier for pseudowords as well as phonologically illegal non-words during rhyme judgements. The late onset of the posterior rhyme effect reported here is thus attributable to the pseudoword status of the critical words. Presumably, the presence of a lexical entry for rhyming words facilitates rhyme judgements (Coch et al., 2005; Rugg, 1984a; Wagensveld, Segers, van Alphen, & Verhoeven, 2013), while the absence of semantic content strengthens the emotional appeal of rhymes (Obermeier et al., 2013).

Functional specification of anterior rhyme effects

In addition to a posterior rhyme effect, the current study showed an anterior negativity for rhymes at around 250 ms, which started earlier (around 130 ms) in the non-metered than in the metered condition. So far, such an effect has not been reported for poetic rhyme processing. Though note that data analysis in previous studies often only started after 200 ms (Obermeier et al., 2016; Vaughan-Evans et al., 2016), rendering it possible that a similar anterior effect was actually present. Previous work has reported anterior negativities for rhymes in later time windows (only starting around 300 ms) during active listening involving rhyme judgements (Coch et al., 2002, 2005; Grossi et al., 2001; Mohan & Weber, 2015) and passive listening (Andersson et al., 2018; Davids et al., 2011). These previous studies already suggested a functional distinction between anterior and posterior rhyme effects (Andersson et al., 2018; Davids et al., 2011; Khateb et al., 2007; Mohan & Weber, 2015). For example, Davids et al. (2011) linked the posterior rhyme effect with explicit rhyme judgements and the anterior rhyme effect with implicit pre-attentive rhyme processing. Related to that, Mohan & Weber (2015) suggested the anterior rhyme effect to index phonological segmentation and rehearsal, which precedes explicit rhyme judgements reflected in the posterior rhyme effect. We will revise this notion in light of the current results.

Anterior ERP effects are generally associated with working memory load (Coulson, King, & Kutas, 1998; Kluender & Kutas, 1993a, 1993b; Rösler, Pechmann, Streb, Röder, & Hennighausen, 1998), with more negative ERPs representing an increase in cognitive effort, for example as a result of resolving a filler-gap dependency. The verse line in poetry with its metrical scaffold represents a crucial unit for working memory (Fabb, 2014) and rhyme and meter in verse strongly constrain the phonological form of line-final words (Wallot & Menninghaus, 2018). We suggest the working memory involvement in phonological form-regularity processing to be reflected in anterior rhyme effects. Specifically, poetic rhymes could be considered a phonological long-distance dependency, as they require matching of phonological overlap across two verse lines. Upon reaching the end of a verse line, listeners engage in the working memory operation of re-evaluating previously processed verse lines to find the corresponding rhyme word. This “back association” (Barkley et al., 2015) is only necessary for rhyming words, hence the negativity for rhymes. Previous fMRI results (Hurschler et al, 2013;2015) already pointed towards the involvement of the phonological loop and subvocal rehearsal, subcomponents of verbal working memory (Baddeley et al., 1984), during rhyme judgements in sentence and verse contexts. Moreover, the CNV in response to primes during rhyme judgement tasks has been associated with working memory involvement, specifically phonological rehearsal (Rugg et al., 1984a, b). Possibly, the anterior rhyme effect observed here is specific to the context in which rhymes occurred, i.e., sentences and poetry.

The phonological working-memory effort involved in resolving rhyme-dependencies across verse lines is apparently not maintained until the end of each nursery rhyme stimulus, given that the anterior rhyme effect was not elicited at the end of line 4. We tentatively suggest attention and the specific rhyme scheme used here to influence occurrence of the anterior rhyme effect. Only rhyme schemes of the form AAAA were used in the current study, meaning that in a rhyming stimulus always the same rhyme was repeated at the end of each verse line (Table 5.1). This might lead listeners to engage more attentional resources during the first two than during the last lines of each nursery rhyme, as they knew that no new type of rhyme would occur. Once the first line-final rhyme is encountered, following rhymes might be processed with less attention and without backward matching of previously heard verse lines, resulting in the anterior rhyme effect to disappear. Crucially, the phonological repetition of rhyme itself is still being (subconsciously) detected, as indicated by the persistence of the posterior rhyme effect at the end of line 4.

In summary, the anterior rhyme effect could be considered a reflection of the working memory process of matching rhyme words across verse lines. The effect requires attention and working memory effort in the listener and hence disappears once top-down processing (semantics, rhyme patterns) or explicit tasks (Perrin & García-Larrea, 2003; Davids et al., 2011) do not require active rhyme processing. The posterior rhyme effect, however, is more stable to attentional and working memory processes and indexes subconscious perception of phonological repetition in the input.

Why then did the anterior rhyme effect set in earlier in non-metered nursery rhymes? Based on the literature cited above, the opposite would be more intuitive, as regular meter should facilitate prediction of upcoming rhymes (Rothermich & Kotz, 2013; Rothermich et al., 2012). ERP rhyme effects should thus be more pronounced or start earlier in the metered condition (Obermeier et al., 2016). One possible explanation relates to the peculiarity of rhymes in non-metered contexts. Rhymes are a phonological phenomenon that is typically associated with poetic and rhythmic language (Obermeier et al., 2013/2016), as rhymes only rarely occur outside of a metered context. The non-isochronous speech rhythm of the non-metered nursery rhymes is presumably more in line with the rhythm of natural speech (Nolan, 2014), wherein rhymes only rarely occur. The surprising presence of rhyme in non-metered stimuli might thus be more salient to listeners and result in more attentional resources allocated towards rhyme processing in non-metered stimuli.

The early rhyme-meter interaction might also indicate differences in the *processing mode* of metered versus non-metered nursery rhymes. A non-metered stimulus might require more active/engaged listening, as the exact moment of occurrence of the rhyme word is less predictable from the rhythmic context. Metered nursery rhymes, however, might elicit a *rhythmic-attending-mode* (Schroeder & Lakatos, 2009), which in turn might result in anterior rhyme effects to arise later, due to them being bound in onset time to the rhythmic envelop of the speech stream. We will follow up on this notion in chapter 6.

Conclusion

Delayed prototypical posterior N400 rhyme effects were observed for pseudoword rhymes occurring the first time (line 2) and repeating throughout verse lines (line 4) of nursery rhyme stimuli. In previous research, the N400 rhyme effect was mainly reported from explicit rhyme judgements and active listening conditions. Here we showed that in naturalistic rhyming contexts the effect is robust even in the absence of explicit attentional demands and lexico-semantic integration of the rhyme words. This robustness renders the N400 rhyme effect a suitable tool for research on rhyme perception in clinical and developmental populations. In addition, an anterior negativity for rhymes was found that preceded the posterior rhyme effect. The effect set in earlier in non-metered nursery rhymes and its appearance was limited to the first rhyme detection on line 2 of the nursery rhymes. The effect might especially occur in implicit rhyme processing in sentence/poetic contexts and might be associated with working memory processes required to associate rhyming verse lines with each other.

6

Chapter 6

Infants’ neural tracking of rhythm in spoken nursery rhymes relates to their vocabulary size

This chapter is based on:
Hahn, L. E., Snijders, T. M., Infants’ neural tracking of rhythm in spoken nursery rhymes relates to their vocabulary size (*Manuscript in preparation*).

Stimuli for this chapter are available at: https://osf.io/87b6y/

Abstract

Neural tracking of the low frequency amplitude envelope of speech is considered a driving factor for unimpaired language processing and development. This study explored to what extent infants' neural tracking of spoken nursery rhymes is modulated by top-down and bottom-up processing factors, namely vocabulary size and regular speech rhythm. In addition, infants' ERPs for rhyming and non-rhyming pseudowords at the end of each nursery rhyme line were compared for metered versus non-metered nursery rhymes and were related to neural tracking. Dutch 10.5-month-old infants listened to naturalistic rhyming and non-rhyming nursery rhymes that either had a salient 1.5 Hz stress rhythm (metered nursery rhymes) or deviated in one syllable to disrupt the rhythmic pattern (non-metered nursery rhymes). Neural tracking was found for the stress rate as well as for the syllable rate of the nursery rhymes, for both metered and non-metered nursery rhymes. At the syllable rate, neural tracking was higher for non-metered nursery rhymes. Improved syllable tracking of non-metered nursery rhymes might reflect infants' maturing auditory processing abilities in dealing with non-isochronous everyday speech. Syllable rate tracking was negatively correlated with concurrent vocabulary, but positively correlated with future vocabulary, suggesting a developmental change in infants' auditory processing focus as their lexicons increase. ERPs for rhyming line-final pseudowords in the nursery rhymes resembled previous research for rhyming songs (chapter 4). Yet their relationship with speech rhythm, neural tracking and lexical development raise questions for further research.

Acknowledgements

We thank Demi Lamers, Fleur Vissers, Maartje Stroo and Katharina Menn for their assistance during EEG measurements. Angela Khadar, Jessica Askamp, Iris Schmits and Demi Lamers are gratefully acknowledged for their support in lab set-up, recruiting and scheduling.

Introduction

Recent developments in cognitive neuroscience suggest neural activity during speech processing to correspond to the amplitude modulations present in the speech envelope (Giraud & Poeppel, 2012; Obleser & Kayser, 2019; Peelle & Davis, 2012). This phenomenon is referred to as neural tracking / entrainment (see Haegens & Zion Golumbic (2018) for terminology) and is assumed to enhance speech comprehension (Ahissar et al., 2001; Peelle & Davis, 2012) and intelligibility (Keitel, Gross, & Kayser, 2017; H. Luo & Poeppel, 2007) in adult listeners. There is accumulating evidence for infants to show neural speech tracking just like adults (Kalashnikova et al., 2018; Jessen et al., 2019; Choi et al., 2020; Lang et al., 2020; Menn et al., 2020; Snijders, 2020; Ortiz Barajas et al., 2021). Yet this research so far leaves open whether speech tracking during infancy is a purely bottom-up stimulus driven process, or whether and to what extent it is modulated by infants' current phonological processing and lexical knowledge. The current study investigates infants' ability to follow the amplitude envelope of spoken nursery rhymes with varying rhythmicity, and relates this ability to their rhyme perception and vocabulary size.

Neural tracking refers to the excitability of neurons aligned to relevant moments in an auditory signal (Gomez-Ramirez et al., 2011; Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008). It represents no single cognitive processing mechanism, but rather reflects a "cyclic processing constraint" of the brain on cognitive processes (Lakatos et al., 2005; Rimmele, Morillon, Poeppel, & Arnal, 2018). Neural tracking is proposed to account for predictive timing and attention selection (Meyer, 2018; Rimmele et al., 2018). According to the dynamic-attending in time model (DAT, Large & Jones (1999)), rhythmic pulses present in the amplitude envelope of an auditory signal lead to shifts in excitability cycles of neurons: listeners enter a stimulus driven "rhythmic-attending mode" (Schroeder & Lakatos, 2009), wherein neural excitability is adjusted (phase reset) to fit the rhythmic pattern in the auditory signal. In line with this assumption, adults' neural tracking of speech has been found to benefit from a regular versus less regular rhythmic context (e.g., a metrical pattern induced by melodies or rhythmic speech (Falk, Lanzilotti, & Schön, 2017; Kayser, Ince, Gross, & Kayser, 2015; van den Bosch der Nederlanden, Joanisse, & Grahn, 2020).

Yet, everyday speech can at most be considered quasi-rhythmic (Jadoul, Ravignani, Thompson, Filippi, & de Boer, 2016; Turk & Shattuck-Hufnagel, 2013). The rhythmic attending mode alone thus cannot explain tracking of an aperiodic acoustic stimulus such as speech. Consequently, accounts on neural speech tracking in adults have been updated to incorporate top-down cognitive processing on the rhythmic attending mode. Specifically, semantic and syntactic knowledge of the listener together with motor processing are proposed to allow for anticipatory phase-resets and complement predictions originating solely from acoustic stimulus features (Meyer, 2018; Rimmele et al., 2018). So far, it is an unsolved issue whether infants' neural tracking as well is influenced by concurrent linguistic knowledge, and whether their neural tracking benefits from a regular versus less regular speech rhythm.

Linguistic rhythm perception during infancy

A long line of research provides evidence that infants are sensitive to rhythmic differences between languages from early on in their lives (Nazzi, Bertoncini, & Mehler, 1998; Ramus, 1999; Ramus, Nespor, & Mehler, 1999) and that infants use speech rhythm to identify essential phonological units in the speech stream (Cutler, 1994; Johnson & Jusczyk, 2001; Jusczyk, Houston, & Newsome, 1999; Langus, Mehler, & Nespor, 2017). For example, stressed syllables are marked by heightened speech amplitude and constitute "acoustic edges" (Doelling, Arnal, Ghitza, & Poeppel, 2014), which in turn often signal word boundaries in languages with lexical stress (see e.g., Cutler & Carter (1987) for English and Schreuder & Baayen (1994) for Dutch). Infants' ability to recognize stressed syllables in fluent speech enables their word segmentation and individual differences therein relate to their vocabulary size (Junge et al., 2012; Kooijman, Hagoort, & Cutler, 2009; Newman, Ratner, Jusczyk, Jusczyk, & Dow, 2006).

The acoustic expression of stress- and phrase-level prosody is particularly salient in rhythmic infant-directed speech, for example in songs and spoken nursery rhymes (Leong et al., 2017, Leong & Goswami, 2015; Falk & Kello 2017). In such oral language play, phonological "landmarks", organized within the prosodic hierarchy (Nespor & Vogel, 2007), can be associated with designated nested frequency ranges: intonational phrase boundaries tend to occur in rhythmic cycles of approximately one per second, resembling a 1 Hz low delta rhythm. Phrases in turn carry prosodic feet, occurring at around 2 Hz and syllables at around 3 Hz, representing the delta and theta rhythm (Leong & Goswami, 2015; Leong, Kalashnikova, Burnham, & Goswami, 2017).

A number of recent studies provide evidence for infants' neural tracking of the speech rhythm at the phrasal and syllable level within the delta and theta frequency range (Kalashnikova et al., 2018; Jessen et al., 2019; Lang et al., 2020; Menn et al., 2020; Snijders, 2020; Ortiz Barajas et al., 2021). While these studies differ substantially in the type of stimuli used (ID-speech, cartoon movie, sung nursery rhymes, rhyming story books and an artificial language, respectively) and the age and language background of the participants (newborns up to 12-month-olds listening to Australian English; German; Dutch; French or Spanish stimuli), they converge on the fact that infants do indeed track the speech envelope at multiple timescales.

Yet, none of these previous studies directly compared neural speech tracking during infancy for more versus less rhythmic infant-directed speech (henceforth: ID-speech) and in how far this process is influenced by infants’ maturing linguistic abilities. Solving these questions would further our understanding of neural speech tracking during infancy as a top-down versus bottom-up / stimulus driven process. Kalashnikova et al. (2018), observed higher neural tracking in the theta-band in 7-month old infants listening to ID-speech compared to adult-directed speech, the former speech register being associated with a particularly strong delta rhythm and a larger number of stressed syllables compared to the latter (Leong, Kalashnikova, et al., 2017). Infants’ auditory attention might thus be more strongly engaged by a more rhythmic speech signal, especially by rhythmically occurring stressed syllables. Ortiz Barajas et al. (2021) report neural tracking to a repeated sentence in native as well as non-native languages, for 6-month-olds as well as asleep newborns, suggesting that no previous (lexical) knowledge or attention is required for neural tracking to occur. These authors differentiate between phase tracking on the one hand, as a bottom-up stimulus-driven perceptual tool and amplitude tracking, which operates more on spectral detail of the speech envelope and may require more exposure to the native language. The current study will follow up on this differentiation by investigating in how far infants’ neural phase tracking is modulated by speech rhythm, induced by regularly occurring stressed syllables and their lexical knowledge.

The role of neural tracking for phonological development

The quality of behavioural and neural rhythmic tracking relates to the quality of language perception and literacy. Pre-schoolers’ accuracy in drumming to a beat predicts their neural syllable envelope encoding and phonological awareness (Woodruff Carr, White-Schwoch, Tierney, Strait, & Kraus, 2014). Moreover, individual differences in music rhythm discrimination explain variance in phonological awareness and grammar skills (Gordon, Shivers, et al., 2015). Deficient neural tracking of speech rhythm is associated with impaired language development (Lallier et al., 2018). More specifically, delta tracking has been found to be lower in children with dyslexia (Molinaro, Lizarazu, Lallier, Bourguignon, & Carreiras, 2016; Power, Mead, Barnes, & Goswami, 2013) and is considered a risk factor for developmental speech and language disorders (see Ladányi, Persici, Fiveash, Tillmann, & Gordon, 2020 for review).

In light of this research, it has been suggested that the development of phonological representations requires oscillatory tracking of rhythmic units in speech (Goswami, 2019a, 2019b). Especially tracking of delta and theta rhythms in speech is considered of profound importance, as these reflect the phrasal-, stress- and syllable envelope of the speech stream (Doelling et al., 2014). The ability to track linguistic units at these low frequency bands is proposed to have cascading effects into language processing at higher frequencies (Myers, Lense, Gordon, 2019).

The current study will elucidate whether neural tracking indeed plays a role in establishing phonological representations. Specifically, 10.5-month-old infants' tracking of stress and syllable rhythm in speech will be related to their rhyme perception in spoken nursery rhymes. Being able to differentiate the odd one out in a set of rhyming and non-rhyming words is considered one of the earliest forms of phonological awareness children develop (Ziegler & Goswami, 2005). Rhyme awareness, as assessed from around three years onwards, also serves as a standard predictor for future literacy (Melby-Lervåg et al., 2012; Wood & Terrell, 1998). While tentative evidence for infants' implicit rhyme perception in language play is accumulating (chapter 3 and 4), this ability has so far not been related to neural speech tracking. Applying the classical DAT theory (Large & Jones, 1999), the perception of rhymes should be boosted in rhythmic speech like nursery rhymes, due to the regular speech rhythm directing auditory attention towards the stressed rhyme word. Computational modelling results suggest tracking of the delta rhythm might enable pre-schoolers to make rhyme judgements in the rhyme oddity task, specifically by delta-beta coupling (Leong & Goswami, 2016). Moreover, connectivity in the brain network underlying delta tracking correlates with phonological awareness in dyslexic readers (Molinaro et al., 2016). These studies both speak in favour of a relationship between low frequency speech tracking and rhyme perception.

The role of neural tracking for lexical development

Word stress functions as a reliable cue to word boundaries (see references above) and infants' ability to segment words based on stressed syllables predicts concurrent and future vocabulary (Junge et al., 2012; Kooijman et al., 2013; Newman et al., 2006). These findings might be grounded in neural tracking of rhythmic units like word stress in the speech amplitude envelope. Indeed, infants who were better able to track the stress rhythm of ID-speech at 7.5 months-old showed the more mature novelty preference during a behavioural word segmentation experiment at 9 months-old (Snijders, 2020). In accordance with these findings, Choi et al. (2020) observed a correlation between 6-month-olds' word-rhythm tracking at around 1 Hz in an artificial language and subsequent segmentation of these words in a behavioural experiment.

Infants' observed ability to track the speech rhythm at the stress and syllable rate might be further enhanced by a speech signal that carries a particularly salient stress- and syllable rhythm, such as oral language play provided in the form of songs and nursery rhymes. Menn et al. (2020) report a positive relationship between 10 to 14 months-olds' syllable tracking in sung nursery rhymes and their vocabulary size at two years old. The current study will compare infants' neural tracking specifically for less versus more rhythmic speech and assess whether the relationship with vocabulary holds up.

The current study

We aim to answer the following main research questions: do infants show neural tracking of the stress- and syllable rhythm of ID-speech and do they, similarly to adult listeners, benefit from a particularly rhythmic speech signal? Speech-Brain-Coherence (henceforth: SBC) will be used as an index of neural phase tracking (Peelle, Gross, & Davis, 2013), representing the consistency of the phase-difference between the speech amplitude envelope and the EEG signal. Dutch infants of 10.5-months of age will be exposed to spoken nursery rhymes manipulated to carry either a regular or irregular alternation of stressed and unstressed syllables (metered versus non-metered nursery rhymes). Based on previous research, we hypothesize that infants track the overall rhythm of the nursery rhyme stimuli at the syllable and stress rate regardless of whether they are regularly metered or not. Yet, we expect higher SBC values for metered nursery rhymes compared to non-metered nursery rhymes at the stress frequency rate of our stimuli, as the stress rate was strongly controlled during recording and this regular rhythm should result in enhanced stimulus driven neural tracking (DAT, Large & Jones, (1999)).

Furthermore, we will explore whether individual differences in infant SBC are associated with variance in lexical and phonological development. SBC values will be related to 1) concurrent and later vocabulary size and to 2) rhyme perception (rhyme ERPs). We expect infants with higher stress rate SBC values to have larger concurrent and later vocabularies, due to these infants being able to better represent the stress rate of the speech stream, which in turn should enable them to more easily segment words from fluent speech (Snijders 2020; Choi, et al., 2020; Menn et al., 2020). Applying the classical DAT theory (Large & Jones, 1999) to our paradigm, we expect the perception of rhymes to be boosted in metered stimuli, due to the regular speech rhythm directing auditory attention towards the stressed line-final rhyme word.

Rhyme ERPs will be measured in response to rhyming and non-rhyming stressed CVC pseudowords at the end of each line of the nursery rhymes. In accordance with previous research (chapter 4 and 5), rhyme ERPs will be separated for first occurrences of rhymes in a given nursery rhyme (*rhyme sensitivity*) and rhymes at following lines within the same nursery rhyme (*rhyme repetition*). Based on these previous studies (chapter 4 in particular), we expect a rhyme positivity for the ERP rhyme sensitivity effect and a rhyme negativity for the ERP rhyme repetition effects at central electrodes. Due to differences in speech rate between the song (chapter 4) and speech stimuli used here, we have no clear predictions about precise timing of these effects. To adhere to previous research (chapter 4), rhyme ERPs will also be correlated with vocabulary scores.

Method

Participants

Sixty-two 10-to-11 months old infants were tested. Data from 41 infants was used for the analyses of SBC and the ERP rhyme repetition effect (mean age: 322 days, range: 298 – 344 days, 19 females). Sample size was determined prior to data collection based on similar ERP studies by Kooijman and colleagues (2005, 2013) and Junge and colleagues (2014, 2012). Infants were recruited via the Baby and Child Research Center at Radboud University, Nijmegen, the Netherlands, were born full-term, had no current ear infection and came from monolingual Dutch households. Caregivers of two infants reported to have dyslexia. These infants were not excluded from analysis. Caregivers received money (20€) or a small gift for their participation. All caregivers signed an informed consent form prior to starting the measurement. The study was approved by the Ethical Board of the Faculty of Social Sciences, Radboud University Nijmegen, CMO 2012/012.

One infant was excluded due to refusal to wear the cap. Additionally, infants were excluded from statistical analysis because of experimenter error (n = 1), excessive movement artefacts (n = 1) or more than three adjacent noisy channels (n = 4). Another 14 infants were excluded due to too few trials remaining after artefact rejection for the rhyme repetition effect, and three additional infants were excluded for the same reason from analysis of the rhyme sensitivity effect. See the analysis section below for more detail on infant and trial exclusion.

Nursery rhyme stimuli

Stimuli were the same as in the adult version of the experiment (chapter 5), 52 nursery rhymes, reminiscent of typical Dutch nursery rhymes in terms of structure and linguistic content (see Table 6.1 for an example stimulus). The entire stimulus set is provided here: https://osf.io/87b6y/. Each nursery rhyme consisted of four lines, each ending in a monosyllabic phonologically legal CVC pseudoword. Nursery rhymes were manipulated for the factors rhyme (rhyming vs. non-rhyming) and meter (metered vs. non-metered), resulting in four different stimulus conditions (metered_rhyming; non-metered_rhyming; metered_non-rhyming; non-metered_non-rhyming) and a stimulus set of 208 unique nursery rhymes (52 original nursery rhymes * 4 experimental conditions).

Table 6.1 Example stimulus for each experimental condition.

Metered_Non-rhyming	**Metered_Rhyming**
GiRAFfen Eten VERse FIP.	GiRAFfen Eten VERse BIJP.
Ze KNAgen TELkens GROEne **SOS**.	Ze KNAgen TELkens GROEne **FIJP**.
BeHOORlijk KAAL is DUS die **LIJG**,	BeHOORlijk KAAL is DUS die **DIJP**,
Na EEN beZOEK van VIJFtien **FAAF**.	Na EEN beZOEK van VIJFtien **KIJP**.
Non-metered_Non-rhyming	**Non-metered_Rhyming**
GiRAFfen Eten graag VERse FIP.	GiRAFfen Eten graag VERse BIJP.
Ze KNAgen ook TELkens GROEne **SOS**.	Ze KNAgen ook TELkens GROEne **FIJP**.
BeHOORlijk KAAL is DUS nu die **LIJG**,	BeHOORlijk KAAL is DUS nu die **DIJP**,
Na TWEE beZOEKen van VIJFtien **FAAF**.	Na twee beZOEKen van VIJFtien **KIJP**.
Literal translation (metered)	***Literal translation (non-metered)***
Giraffes eat fresh fip/bijp.	*Giraffes like to eat fresh fip/bijp.*
Again and again they gnaw green sos/fijp.	*Again and again they also gnaw green sos/fijp.*
Pretty bald is thus the lijg/dijp.	*Pretty bald is thus now the lijg/dijp.*
After one visit of fifteen faaf/kijp.	*After two visits of fifteen faaf/kijp.*

Note: Lyrics consist of real Dutch words, line-final words are legal pseudowords. Stressed syllables in CAPitals. Non-metered version created by adding/removing a syllable (counterbalanced). Critical words used to assess the rhyme sensitivity effect at the end of phrase 2 of the nursery rhymes are highlighted in green (rhyme) and purple (non-rhyme); Critical words to assess the rhyme repetition effect from phrase 3 and 4 highlighted in red (rhyme) and blue (non-rhyme). For the analysis, the line-final word of line 1 was disregarded, as this cannot rhyme.

Rhyme was manipulated by placing either rhyming or non-rhyming CVC pseudowords (the *critical word*) at the end of each line of the nursery rhymes. Critical words were approximately 500 ms long (see Table 6.2) and the same set of pseudowords was used across conditions. Meter was manipulated by varying the number of syllables per nursery rhyme line across conditions: metered nursery rhymes had 8-syllable lines in iambic meter, thus starting each line with an unstressed syllable (Table 6.1). This iambic meter was disrupted in the non-metered versions of each nursery rhyme by adding or removing a single syllable from each line. Thus, the corresponding non-metered version of a nursery rhyme either had 7-syllable lines (short) or 9-syllable lines (long), counterbalanced across the set of 52 original nursery rhymes. Changing the number of syllables was achieved by adding or removing particles and function words or by replacing content words with ones of similar meaning and leaving the rest of the nursery rhyme intact. The pre-final word of each line of each nursery rhyme was the same across conditions.

Table 6.2 Mean and standard deviation of line and critical word duration across conditions.

	Line (s)	Critical word (ms)
Condition	M (SD)	M (SD)
metered_rhyming	2.55 (0.21)	521 (82)
non-metered_rhyming (short)	2.23 (0.24)	520 (81)
non-metered_rhyming (long)	2.75 (0.27)	523 (80)
metered_non-rhyming	2.64 (0.19)	541 (80)
non-metered_non-rhyming (short)	2.27 (0.22)	528 (88)
non-metered_non-rhyming (long)	2.77 (0.24)	528 (75)

Stimuli were recorded from a female Dutch native speaker (the last author of the study), reading the stimuli in a child-directed manner. Recitation of the nursery rhymes was paced to a visual metronome set at 90 bpm to achieve a regular 1.5 Hz stress rate and 3 Hz syllable rate in the stimulus amplitude envelope (Figure 6.1). Silences were inserted at the beginning (50 ms) and end (30 ms) of every nursery rhyme and silences of 666 ms were inserted between the consecutive lines using Praat (Boersma & Weenink, 2014), resulting in nursery rhymes between 10 and 15 seconds long (mean = 12.7 sec, SD = 987 ms).

Visual inspection of the Frequency-Power Spectrum for each line of the 52 nursery rhymes in their metered and non-metered condition revealed that metered nursery rhymes had clear peaks in acoustic power at around 1.5 Hz, reflecting their regular stress rate. Moreover, both metered and non-metered nursery rhymes carried peaks in power at around 3 Hz, the syllable rate, which was the same for both conditions. In summary, the stimuli reflected the most common stress- and syllable rates for songs and infant-directed speech (Ding et al., 2016; Leong et al., 2014).

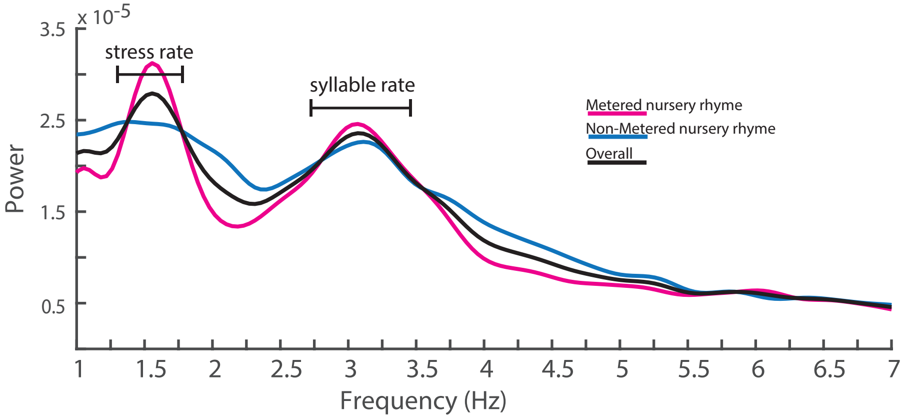


Figure 6.1 Frequency-Power spectrum from the lines of the nursery rhyme stimuli.

Procedure

Preparation and measurement were the same as in chapter 4. Each test session was run by two experimenters. One experimenter briefed the parents and ran the measurement, while the other entertained the infant during placement of the electrode cap and data collection. Data collection lasted around 25 minutes; the entire experimental session lasted around 1 hour. During the test session, the infant sat on a caregiver's lap in an electrically shielded room. Silent baby-friendly movie clips were played at a PC screen in front of the infant. One experimenter sat next to the caregiver to silently entertain the infant during the measurement if necessary. Both caregiver and experimenter listened to masking music over headphones throughout data collection. Stimulus presentation was controlled by Presentation software (Neurobehavioral Systems, Inc., Berkeley, CA, USA).

Parents filled in questionnaires on their musical and demographic background and the vocabulary of their child (N-CDI 1 at 10.5-months-old, concurrent vocabulary; and N-CDI 2B at 18 months-old; later vocabulary). Concurrent vocabulary was usually assessed around the day of testing. For some infants, however, more time passed before questionnaires were returned (average amount of days between date of testing and N-CDI return date: 3; range: 0 – 49 days). For the present study, only the vocabulary questionnaires have been analysed.

EEG activity was collected from 32 Ag/AgCl electrodes (ActiCAP) using BrainAmp DC and BrainVision Recorder Software (Brain Products GmbH, Germany). Electrode locations were in accordance with an extended 10/20 system: F7/3/4/8, FC5/1/2/6, C3/4, CP5/1/2/6, P7/3/4/8, Fz, FCz, Cz and Pz, for collection of EEG activity. Electro-oculogram (EOG) was recorded using an electrode on the left or right cheek and above the eye (Fp1/2) for vertical EOG, and left and right of the eyes (FT9/10) for horizontal EOG. AFz served as Ground, FCz as online reference. Impedance was typically kept below 25 kΩ. Data was collected with a sampling rate of 500 Hz using an online low-cut off filter of 10 Hz and high-cut off of 1000 Hz.

The 52 nursery rhymes in their four conditions were divided over four stimulus lists. Within each list, stimuli were randomized and presented within two experimental halves, resulting in 104 nursery rhymes presented within the full experiment duration. Across stimulus lists we counterbalanced for short and long non-metered versions. Within each stimulus list, the factors Meter and Rhyme were controlled to occur not more than twice in a row (e.g., not more than two non-rhyming nursery rhymes in a row) and the same condition was never immediately repeated (e.g., not more than one non-metered_rhyming in a row). The nursery rhymes were always played in their entirety without interruption, with 1300 ms silence between nursery rhymes.

Data preprocessing

EEG data were analysed in MATLAB (version 2018 and following; TheMathWorks, Natick, MA, USA) using the Fieldtrip toolbox (Oostenveld et al., 2011). Data was filtered offline at 0.1 - 30 Hz. Four channels (T7/8, P7/8) were removed already before artefact rejection, due to being noisy in many datasets. Bad channels in individual datasets were removed manually. Data were first cut in 1-sec data snippets, and data segments with flat channels or large artefacts (>150 µV for EEG channels, >250 µV for EOG channels) were removed. Eye/noise artefact components were identified (average of 2.6 noise components and 3.4 eye components) using Independent Component Analysis (Makeig et al., 1996) as implemented in the EEGlab toolbox (Delorme & Makeig, 2004) with infomax ICA (Bell & Sejnowski, 1995) on the 1-sec data snippets.

Speech-brain coherence specific preprocessing. EEG data was segmented into 4-second snippets for every second after the onset of line 1 of every nursery rhyme and low-pass filtered at 45 Hz and high-pass filtered at 1 Hz. The 4-second snippets were re-referenced, either to linked-mastoids (48 infants) or to a single mastoid electrode (7 infants) if one of the mastoid electrodes was noisy, and the eye- and noise-components identified with ICA were removed. The acoustic envelope of the nursery rhymes was obtained using a Hilbert transform with a 2nd-order Butterworth filter and was added as a separate channel to each EEG dataset. Remaining artefacts were removed using semi-automatic artefact rejection (cut-off +/-150mV for EEG electrodes) and noisy/missing channels were repaired using spline interpolation and a custom neighbourhood structure (32 channels in 21 infants). EEG data and speech envelope were fourier-transformed from 1 to 7 Hz, with a frequency resolution of .25 Hz due to the 4-second data segments. The cross-spectral density was computed for each combination of electrodes and the speech signal. Coherence between brain and speech signal was then computed as the absolute value of the mean cross-spectrum between the EEG signal and the speech signal, divided by the square root of the product of the mean power spectra of those signals (Rosenberg, Amjad, Breeze, Brillinger, & Halliday, 1989). Thus, speech brain coherence (SBC) gives, for every frequency, the phase-consistency between EEG data and the corresponding stimulus envelope, resulting in one coherence value between 0-1 per electrode and frequency (see for a similar approach e.g., Peelle et al., 2013).

To remedy differences in the number of trials each infant or condition provided for the analysis, speech-brain coherence was standardized. Random SBC was obtained by shuffling the speech envelope across EEG epochs and computing the average coherence over 100 pairings of a random speech envelope with the EEG data for Overall SBC, SBC metered and SBC non-metered. Standardized SBC was calculated by subtracting the Observed SBC from Random SBC and dividing the result by the sum of Observed SBC and Random SBC.

ERP specific preprocessing. Raw EEG data was epoched into the critical words (-200 to 900 ms around critical word onset, see note Table 6.1) using 0.1 - 30 Hz filters. Baseline correction was applied to the time-locked data with waveforms normalized relative to the 200 ms epoch preceding the onset of the critical word. Rereferencing and ICA-component rejection was applied with the same settings as for speech-brain specific preprocessing above. Remaining artefacts were removed using automatic artefact rejection (cut-off +/-150mV) and the same channels were repaired in the same data sets as in the speech-brain coherence specific preprocessing above. Infants were excluded from further analysis because they had fewer than 10 of the 52 possible trials per condition remaining in line 3 and 4 of the nursery rhymes (n = 14). From the remaining sample, another 3 infants were removed due to not contributing a minimum of 10 of the 52 trials from line 2 of the nursery rhymes for the rhyme sensitivity effect (see Table 6.1). Further noisy/missing channels were repaired using spline interpolation and a custom neighbourhood structure (32 channels in 21 infants). Event-related potentials were computed for the remaining 22 channels (Fp1/2, F7/8, F3/4, Fz, FC5/FC6, FC1/FC2, FCz, C3/C4, Cz, CP5/CP6, CP1/CP2, P3/P4, Pz) by averaging the amplitude time-locked to the onset of the critical word: 1) for the rhyme sensitivity effect, rhyming and non-rhyming trials from line 2 (Mean (SD) number of trials rhyme: 22.00 (8.68), non-rhyme: 23.32 (7.63), n = 38 infants) and 2) the rhyme repetition effect, averaged over ERPs from line 3 and 4 of nursery rhymes, Mean (SD) number of trials rhyme: 44.88 (15.12), non-rhyme: 45.30 (14.75), n = 41 infants). To analyse the interaction between rhyme and meter, trials from all four conditions from line 3 and 4 were considered (Mean (SD) number of trials for all four conditions: 22/23 (8/7), n = 41 infants.

Statistical Analysis

Speech-Brain Coherence. Speech-Brian coherence was used to evaluate differential processing of metered and non-metered nursery rhymes. Here the effect of Rhyme was disregarded, as it only occurred on the last word of every line. The analysis of SBC was limited to infants that were also included in the analysis of the ERP rhyme repetition effect (n = 41, Mean (SD) number of SBC trials per infant: 577 (208)). The difference in coherence was investigated at stress (1.25 – 1.75 Hz) and syllable (2.5 – 3.5 Hz) frequency rates of the experimental stimuli (see Figure 6.1). In addition, exploratory tests were run across the whole available frequency range of 1 to 7 Hz. SBC values were compared using cluster-based permutation tests.

First, to establish whether speech-brain coherence occurred, Observed SBC was compared to Random SBC (data with a shuffled speech envelope) for all nursery rhymes, regardless of experimental condition. Second, to assess the effect of Meter on SBC, standardized SBC of metered and non-metered nursery rhymes were first compared to each other and then their Observed SBC was compared to their corresponding Random SBC.

ERP rhyme sensitivity and rhyme repetition effects. The ERP analysis evaluated the effect of Rhyme, by assessing the difference between rhyme/non-rhyme at the critical words (the final words) from line 2 to 4 of the nursery rhymes. Two ERP effects were investigated in analogy with chapter 4 and 5: 1) the rhyme sensitivity effect, only on ERPs for rhyme/non-rhyme from line 2 (Table 6.1), n = 38 infants; and 2) the rhyme repetition effect, only on ERPs for rhyme/non-rhyme from line 3 and 4 (Table 6.1), n = 41 infants. To assess whether the effect of rhyme was different between the non-metered and metered condition, the difference score between rhyme and non-rhyme was compared for the two Meter conditions. This analysis of the interaction between Rhyme and Meter was limited to line 3 and 4 of the nursery rhymes, due to too many missing trials for line 2. Non-parametric cluster-based permutation tests (Maris & Oostenveld, 2007) were used to evaluate differences between conditions in the ERP waveform across 22 electrodes (Monte-Carlo permutation distribution, 1000 permutations as standard, 10000 permutations where p-values were around 0.05). The rhyme sensitivity effect was analysed on the average of single electrodes between 200 and 500 ms post onset of the critical words for line 2, based on previous research (chapter 4, Junge et al., 2014, Snijders et al., 2020). The rhyme repetition effect was analysed on the average of single electrodes between 0 and 200 ms post onset of the critical words of lines 3 and 4, also based on previous research (chapter 4). Correlations with vocabulary will be calculated based on averaged difference scores (rhyme – non-rhyme) from significant rhyme clusters only. In a similar manner, the interaction between Rhyme and Meter will be assessed within significant rhyme clusters only. There the average difference score for the factor Rhyme (rhyme – non-rhyme) will be compared between metered and non-metered nursery rhymes using paired *t*-tests.

Relating SBC to infant vocabulary. Vocabulary scores were obtained for comprehension at 10.5 months (data from 1 infant missing) and at 18 months (data from 1 infant missing). For concurrent vocabulary, percentiles of raw comprehension scores were used, due to a relatively wide spread in infant age when parents returned N-CDIs (range in age for concurrent vocabulary questionnaire: 46 days; for vocabulary at 18 months-old: 19 days). The relationship between SBC and concurrent (10.5 months) and later (18 months) vocabulary was investigated using cluster-based permutation correlations in Fieldtrip and robust regression models using the robust package (Maechler et al., 2021) in R (version 1.2.5033, R Development Core Team, 2020).

Relating SBC and infant vocabulary to ERPs. Regression models were set up to predict the size of the ERP rhyme repetition and ERP rhyme sensitivity effect based on concurrent vocabulary or overall SBC. Due to outliers in several factors and dependent variables, again robust regression was used.

Results

Speech-Brain Coherence

Overall SBC

Observed Speech-Brain Coherence was compared to data with a shuffled speech envelope, at 1.25 to 1.75 Hz for stress rate and 2.5 to 3.5 Hz for syllable rate (based on the stimulus characteristics, see Figure 6.1). This resulted in significant positive clusters at both frequency ranges (cluster *p* = .002 for both, Figure 6.2AB). To explore whether SBC was present outside of the stimulus-based stress- and syllable frequencies, overall SBC was compared to SBC from a shuffled speech envelope from 1 to 7 Hz, without averaging over any frequency ranges or channels. Three significant positive clusters were identified (Figure 6.3C): one cluster ranged from 1.5 to 4.5 Hz and involved all 22 electrodes, cluster *p* = .002; another cluster ranged from 1 to 1.25 Hz, and involved 17 electrodes (Fp1, F7/8, F3/F4, Fz, FC5/6, FC1/2, FCz, Cz, C4, CP1/2, CP6, P4), cluster *p* = .02; and a third cluster ranged from 5.75 to 6.5. Hz, involving 12 electrodes (Fp2, F7, F3/4, FC5/6, FC2, FCz, C4, CP2, CP6, Pz), cluster *p* = .03.

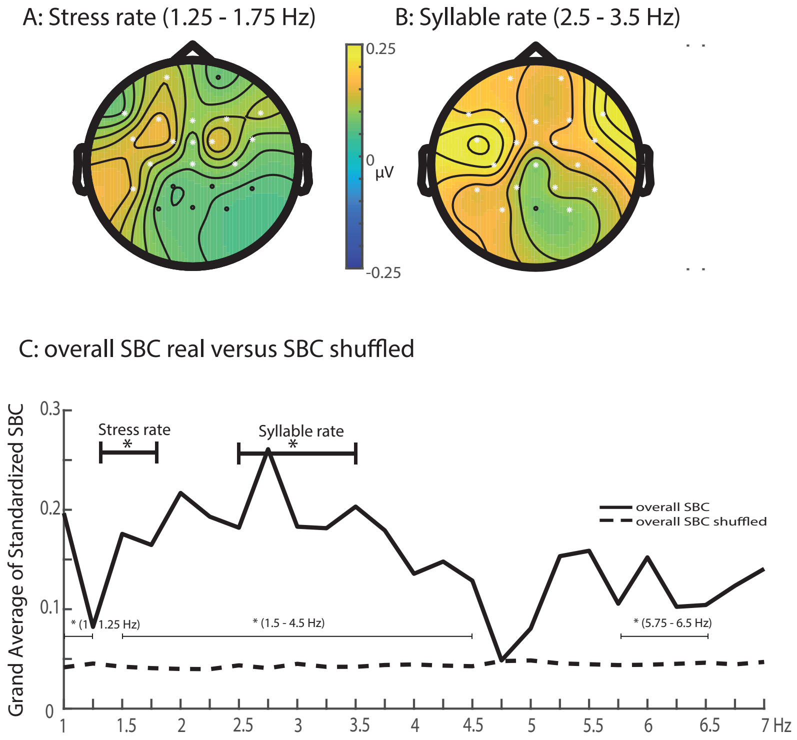


Figure 6.2 SBC in response to nursery rhymes. A, B: Topographic isovoltage maps of the difference between overall SBC and shuffled-envelope SBC at stress and syllable frequency rates; electrodes involved in significant clusters highlighted with white *; C: standardized overall SBC (solid) and shuffled-envelope SBC (dashed) averaged over all 22 electrodes; n = 41 infants.

SBC for metered and non-metered nursery rhymes

Observed SBC was compared to shuffled-envelope SBC for both metered and non-metered nursery rhymes separately (see Figure 6.3). Observed SBC was higher than shuffled-envelope SBC for both the metered and non-metered condition, for both stress (cluster *p* for metered = .006; for non-metered = .01) and syllable frequency rates (cluster *p* = .002 for both metered and non-metered).

The difference in SBC between metered and non-metered nursery rhymes was statistically compared at stress and syllable frequency ranges, as well as for the whole frequency range from 1-7 Hz. For stress rate, no difference was found in SBC between metered and non-metered conditions (no clusters identified). For syllable rate, SBC for metered nursery rhymes was significantly lower than for non-metered nursery rhymes over central electrodes (Figure 6.3B bottom row, cluster *p* = .03).

Comparing SBC for metered versus non-metered nursery rhymes for the whole frequency range from 1 to 7 Hz resulted in no significant clusters (lowest cluster *p* = .4).

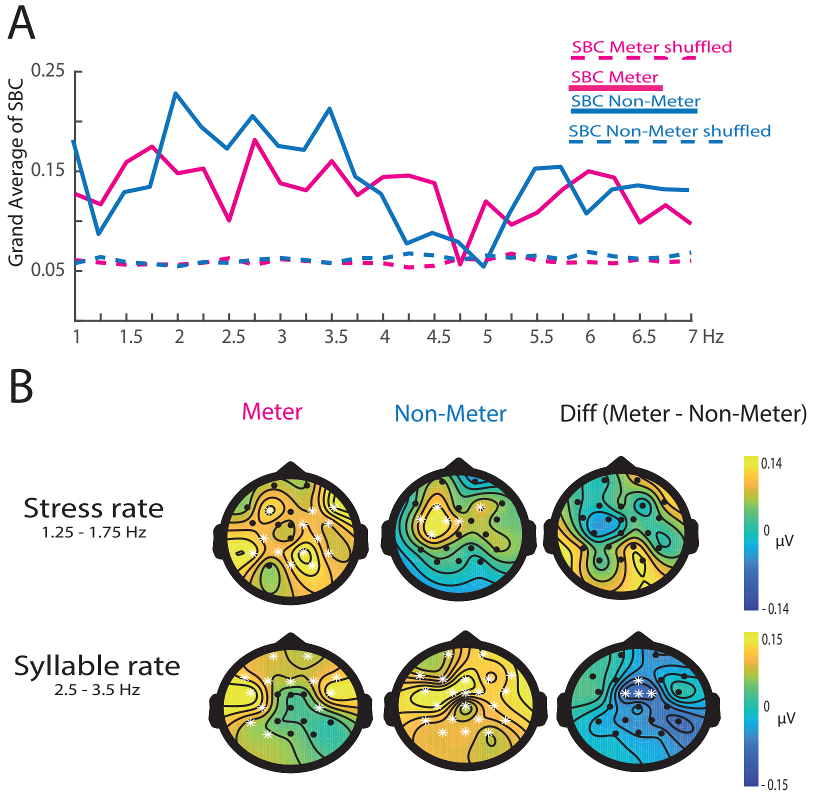


Figure 6.3 SBC for metered and non-metered nursery rhymes. A: Standardized SBC of all 22 electrodes for metered (pink) and non-metered (blue) nursery rhymes and corresponding shuffled-envelope SBC (dashed); B: Topographic isovoltage map of SBC at Stress (top row) and Syllable frequency rates (bottom row) for metered, non-metered and the difference between metered and non-metereded nursery rhymes; electrodes involved in significant clusters highlighted with white *; n = 41 infants.

ERP Results

Rhyme sensitivity effect

In response to the first occurrence of the Rhyme/Non-Rhyme (line 2 of each nursery rhyme), rhyming words induced a more positive ERP waveform compared to non-rhyming words (Figure 6.4A). The cluster-based permutation test on the 200-500 ms time-window resulted in 1 positive cluster at electrode F7, which did not survive multiple comparison correction (cluster *p* = .5, *Mean (SD)*$_{\text{Rhyme}}$ = 0.86 (4.18); *Mean (SD)*$_{\text{Non-Rhyme}}$ = -0.55 (3.47) within the cluster).

Rhyme repetition effect

ERPs for rhyming words occurring at the end of line 3 and 4 of the nursery rhymes were more negative than ERPs for non-rhyming words within the first 200 ms after word onset (Figure 6.4B). For the 0-200 ms time-window, the cluster-based permutation test revealed one negative cluster at electrodes FC5, C3, CP5 and CP1/2 (cluster p = .048, 10000 permutations), $Mean\ (SD)_{\text{Rhyme}}$ = -0.51 (1.68), $Mean\ (SD)_{\text{Non-Rhyme}}$ = 0.85 (2.40). The test for an interaction between Rhyme and Meter within the rhyme repetition cluster resulted in no significant effects (t = -0.57, p = .57, $Mean(SD)_{RhymeDiff\ Meter}$ = -1.62 (3.97), $Mean(SD)_{RhymeDiff\ Non\text{-}Meter}$ = -1.09 (4.32)).

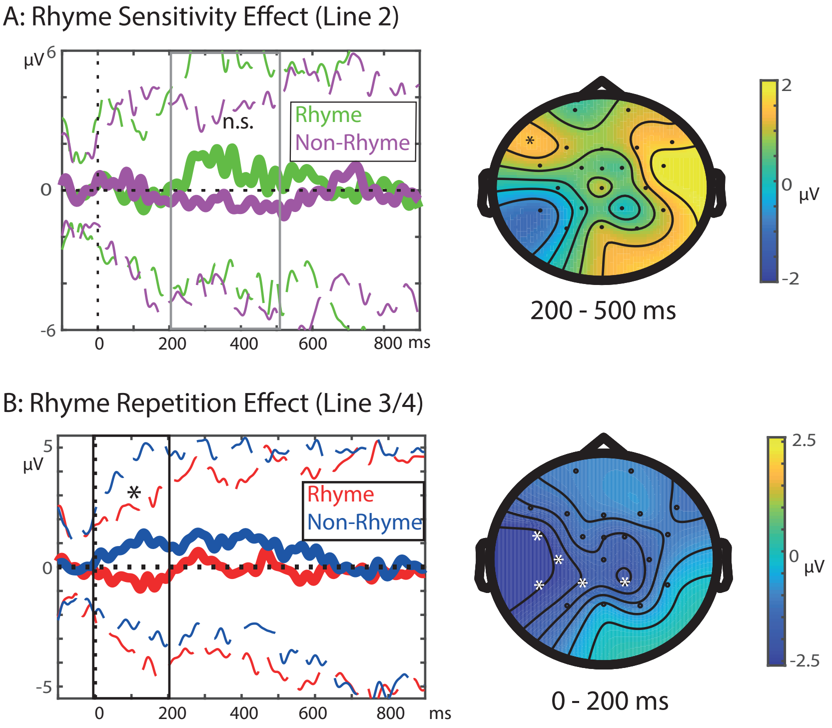


Figure 6.4 ERPs for rhyme sensitivity and rhyme repetition in nursery rhymes. A, Rhyme sensitivity effect, in response to the final word of line 2 of the nursery rhymes; left: ERPs averaged at electrode F7, solid lines = rhyme and non-rhyme; dotted lines = ± 1SD; right: Topographic isovoltage map of the difference between rhyme and non-rhyme between 200 and 500 ms; * indicates cluster electrode (p = .5; n=38 infants); B, Rhyme repetition effect, in response to the final word of line 3 and 4 of the nursery rhymes; left: ERPs averaged within cluster electrodes, solid lines = rhyme and non-rhyme; dotted lines = ± 1SD; right: Topographic isovoltage map of the difference between rhyme and non-rhyme from 0 to 200 ms; cluster electrodes highlighted with white * (*p* = .048; n=41 infants).

Relationship between ERP rhyme effects and Speech-Brain Coherence

The robust regression models revealed no relationship between SBC and the ERP rhyme sensitivity or rhyme repetition effects (all *p*'s $> .1$).

Relationship between ERP rhyme effects and vocabulary

Previous research indicated a marginally significant negative correlation between the ERP rhyme repetition effect in child songs and productive vocabulary at 18 months-old (chapter 4). In the current study, no relationship between rhyme repetition with concurrent or future comprehension scores was found (all *p*'s $> .1$). Testing explicitly for a relationship with productive vocabulary, as reported in the previous study, again resulted in no relationship with rhyme sensitivity (spearman's $\rho = -0.17$; $p = .3$) or rhyme repetition (spearman's $\rho = .02$; $p = .9$). For rhyme sensitivity, there was no relationship with concurrent vocabulary as well, but a marginal correlation with future comprehension scores was found. Infants with a larger ERP difference score for rhyme sensitivity had larger vocabularies at 18 months (spearman's $\rho = .30$; $p = .08$).

Relationship between infant vocabulary and Speech-Brain Coherence

Associations between SBC and vocabulary were first assessed using non-parametric cluster-based permutations in Fieldtrip using all 22 channels. To explore the influence of possibly confounding factors, robust regression models were run in addition, see text and footnotes for model-specifications.

Syllable rate SBC was negatively correlated with concurrent receptive vocabulary in a cluster at left-frontal electrodes sites (cluster $p = .055$; 10000 permutations using all 22 channels, Figure 6.5A). The relationship remained after controlling for infant age ($\beta_{SBC\_SyllRate} = -63.77$; $t = -2.39$; $p = .02$) or gender ($\beta_{SBC\_SyllRate} = -56.09$; $t = -2.45$; $p = .02$)[3], in a robust regression model using syllable rate SBC averaged over all 21 cluster electrodes showing syllable rate SBC (see Figure 6.2B for electrodes, model without any control variables: $\beta_{SBC\_SyllRate} = -64.04$; $t = -2.39$; $p = .02$). The negative correlation indicates that infants with less SBC at syllable frequency rate had higher concurrent vocabulary perception percentiles. For stress rate, no correlation was found between SBC and concurrent receptive vocabulary (smallest cluster $p = .4$). There were no correlations between vocabulary at 18 months-old and stress rate SBC, using all 22 channels (smallest cluster $p = .1$) or syllable rate SBC (no clusters found). However, after taking vocabulary at 10.5 months-old into account, average SBC at syllable rate (electrodes again from syllable cluster Figure

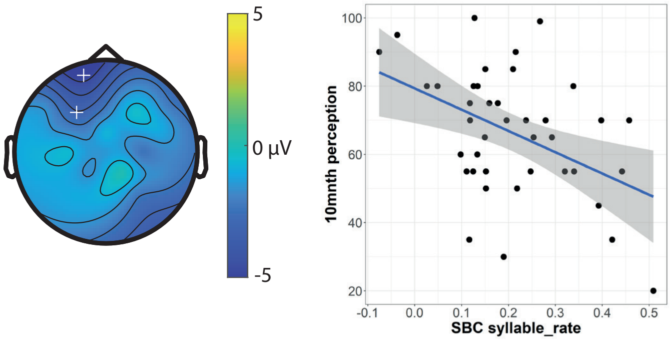
---

[3] lmrob(Percentiles_Comprehension10mnth ~ Age_atTest(days) + SBC_Syll); lmrob(Percentiles_Comprehension10mnth ~ Gender + SBC_Syll)

6.2B) became a significant positive predictor in the regression model (Figure 6.5B; $\beta_{SBC\_SyllRate}$ = 38.74; *t* = 2.30; *p* = .03[4]; partial correlation spearman's ρ = .28; *p* = .09).

Exploring the entire frequency range from 1 to 7 Hz for a correlation with concurrent or later vocabulary resulted in a significant negative correlation between SBC at 2.5 to 3.75 Hz at a left fronto-central cluster and concurrent vocabulary, confirming the results reported above (cluster *p* = .01; cluster electrodes: Fp1/2, F3, Fz, FC5, FC1/2, Cz, FCz). For later vocabulary, no significant clusters were found (lowest cluster *p* = .4).

A: Correlation syllable_rate SBC and concurrent vocabulary



B: Partial correlation syllable_rate SBC and later vocabulary

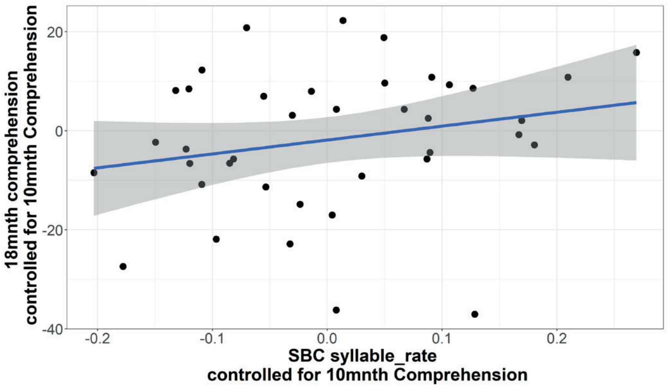


Figure 6.5 Correlation between syllable rate SBC and vocabulary. A: Topographic isovoltage maps of the significant correlation cluster (left side) and scatterplot for 21 electrodes (right side) of the correlation between syllable rate SBC and concurrent (10.5 months) vocabulary; cluster channels highlighted with white *; n = 40 infants; B: partial correlation between later vocabulary (18 months) and syllable rate SBC at all 21 electrodes, after controlling for concurrent vocabulary; n = 39 infants.

---

[4] lmRob(Comprehension18mnd ~ Percentiles_ Comprehension10mnth + SBC_Syll)

Discussion

Neural speech tracking has been suggested to play a key role in language development (Goswami, 2019ab) and speech comprehension. While first reports on infants' neural speech tracking are accumulating (see references below), it remains open which role individual linguistic abilities (top-down) and stimulus features (bottom-up) play in this phenomenon (Rimmele et al., 2018; Meyer, 2018). Here we assessed whether infants benefit from a salient stress rhythm in their neural tracking of ID-speech. 10.5-month-old Dutch infants passively listened to spoken nursery rhymes, manipulated for Rhyme (rhyming/non-rhyming) and Meter (metered/non-metered; 1.5 Hz stress rate). Nursery rhyme stimuli had a 3 Hz syllable rate (Figure 6.1). SBC was used as an EEG index for neural phase tracking of the stress and syllable rhythm and individual differences in infants' phase tracking were related to infants' vocabulary size. As a secondary analysis, neural speech tracking has been linked to infants' perception of line-final rhymes in the nursery rhyme stimuli, to elucidate whether rhythmic tracking also plays a role in establishing phonological representations. Rhyme ERPs were analysed in accordance with previous research for the pseudowords occurring at the line-final positions in the nursery rhyme stimuli.

SBC was observed for the stress and syllable rhythm of the nursery rhymes in bilateral fronto-temporal electrode clusters (Figure 6.2). Surprisingly, the reliable stress rhythm in metered nursery rhymes did not result in improved neural tracking. Instead, SBC was higher for non-metered nursery rhymes throughout the entire available frequency spectrum (1-7 Hz), with a significant difference emerging at the 3 Hz syllable frequency rate (Figure 6.3B). Infants' neural tracking was related to lexical development, as is evident from significant correlations between syllable rate SBC and vocabulary: a negative correlation with concurrent vocabulary at 10.5 months-old and a positive correlation at 18 months-old after accounting for concurrent vocabulary.

Rhyme ERPs for line-final rhymes occurring in each nursery rhyme were analysed in accordance with previous research (chapter 4 and 5). For the first occurrence of rhymes within each nursery rhyme (at the end of line 2, rhyme sensitivity effect) no significant difference between ERPs for rhyming versus non-rhyming pseudowords were found. Rhymes aggregated across line 3 and 4 (rhyme repetition) resulted in a sustained negativity. The polarity and timing of both ERP rhyme effects was in accordance with our previous study (chapter 4). Moreover, a tentative positive correlation between rhyme sensitivity and vocabulary at 18-months-old was found.

Discussion of SBC results

Infants' neural tracking at the delta and theta range (1-7 Hz) of the nursery rhymes converges with the sparse previous research that so far reports neural speech tracking for infants (Kalashnikova et al., 2018; Jessen et al., 2019; Choi et al., 2020; Lang et al., 2020; Menn et al., 2020; Snijders, 2020; Ortiz Barajas et al., 2021). Taken together, these results indicate that the prosodic envelope transmitted in low acoustic frequency bands is mirrored in infants' corresponding neural oscillations (delta/theta band). Here, a series of significant frequency clusters emerged from the available range from 1 to 7 Hz (Figure 6.2C, 1-1.25 Hz; 1.5-4.5 Hz, 5.75-6.5 Hz), roughly overlapping with the stress and syllable frequency of our stimuli (Figure 6.1). Stronger neural tracking within these clusters might reflect more detailed encoding of prosodic/phonological information provided at these frequencies, e.g., phrasal- and word-stress, syllable and phonological rate (Gross et al., 2013; Doelling, 2014; Leong & Goswami, 2015). Yet, arguing for a direct relationship between the frequency clusters found and the amplitude envelope of our stimuli requires a more thorough acoustic analysis of the stimuli and their association with infants' SBC values (Keitel et al., 2017; Menn et al., 2020).

Ortiz Barajas et al. (2021) suggest infants' phase tracking to be stable across development, whereas amplitude tracking shows a developmental effect. Stronger amplitude tracking was observed in six-month-olds for their native compared to non-native languages, while newborns tracked all studied languages equally well. Phase tracking, however, was similar for all languages in both age groups, suggesting it is not modulated by language experience (effect of language) or linguistic/cognitive maturation (though note that the effect of native/non-native language was marginally significant in newborns' phase tracking). In the current study infants' phase tracking – as measured with speech-brain coherence - was more precise for less rhythmic speech and an association of syllable tracking and lexical development was found. At the end of the first year of life, infants' phase tracking is thus modulated by top-down factors such as vocabulary size and bottom-up factors such as speech rhythm, suggesting that infants' phase tracking is more versatile than previously assumed. The differences in results might be attributable to the older infant age as well as the more varied stimulus materials used in the current study: Ortiz Barajas et al. presented each infant with one of three possible sentences per language, repeated 100 times in a block design, while in our study each infant listened to many different nursery rhymes. The larger individual differences in phase tracking observed here might thus at least in part be attributed to the more varied stimulus set.

Contrary to our hypothesis, non-metered nursery rhymes elicited higher SBC values than metered nursery rhymes across the whole available frequency spectrum (Figure 6.3A). While stress frequency was successfully manipulated, as indicated by the power spectrum (Figure 6.1), and SBC around the stress rate appeared to be numerically higher for metered nursery rhymes (Figure 6.3A), no significant difference between SBC for metered versus non-metered nursery rhymes was found at the stress specific frequency range. Instead, higher SBC for non-metered nursery rhymes was attested at the syllable rate.

The finding of overall less accurate neural phase tracking of metered nursery rhymes, especially around the syllable rate was unexpected. The regular meter found in songs and spoken nursery rhymes is generally associated with positive attentional effects in infant behaviour (Cirelli et al., 2019), and adults' neural tracking (DAT theory, but also Falk et al., 2017; Kayser et al., 2015; van den Bosch der Nederlanden et al., 2020). An explanation might lay in the specific implementation of the meter manipulation (see also discussion in chapter 5). In metered nursery rhymes, stress was alternating between adjacent syllables, resulting in a stress rate of 1.5 Hz. Yet, word stress was occurring in a wider frequency range around this stress rate, due to words varying in their number of syllables and position of lexical stress (see also Langus et al., (2017)). In this specific stimulus set, tracking the stress rate of the stimuli might therefore not have come with a processing benefit for the infant listeners.

Another, possibly related, explanation might lay in different amounts of attention dedicated to metered versus non-metered nursery rhymes. Possibly, infants start out with a rhythmic attending mode (Schroeder & Lakatos, 2009) on the metered nursery rhymes, due to their captivating rhythm and predictable structure. As the experiment continues, however, infants might shift into a continuous processing mode, originating from the less predictable, presumably more interesting non-metered nursery rhymes (Hunter & Ames, 1988; C. Kidd, Piantadosi, & Aslin, 2014). In general, neural tracking improves with active listening (Calderone, Lakatos, Butler, & Castellanos, 2014; Zion Golumbic, Poeppel, & Schroeder, 2012). In fact, the non-metered nursery rhymes might better align with speech in general, which has a non-isochronous rhythm and is therefore less predictable than rhythmic nursery rhymes. The more precise neural tracking of the less rhythmic stimulus observed here might thus be beneficial for processing of natural (ID)-speech. Future research could make use of a block design in combination with silent pauses between blocks (see e.g., Ortiz Barajas et al., 2021) to assess a rhythmic processing benefit. A block design might allow for differences in neural tracking for different types of stimuli to accumulate over longer stretches of time and might enable phase-reset between stimulus blocks of different experimental conditions.

Finally, another reason for more precise neural tracking of non-metered nursery rhymes might lie in the specific listening conditions in this study, as infants were exposed to the stimuli in a quiet environment. The processing benefit for more regularly structured nursery rhymes might only arise in adverse listening conditions, for example in noise (see e.g., van den Bosch der Nederlanden et al., 2020) or in children with hearing or language impairments.

Future research should take length of stimulus presentation into account in analysis of infant phase coherence, for example to explore differences between more versus less rhythmic stimuli over time as a result of auditory habituation. Choi et al., (2021) reported a rapid increase followed by a plateau in infants' syllable tracking in response to an artificial language already within the first 90 ms of stimulus exposure. Jessen et al., (2020) obtained reliable estimates of infant neural tracking from only 5 minutes of EEG recordings, indicating that shorter and therefore potentially less demanding experimental procedures are sufficient to gain reliable EEG indices of neural speech tracking.

Relationship between infants' speech tracking and their lexical development

Our study is among the first to report a relationship between infants' neural speech tracking and lexical processing abilities (Menn et al., 2020; Snijders, 2020; Choi et al., 2020). For concurrent vocabulary (10.5 months) a negative correlation was found with syllable rate SBC at left and fronto-central electrodes (Figure 6.5A). For later vocabulary (18 months) a positive correlation with syllable rate SBC was found after taking concurrent vocabulary into account (Figure 6.5B). The correlation at 10.5 months-old indicates a top-down influence of linguistic knowledge on auditory processing (Meyer, 2018; Rimmele et al., 2018): infants with larger vocabularies depend less on syllable tracking. These linguistically more mature infants need to rely less on acoustic stimulus features when listening to speech as they have more lexical representations onto which they can map the incoming speech signal. This interpretation is backed by higher phase coherence values in adults processing a second compared to their native language (Song & Iverson, 2018), suggesting that an increase in coherence reflects higher processing effort and auditory attention, which in turn is modulated by concurrent linguistic knowledge.

Infants who focus more on syllables while processing speech at 10.5 months, have smaller vocabularies around that age, but larger vocabularies halfway through their second year. Thus, a syllabic processing focus at 10.5-months-old might come with a lexical disadvantage at this age (smaller vocabularies). Later on, however, having spent more time on encoding of syllables around 10.5 months old might pay off in more sophisticated phonological and lexical processing abilities, and thus larger vocabularies at 18-months-old. Note that a reversed interpretation of these effects might be intriguing as well: those infants who have larger early vocabularies might focus more on whole lexical items, and less on syllables, while processing speech. This might give them a disadvantage later on, because they develop less sophisticated phonological representations.

Hemispheric differences in SBC for metered versus non-metered nursery rhymes

Hemispheric differences in SBC values were not tested explicitly, but will be discussed briefly for potential future research interest. Coherence clusters in non-metered nursery rhymes had a left frontal peak of neural activity (Figure 6.3B), a cortical region that is typically associated with linguistic processing in infants (see e.g., Altvater-Mackensen & Grossmann, 2016) and which has previously been reported for infants' phase tracking of rhythmic stories (Snijders, 2020) and the theta rhythm of speech (Kalashnikova et al., 2018). Moreover, left-frontal activity has been associated with top-down processes influencing neural tracking of the delta and theta rhythm in adult listeners (Kayser et al., 2015; Park et al., 2015). Infants' phase coherence for metered nursery rhymes, however, was more broadly distributed across both hemispheres and is in line with a widely distributed fronto-central activation pattern reported for adults listening to rhythmic songs (van den Bosch der Nederlanden et al., 2020). Future research should elaborate on the potential relevance of lateralization differences in neural tracking in response to more versus less rhythmic speech, especially with regards to impaired language acquisition (Lallier et al., 2018).

Rhyme ERPs and relationship with regular meter, phase coherence and lexical development

ERP rhyme effects only partially rendered significant results. At the first moment rhyming and non-rhyming nursery rhymes differed (rhyme sensitivity effect), rhymes elicited a numerically but not significantly more positive ERP than non-rhymes between 200-500 ms. Polarity and latency resemble the tentative rhyme sensitivity effect in rhyming songs (chapter 4). Whether infants already recognize the first instance of rhymes in poetic speech thus requires further corroboration by future research, ideally with more statistical power.

For rhymes recurring at subsequent lines of the nursery rhymes (rhyme repetition), a sustained negativity for rhymes was observed, with a significant difference between ERPs in response to rhymes versus non-rhymes attested from 0 to 200 ms (time window was based on chapter 4). These results again parallel the effects reported for repeating rhymes and entire words in child songs (chapter 4; Snijders et al., 2020). In adults, both rhyme sensitivity and rhyme repetition effects emerged much later, around 400 ms after onset of the line-final pseudoword (see chapter 5), rendering it difficult to compare infants' developing implicit rhyme abilities with mature responses. There was no evidence for regular speech rhythm influencing infants' rhyme perception, as SBC was not related to rhyme ERPs, nor did rhyme ERPs differ between metered versus non-metered nursery rhymes. Absence of evidence for regular meter affecting infants' rhyme processing might again be attributable to low statistical power, as even the meter effects on adult rhyme processing were subtle (see the discussion section in chapter 5 for possible reasons).

ERP rhyme repetition effects elicited by rhyming songs were negatively correlated with productive vocabulary in the comparable previous study (chapter 4). Here, in contrast, ERP rhyme sensitivity was positively correlated with 18-months comprehension scores. In both cases, the correlation was only marginally significant and was not based on robust ERP rhyme effects. The question of how early implicit rhyme processing relates to lexical development thus still requires future research.

Conclusion

Neural tracking of the low frequency speech amplitude envelope provides infants with a temporal processing constraint while deriving lexical and phonological information from the speech stream. In this study, 10.5-month-olds closely followed the stress- and syllable rhythm of spoken nursery rhymes, and cortical tracking of the syllable rate correlated with concurrent and later vocabulary. Moreover, ERP rhyme effects resembled previous tentative effects reported for rhyming songs (chapter 4). No processing benefit for regularly metered nursery rhymes was found for rhyme perception or neural tracking. Instead, neural tracking was higher for non-metered nursery rhymes, suggesting that infants' envelope phase tracking during the second half of the first year of life is well equipped to cope with rhythmically challenging ID-speech. These results indicate interactive effects of bottom-up acoustic stimulus features (here speech rhythm) and top-down linguistic processing abilities (here vocabulary size) on infants' cortical tracking, furthering our understanding on low frequency speech tracking during the development of lexical and phonological representations.

7

Chapter 7

General discussion

The present dissertation provided evidence for infants' recognition of phonological and prosodic sound patterns in oral language play. So far, the focus of previous research has primarily been on the socio-emotional benefit of oral language play during early childhood, especially the effect of ID-song on infant attention and arousal. Previous research has also elaborated on the rich structural cues provided in songs and nursery rhymes and its potential role in auditory scaffolding and second language vocabulary acquisition (see references at the beginning of the General Introduction). Yet, only few studies so far had asked whether infants actually process the linguistic information provided in language play, which was the topic of this dissertation. In addition, the results presented here also offered suggestions for how recognition of rhymes and rhythm in language play might relate to future lexical development and emerging phonological skills. None of our studies required training or extensive previous exposure to language play. Rather, our paradigms allowed for spontaneous processing. Moreover, naturalistic songs and nursery rhymes were used throughout all studies. This approach should encourage generalizability of our findings towards naturalistic infant-caregiver interactions and might inform future research with clinical, developmental or other vulnerable populations.

Detecting sound patterns in oral language play - summary of findings

Chapters 3, 4 and 6 delivered tentative evidence that **infants process rhymes in songs and spoken nursery rhymes** using behavioural as well as neurophysiological paradigms. In chapter 3, infants' discrimination between rhyming and non-rhyming songs was indicated by overall longer looking times to non-rhyming songs in the Headturn Preference Procedure. Moreover, repeating rhymes in songs (chapter 4) and spoken nursery rhymes (chapter 6) resulted in a central negativity for rhymes compared to non-rhymes. First occurrences of rhymes (compared to non-rhymes; rhyme sensitivity) elicited a left-frontal positivity, in both songs (chapter 4) and spoken nursery rhymes (chapter 6). Both the ERP and the behavioural results require stronger statistical corroboration, as only for rhyme repetition in spoken nursery rhymes a significant effect was found, while all other effects were marginally significant. Despite the many questions that remain open from these studies, the research base has been enriched by ERP rhyme effects from naturalistic stimuli and a task-free paradigm, rendering our results potentially relevant for clinical and developmental populations.

In chapter 5, the developmental target for the infant ERP rhyme effects was reported in an adult study using the exact same nursery rhyme stimuli as in chapter 6. This study revealed that **adult listeners implicitly perceive line-final rhymes in spoken nursery rhymes and rhyme perception is only subtly modulated by regular meter.** Adults' ERP rhyme responses resembled a delayed posterior phonological N400 effect, previously reported for rhymes in single word priming studies. The study extends previous accounts of the N400 rhyme effect to poetic pseudoword rhymes, which were reliably detected even in the rather shallow passive processing conditions of our paradigm. Moreover, an early anterior negativity for rhymes was observed, which might reflect working memory involvement in perceptually binding rhyming verse lines. Regular meter, a prototypical feature of poetry, only subtly influenced rhyme perception, as the anterior rhyme effect occurred slightly earlier in non-metered verses. This facilitating effect of irregular meter is surprising given previous accounts of rhythmic benefits for phonological processing. We will briefly reiterate possible reasons for this counterintuitive effect below, where we combine these findings with the results from the complementary infant study (chapter 6).

In chapter 2, a behavioural study attested that **infants can segment songs into phrases**, replicating and extending previous research on ID-speech. Our results suggest that six-month-olds are capable of domain-general parsing, as they processed the same acoustic cues to segment speech as well as song into the respective underlying phrases. This extension of the prosodic parsing account onto a new domain of infant-caregiver communication indicates that infants derive structure from song melodies and lyrics, an important prerequisite for syntactic development.

Finally, chapter 6 provides evidence that **infants track the speech rhythm of nursery rhymes**. Infants' neural activity mirrored the stress and syllable rhythm of the nursery rhymes and this resemblance was particularly strong for non-metered nursery rhymes. Infants' more accurate encoding of the less regular speech rhythm might resemble the facilitated rhyme perception in non-metered stimuli observed in adults (chapter 5). In both cases, processing benefits for less regular stimuli were attributed to a shift from a rhythmic to a continuous processing mode. This mode shift might have evolved over the course of the experiment and potentially originates from a perceptual generalization from the typically non-isochronous rhythm of (ID)-speech.

Relationship with general language development

Infants' perception of rhymes in language play correlates with concurrent and later vocabulary size. This link was attested for the ERP rhyme effects reported for both songs (chapter 4) and spoken nursery rhymes (chapter 6), but both require confirmation by future research. In chapter 4, we elaborated in more detail on the potential of language play for boosting early implicit rhyme awareness during infancy and a possible influence thereof on future language outcomes. Below these suggestions will be combined into a more general framework.

Infants' neural tracking of the syllable rhythm in nursery rhymes has a negative relationship with concurrent vocabulary (10.5 months), but is positively correlated with future vocabulary (18 months). A syllabic processing focus at the end of the first year of life might thus be associated with different effects on concurrent and future vocabulary size. In chapter 6 we suggested possible consequences of this finding.

Limitations

Limitations were mainly discussed in respective chapters, but a few general points will be reiterated here. From a methodological point of view, **the use of pseudowords should be critically reviewed**. The entire song lyrics of chapters 3 and 4 and the line-final target words in chapters 3 to 6 consisted of legal Dutch pseudowords. This design decision enabled control for phonological features of these items (rime frequency, CVC structure) and allowed us to use the same target words for rhyming and non-rhyming stimuli. Yet, the constant use of unfamiliar words might have increased fatigue in the infant and adult (chapter 5) listeners and result in a shift of attention away from the experimental stimulus.

At the age of testing (between 9 and 11 months old), infants already know many highly frequent words (Johnson, 2016) and are thus used to encountering these familiar words in the speech input. Highly familiar words might function as lexical anchors in speech processing (Frost, Dunn, Christiansen, Gómez, & Monaghan, 2020; Frost, Monaghan, & Christiansen, 2019), binding infants' attention to the stimulus and motivating them to explore the signal even further. For example, 9-month-olds benefit in their word segmentation from words that overlap in rimes with highly familiar words (Altvater-Mackensen & Mani, 2013). The same rationale applies to sound sequences below the word unit: at 11-months-old, highly frequent phoneme patterns (that cross word boundaries) are differentiated from low-frequency patterns, indicating storage of words as well as non-words containing these highly-frequent sound patterns in infants' proto-lexicons (Ngon et al., 2013). In chapters 3 and 4, lexical anchors were limited to function words occurring in the song lyrics. The remainder of the speech signal consisted of unfamiliar pseudowords, potentially resulting in attentional decay and increased data loss. Including a few highly familiar words as line-final target words in chapters 3, 4 and 6 might have provided us with more reliable estimates of infants' rhyme processing.

Another methodological limitation might be **the use of audio-only stimuli**. In the behavioural studies reported in chapter 2 and 3, infants were fixating at blinking lights during stimulus exposure. In chapters 4 and 6, stimulus-unrelated silent cartoons were presented on a TV screen during stimulus presentation and

EEG recording. While exclusion rates were not exceptionally high in any of the infant studies reported here, they might have been even lower by using audio-visual stimuli (Junge et al., 2020; van der Velde & Junge, 2020) and this might again have made for more robust effects and would strengthen the ecological validity of our approach even more. Adults' speech perception is known to benefit from the complementary information of both auditory and visual cues (Ross, Saint-Amour, Leavitt, Javitt, & Foxe, 2007). Captivating multimodal stimuli might allow for inter-sensory redundancy (Bahrick & Lickliter, 2014), for example, caregivers' facial cues in language play reinforce the hierarchical structure (e.g., Longhi, 2009) and contingent eye-gaze between infant and caregiver during language play has been suggested to allow for increased encoding of the stimulus in infant listeners (Leong, Byrne, et al., 2017).

The current studies only pertained to infants' perceptual processing abilities and how they possibly interact with concurrent and future lexical development. For a full-fledged model of the potential contribution of early language play to (first) language acquisition, **more background information about at-home use of music and language play as well as a more heterogeneous infant sample is necessary**. Such an extension of our analyses would also be in accordance with recent calls for multi-factorial models to predict language outcome (Cristia et al., 2014; Seidl, French, Wang, & Cristia, 2014). In the ERP studies reported in chapter 4 and 6, parents filled in an extensive questionnaire on their SES, musical education, use of music and rhythms at home and how often they read to their children, but these data have not been analysed for this dissertation. Such additional analysis would be valuable, for example to elucidate whether infants with more at-home exposure to rhythmic language play show different neural tracking of rhythmic and less rhythmic speech compared to infants who receive less language play input. Individual differences in neural tracking might in turn be associated with infants' recognition of structural units in speech, for example phrase boundaries, rhymes or stressed syllables.

Future research

A number of studies report improved phonological processing (*change detection* specifically) in infants exposed to sung or rhythmic speech streams compared to (ID-)speech (François et al., 2017; Lebedeva & Kuhl, 2010; Suppanen et al., 2019; Thiessen & Saffran, 2009; Yamane, Sato, Shimura, & Mazuka, 2021). Rhythmic ID-speech in the form of songs and spoken nursery rhymes might be considered an acoustically super-stimulus, due to its slow pace, repetitiveness, discrete pitch steps and isochronous rhythm (see references in respective chapters). The picture emerging from this dissertation, however, is different. **We found no processing benefit for songs or rhythmic nursery rhymes** for infants' phrase segmentation (chapter 2), rhyme perception or neural tracking (chapter 6 for both). Adults' rhyme perception (chapter 5) was also only subtly influenced by a regular speech rhythm, with an indication of processing facilitation in non-metered nursery rhymes. Infants and adults showed more engaged processing of the non-metered stimulus (chapter 5 and 6). The effects observed here might in fact not be specific to oral language play. A potential perceptual benefit of this communicative register compared to typical (ID)-speech should be studied more thoroughly in future studies (see discussion sections in the respective chapters). These studies should also examine a wider age range and infants with more diverse linguistic abilities.

Phonological processing becomes more language specific towards the end of the first year of life (Gervain & Mehler, 2010), which could result in processing advantages for stimuli that better align with ordinary speech, as observed in chapter 6. Moreover, processing prosodic and phonological information in songs and nursery rhymes might play a different role in infants with regular versus impaired language acquisition. For the former, language play might represent a speech processing playground, which fosters communicative exchange during the pre-verbal stage. For the latter population, the same might apply but then on top of a processing benefit at work in language play, compared to the more challenging listening conditions imposed by speech outside of playful contexts.

This dissertation provided suggestions on the role of oral language play for infants between 6 months and 1.5 years old, in particular with regards to their phonological and lexical development. **Future research should now close the gap between early processing of language play and emerging phonological and lexical skills**, especially the underlying abilities of early literacy. One target theoretical concept to bridge infants' perception of sound patterns in language play and future language development beyond the first and second year of life might be *lexical specificity*: "the richness and specificity of, and distinctness between, phonological representations in the emerging mental lexicon" (van Goch, McQueen, & Verhoeven, 2014). As infants' vocabularies grow larger over time, lexical entries need to become more and more specific with regards to their phonological structure, in order to allow for differentiation of similar sounding words (Fikkert & Levelt, 2008; Metsala & Walley, 1998). For example, the word *bear* has to be differentiated from *pear*, which shares the same syllable rime, but differs in onset consonant.

Caregivers' rhythmic language play might reinforce lexical specificity, by highlighting the hierarchical, phonological and rhythmic structure of speech. For example, rhyming words are a frequent phenomenon in language play, but otherwise occur only rarely in infants' input. Onset-rime differentiation in rhyming words might be supported by language play, where rhymes occur in salient line-final positions. Lexical specificity has been suggested to allow for transfer from infants' implicit unconscious perception of linguistic units in speech to later explicit phonological skills (Janssen, Segers, McQueen, & Verhoeven, 2017). This transfer might build upon the neural tracking of the amplitude modulations in speech, as these were suggested to be particularly salient in rhythmic ID-speech such as songs and nursery rhymes (Falk & Kello, 2017; Goswami, 2019b; Leong & Goswami, 2015). Rhythmic structure in language play might enable infants to enter a rhythmic processing mode, which in turn fosters build-up of specific phonological representations. Appealing as this account may be, empirical evidence so far is lacking. The results provided in chapter 6 indicate that 10.5-month-old infants do not show enhanced neural tracking of rhythmic speech. The lexical specificity hypothesis with regards to rhythmic language play thus requires future research.

So far, emerging explicit phonological and lexical skills were mostly attributed to joined book reading and other informal language activities at home (Krijnen et al., 2020). This dissertation in turn suggests a potential contribution of oral language play to early phonological and lexical skills**.** Recent studies as well suggest a positive relationship between early informal musical activities, especially caregiver singing, and linguistic development in regular (Franco et al., 2021; Politimou et al., 2019) as well as impaired language acquisition (Torppa, Faulkner, Kujala, Huotilainen, & Lipsanen, 2018; Virtala & Partanen, 2018). Infants' processing of songs and nursery rhymes might thus be another stepping stone in transfer of implicit unconscious speech processing into explicit linguistic skills, which in turn underlie literacy.

The results reported here might only apply to the specific cultural tradition of poetry in Germanic languages. As outlined in chapter 1, poetry combines both universal as well as language-specific phonological and lexical properties of ordinary language (Kiparsky, 1973; Fabb, 2010). Germanic languages could be considered rime-based languages, as they contain many minimal pairs that overlap in syllable rime, resulting in a higher proportion of rime neighbours compared to neighbours overlapping in other phonological units (De Cara & Goswami, 2002; Martensen, Maris, & Dijkstra, 2000). Language play by Dutch caregivers might thus overemphasize the onset-rime distinction, due to the saliency of rimes in the Dutch lexicon. Future studies should incorporate data from a more diverse language sample, ideally from non-WEIRD societies, to resolve whether other phonological units are highlighted in language play and whether these units in turn are recognized by infant listeners.

Given the large individual and cross-cultural differences in caregivers use of (ID)-speech in general (Bergelson et al., 2019; Casillas, Brown, & Levinson, 2020) and poetic language in particular, **infants will need to be able to arrive at full linguistic competence even without song and verse input.** A reader playing devil's advocate might now be inclined to repeat the infamous quote from Steven Pinker, who compared music with "auditory cheesecake", as something that is emotionally pleasant but not necessary from an evolutionary perspective (Pinker, 1997, p. 528). Language play as well might thus be considered pleasant and rewarding for both infants and caregivers, yet not providing a unique contribution to linguistic development of children. Just like in the case of music, such critique is inconsistent with the ubiquity of songs and nursery rhymes in caregiving contexts across the world (Ilari, 2005; Ilari et al., 2011; Markova, 2018). Undoubtedly, the functional relevance of language play for early language acquisition requires further research, but the wide-spread use across the worlds' languages warrants a better understanding of infants' linguistic processing of language play.

Conclusion

The studies summarized in this dissertation represent evidence for the perception of rhyme, rhythm and phrase boundaries in oral language play in infants and adults. Our experimental studies converge in the notion that the linguistic information transmitted in oral language play is not merely a sea of sound for the infant (and adult) listener. Instead, infants are sensitive to the phonological, structural and rhythmic patterns transmitted in the acoustic shape of oral language play and this might have consequences for their language development.

References

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(23), 13367–13372. https://doi.org/10.1073/pnas.201400998

Altvater-Mackensen, N., & Grossmann, T. (2016). The role of left inferior frontal cortex during audiovisual speech perception in infants. *NeuroImage*, *133*, 14–20. https://doi.org/10.1016/j.neuroimage.2016.02.061

Altvater-Mackensen, N., & Mani, N. (2013). Word-form familiarity bootstraps infant speech segmentation. *Developmental Science*, *16*(6), 980–990. https://doi.org/10.1111/desc.12071

Andersson, A., Sanders, L. D., Coch, D., Karns, C. M., & Neville, H. J. (2018). Anterior and posterior erp rhyming effects in 3- to 5-year-old children. *Developmental Cognitive Neuroscience*, *30*(February), 178–190. https://doi.org/10.1016/j.dcn.2018.02.011

Arleo, A. (2006). Do children's rhymes reveal universal metrical patterns? In P. Hunt (Ed.), *Children's Literature: Critical Concepts in Literary and Cultural Studies*: *Vol. IV* (pp. 39–56). London, UK: Routledge. Retrieved from http://halshs.archives-ouvertes.fr/halshs-00637386/

Baayen, H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database. Linguistic data consortium.* University of Pennsylvania, Philadelphia.

Baddeley, A. D., Lewis, V. J., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology*, *36A*, 233–252. https://doi.org/10.4324/9781315111261

Bahrick, L. E., & Lickliter, R. (2014). Learning to Attend Selectively: The Dual Role of Intersensory Redundancy. *Current Directions in Psychological Science*, *23*(6), 414–420. https://doi.org/10.1177/0963721414549187

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bebout, J., & Belke, E. (2017). Language play facilitates language learning: Optimizing the input for gender-like category induction. *Cognitive Research: Principles and Implications*, *2*(1), 11. https://doi.org/10.1186/s41235-016-0038-z

Beier, E. J., & Ferreira, F. (2018). The temporal prediction of stress in speech and its relation to musical beat perception. *Frontiers in Psychology*, *9*(431), 1–6.

Bell, A. J., & Sejnowski, T. J. (1995). An Information-Maximisation Approach to Blind Separation and Blind Deconvolution. *Neural Computation*, *7*, 1129–1159. https://doi.org/doi:10.1162/neco.1995.7.6.1129

Bergelson, E., Amatuni, A., Dailey, S., Koorathota, S., & Tor, S. (2019). Day by day, hour by hour: Naturalistic language input to infants. *Developmental Science*, *22*(1), 1–10. https://doi.org/10.1111/desc.12715

Bergeson, T. R., & Trehub, S. E. (2002). Absolute pitch and tempo in mothers' songs to infants. *Psychological Science*, *13*(1), 72–75. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11892783

Bergmann, C., & Cristia, A. (2016). Development of infants' segmentation of words from native speech: a meta-analytic approach. *Developmental Science*, *19*(6), 901–917. https://doi.org/10.1111/desc.12341

Bergmann, C., Piccinini, P. E., Lewis, M. L., Frank, M. C., & Cristia, A. (2018). Promoting Replicability in Developmental Research Through Meta-analyses : Insights From Language Acquisition Research. *Child Development*, *00*(0), 1–14. https://doi.org/10.1111/cdev.13079

Böcker, K. B. E., Bastiaansen, M. C. M., Vroomen, J., Brunia, C. H. M., & De Gelder, B. (1999). An ERP correlate of metrical stress in spoken word recognition. *Psychophysiology*, *36*(6), 706–720.

Boersma, P., & Weenink, D. (2014). *Praat: doing phonetics by computer*. Retrieved from www.praat.org

Bohn, K., Knaus, J., Wiese, R., & Domahs, U. (2013). The influence of rhythmic (ir)regularities on speech processing: Evidence from an ERP study on German phrases. *Neuropsychologia*, *51*(4), 760–771. https://doi.org/10.1016/j.neuropsychologia.2013.01.006

Bradley, L., & Bryant, P. (1983). Categorizing sounds and learning to read. *Nature*, *21*(5899), 159–421. https://doi.org/10.1038/301419a0

Braze, D., McRoberts, G., & Mc Donough, C. (2011). Early pre-cursors of reading-relevant phonological skills. In S. Brady, D. Braze, & C. A. Fowler (Eds.), *Explaining individual differences in reading: Theory and evidence* (pp. 23–44). Psychology Press.

Breen, M. (2018). Effects of metric hierarchy and rhyme predictability on word duration in The Cat in the Hat. *Cognition*, *174*(December 2016), 71–81. https://doi.org/10.1016/j.cognition.2018.01.014

Breen, M., & Fitzroy, A. B. (2019). Event-Related Potential Evidence of Implicit Metric Structure during Silent Reading. *Brain Sciences*, *9*(192).

Brentari, D., González, C., Seidl, A., & Wilbur, R. (2010). Sensitivity to Visual Prosodic Cues in Signers and Nonsigners. *Language and Speech*, *54*(1), 49–72. https://doi.org/10.1177/0023830910388011

Bryant, P., & Bradley, L. (1978). Difficulties in auditory organisation as a possible cause of reading backwardness. *Nature*, *271*, 746–747.

Bryant, P., Bradley, L., MacLean, M., & Crossland, J. (1989). Nursery rhymes, phonological skills and reading. *Journal of Child Language*, *16*(02), 407. https://doi.org/10.1017/S0305000900010485

Bryant, P., Maclean, M., & Bradley, L. (1990). Rhyme, language, and children's reading. *Applied Cognitive Psychology*, *11*(1990), 237–252.

Burling, R. (1966). The Metrics of Children's Verse: A Cross-Linguistic Study. *American Anthropologist*, *68*, 1418–1441.

Calderone, D. J., Lakatos, P., Butler, P. D., & Castellanos, F. X. (2014). Entrainment of neural oscillations as a modifiable substrate of attention. *Trends in Cognitive Sciences*, *18*(6), 300–309. https://doi.org/10.1016/j.tics.2014.02.005

Carroll, J. M., Snowling, M. J., Stevenson, J., & Hulme, C. (2003). The development of phonological awareness in preschool children. *Developmental Psychology*, *39*(5), 913–923. https://doi.org/10.1037/0012-1649.39.5.913

Carvalho, A. De, Dautriche, I., Millotte, S., & Christophe, A. (2018). Early perception of phrasal prosody and its role in syntactic and lexical acquisition. In P. Pilar & N. Esteve-Gibert (Eds.), *The Development of Prosody in First Language Acquisition*. Amsterdam / Philadelphia: John Benhamins Publishing Company (TILAR Series).

Casillas, M., Brown, P., & Levinson, S. C. (2020). Early language experience in a Papuan community. *Journal of Child Language*, 1–23. https://doi.org/10.1017/S0305000920000549

Casteren, M. van, & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods*, *38*(4), 584–589. https://doi.org/10.3758/BF03193889

Chen, Q., Zhang, J., Xu, X., Scheepers, C., Yang, Y., & Tanenhaus, M. K. (2016). Prosodic expectations in silent reading: ERP evidence from rhyme scheme and semantic congruence in classic Chinese poems. *Cognition*, *154*, 11–21. https://doi.org/10.1016/j.cognition.2016.05.007

Choi, D., Batterink, L. J., Black, A. K., Paller, K. A., & Werker, J. F. (2020). Preverbal Infants Discover Statistical Word Patterns at Similar Rates as Adults: Evidence From Neural Entrainment. *Psychological Science*, *31*(9), 1161–1173. https://doi.org/10.1177/0956797620933237

Chow, M., Macnamara, B. N., & Conway, A. R. A. (2016). Phonological similarity in working memory span tasks. *Memory & Cognition*, *44*, 937–949.

Cirelli, L. K., Jurewicz, Z. B., & Trehub, S. E. (2019). Effects of Maternal Singing Style on Mother–Infant Arousal and Behavior. *Journal of Cognitive Neuroscience*, *24*(6), 1275–1285. https://doi.org/10.1162/jocn

Cirelli, L. K., Trehub, S. E., & Trainor, L. J. (2018). Rhythm and melody as social signals for infants. *Annals of the New York Academy of Sciences*, 1–7. https://doi.org/10.1111/nyas.13580

Coch, D., Grossi, G., Coffey-Corina, S., Holcomb, P. J., & Neville, H. J. (2002). A developmental investigation of ERP auditory rhyming effects. *Developmental Science*, *4*, 467–489. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/1467-7687.00241/abstract

Coch, D., Grossi, G., Skendzel, W., & Neville, H. (2005). ERP nonword rhyming effects in children and adults. *Journal of Cognitive Neuroscience*, *17*(1), 168–182. https://doi.org/10.1162/0898929052880020

Coch, D., Hart, T., & Mitra, P. (2008). Three kinds of rhymes: An ERP study. *Brain and Language*, *104*(3), 230–243. https://doi.org/10.1016/j.bandl.2007.06.003

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ, USA: Lawrence Erlbaum Associates.

Conway, C. M., Pisoni, D. B., & Kronenberger, W. G. (2009). The Importance of Sound for Cognitive Sequencing Abilities: The Auditory Scaffolding Hypothesis. *Current Directions in Psychological Science*, *18*(5), 275–279. https://doi.org/10.1111/j.1467-8721.2009.01651.x

Cook, G. (1997). Language play, language learning. *ELT Journal*, *51*(July), 224–231.

Corbeil, M., Trehub, S. E., & Peretz, I. (2013). Speech vs. singing: infants choose happier sounds. *Frontiers in Psychology*, *4*(June), 372. https://doi.org/10.3389/fpsyg.2013.00372

Corbeil, M., Trehub, S. E., & Peretz, I. (2016). Singing Delays the Onset of Infant Distress. *Infancy*, *21*(3), 373–391. https://doi.org/10.1111/infa.12114

Costa-Giomi, E. (2014). Mode of Presentation Affects Infants' Preferential Attention to Singing and Speech. *Music Perception*, *32*(2), 160–169. Retrieved from http://www.jstor.org/stable/10.1525/mp.2014.32.2.160

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain potentials to morphosyntactic violations. *Language and Cognitive Processes*, *13*(1), 21–58.

Cristia, A. (2013). Input to Language: The Phonetics and Perception of Infant-Directed Speech. *Linguistics and Language Compass*, *7*(3), 157–170. https://doi.org/10.1111/lnc3.12015

Cristia, A., Seidl, A., Junge, C., Soderstrom, M., & Hagoort, P. (2014). Predicting individual variation in language from infant speech perception measures. *Child Development*, *85*(4), 1330–1345. https://doi.org/10.1111/cdev.12193

Cross, K., & Fujioka, T. (2019). Auditory rhyme processing in expert freestyle rap lyricists and novices: An ERP study. *Neuropsychologia*, *129*(March), 223–235. https://doi.org/10.1016/j.neuropsychologia.2019.03.022

Csibra, G., Hernik, M., Mascaro, O., Tatone, D., & Lengyel, M. (2016). Statistical Treatment of Looking-Time Data. *Developmental Psychology*, *52*(4), 521–536.

Curtin, S., & Zamuner, T. S. (2014). Understanding the developing sound system: Interactions between sounds and words. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(5), 589–602. https://doi.org/10.1002/wcs.1307

Custodero, L. A., & Johnson-Green, E. A. (2008). Caregiving in counterpoint: Reciprocal influences in the musical parenting of younger and older infants. *Early Child Development and Care*, *178*(1), 15–39. https://doi.org/10.1080/03004430600601115

Custodero, L. A., Rebello Britto, P., & Brooks-Gunn, J. (2003). Musical lives: A collective portrait of American parents and their young children. *Journal of Applied Developmental Psychology*, *24*(5), 553–572. https://doi.org/10.1016/j.appdev.2003.08.005

Cutler, A. (1994). Segmentation problems, rhythmic solutions. *Lingua*, *92*, 81–104. https://doi.org/10.1016/0024-3841(94)90338-7

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, *2*(3–4), 133–142. https://doi.org/10.1016/0885-2308(87)90004-0

Davids, N., van den Brink, D., van Turennout, M., & Verhoeven, L. (2011). Task-related influences on neurophysiological assessment of auditory rhyme: implications for clinical studies. *Clinical Neurophysiology*, *122*(8), 1629–1636. https://doi.org/10.1016/j.clinph.2010.12.058

Davis, G. M. (2017). Songs in the young learner classroom : a critical review of evidence. *ELT Journal*, *71*(July 2018), 445–455. https://doi.org/10.1093/elt/ccw097

De Cara, B., & Goswami, U. (2002). Similarity relations among spoken words: the special status of rimes in English. *Behavior Research Methods, Instruments, & Computers*, *34*(3), 416–423. https://doi.org/10.3758/BF03195470

De Cara, B., & Goswami, U. (2003). Phonological neighbourhood density: Effects in a rhyme awareness task in five-year-old children. *Journal of Child Language*, *30*(3), 695–710. https://doi.org/10.1017/S0305000903005725

de Carvalho, A., Dautriche, I., & Christophe, A. (2017). Phrasal prosody constrains online syntactic analysis in two-year-old children. *Cognition*, *163*, 67–79. https://doi.org/10.1016/j.cognition.2017.02.018

De Diego Balaguer, R., Martinez-Alvarez, A., & Pons, F. (2016). Temporal Attention as a Scaffold for Language Development. *Frontiers in Psychology*, *7*(February), 1–15. https://doi.org/10.3389/fpsyg.2016.00044

DeCastro-Arrazola, V., & Kirby, S. (2018). The emergence of verse templates through iterated learning. *Journal of Language Evolution*, *4*(1), 28–43. https://doi.org/10.1093/jole/lzy013

DeCastro-Arrazola, V., Kranenburg, P. Van, & Janssen, B. (2015). Computational textsetting analysis of Dutch folk songs. In O. Adam & D. Cazau (Eds.), *Proceedings of the 5th International Workshop on Folk Music Analysis* (pp. 51–55). Paris: Association Dirac.

Delavenne, A., Gratier, M., & Devouche, E. (2013). Expressive timing in infant-directed singing between 3 and 6 months. *Infant Behavior & Development*, *36*(1), 1–13. https://doi.org/10.1016/j.infbeh.2012.10.004

Dell, F., & Halle, J. (2009). Comparing musical textsetting in French and in English songs. *Aroui & Arleo (Eds.)*. Retrieved from http://books.google.com/books?hl=en&lr=&id=FSA95IXWjaAC&oi=fnd&pg=PA63&dq=Comparing+musical+textsetting+in+French+and+English+songs&ots=tAW7EBFO3v&sig=1p_TJmizSqN9P7Tz65Mznu6soDM

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *13*, 9–21. https://doi.org/http://dx.doi.org/10.1016/j.jneumeth.2003.10.009

Desroches, A. S., Newman, R. L., Robertson, E. K., & Joanisse, M. F. (2013). Electrophysiological indices of phonological impairments in dyslexia. *Journal of Speech, Language, and Hearing Research*, *56*(1), 250–264. https://doi.org/10.1044/1092-4388(2012/10-0351)

Deutsch, D., & Feroe, J. (1981). The internal Representation of Pitch Sequences in tonal Music. *Psychological Review*, *88*(6), 503–522. https://doi.org/10.1037/0033-295X.88.6.503

Dissanayake, E. (2000). Antecedents of the temporal arts in early mother-infant interaction. In N. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 389–410). Cambridge, MA: MIT Press.

Doelling, K., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Speech Comprehension By Facilitating Perceptual Parsing. *Neuroimage*, *85*(15). https://doi.org/10.1016/j.neuroimage.2013.06.035.Acoustic

Domahs, U., Wiese, R., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). The processing of German word stress: Evidence for the prosodic hierarchy. *Phonology*, *25*(1), 1–36. https://doi.org/10.1017/S0952675708001383

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition*, *77*(3), 251–288. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11018511

Dumay, N., Benraïss, A., Barriol, B., Colin, C., Radeau, M., & Besson, M. (2001). Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *Journal of Cognitive Neuroscience*, *13*(1), 121–143. https://doi.org/10.1162/089892901564117

Dunst, C., Meter, D., & Hamby, D. (2011). Relationship between young children's nursery rhyme experiences and knowledge and phonological and print-related abilities. *Center for Early Literacy Learning Reviews*, *4*(1), 1–12. Retrieved from http://earlyliteracylearning.org/cellreviews/cellreviews_v4_n1.pdf

Endress, A. D., & Hauser, M. D. (2010). Word segmentation with universal prosodic cues. *Cognitive Psychology*, *61*(2), 177–199. https://doi.org/10.1016/j.cogpsych.2010.05.001

Engh, D. (2013). Why use music in English language learning? a survey of the literature. *English Language Teaching*, *6*(2), 113–127. https://doi.org/10.5539/elt.v6n2p113

Fabb, N. (1999). Verse constituency and the locality of alliteration. *Lingua*, *108*, 223–245.

Fabb, N. (2010). Is literary language a development of ordinary language? *Lingua*, *120*(5), 1219–1232. https://doi.org/10.1016/j.lingua.2009.07.007

Fabb, N. (2014). The Verse-Line as a Whole Unit in Working Memory, Ease of Processing, and the Aesthetic Effects of Form. *Royal Institute of Philosophy Supplement*, *75*(1995), 29–50. https://doi.org/10.1017/s1358246114000186

Fabb, N. (2015). Verse. In V. Sotirova (Ed.), *The Bloomsbury Companion to Stylistics* (pp. 523–536). London, UK: Bloomsbury.

Falk, S., & Kello, C. T. (2017). Hierarchical organization in the temporal structure of infant-direct speech and song. *Cognition*, *163*, 80–86. https://doi.org/10.1016/j.cognition.2017.02.017

Falk, S., Lanzilotti, C., & Schön, D. (2017). Tuning Neural Phase Entrainment to Speech. *Journal of Cognitive Neuroscience*, (December), 139. https://doi.org/10.1162/jocn

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17695343

Fikkert, P., & Levelt, C. (2008). How does Place fall into Place? The lexicon and emergent constraints in children's developing phonological grammar. In P. Avery, B. E. Dresher, & K. Rice (Eds.), *Contrast in Phonology: Theory, Perception, Acquisition* (pp. 231–268). Berlin, Germany: Mouton.

Fitzroy, A. B., & Breen, M. (2019). Metric Structure and Rhyme Predictability Modulate Speech Intensity During Child-Directed and Read-Alone Productions of Children's Literature. *Language and Speech*, 002383091984315. https://doi.org/10.1177/0023830919843158

Floccia, C., Nazzi, T., Delle Luche, C., Poltrock, S., & Goslin, J. (2014). English-learning one- to two-year-olds do not show a consonant bias in word learning. *Journal of Child Language*, *41*(5), 1085–1114. https://doi.org/10.1017/S0305000913000287

Foy, J. G., & Mann, V. (2001). Does strength of phonological representations predict phonological awareness in preschool children? *Applied Psycholinguistics*, *22*(3), 301–325. https://doi.org/10.1017/S0142716401003022

Franco, F., Suttora, C., Spinelli, M., Kozar, I., & Fasolo, M. (2021). Singing to infants matters: Early singing interactions affect musical preferences and facilitate vocabulary building. *Journal of Child Language*, *48*, 1–26. https://doi.org/10.1017/S0305000921000167

François, C., Teixidó, M., Takerkart, S., 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. https://doi.org/10.1038/s41598-017-12798-2

Frank, M. C., Alcock, K. J., Arias-Trejo, N., Aschersleben, G., Baldwin, D., Barbu, S., … Soderstrom, M. (2020). Quantifying Sources of Variability in Infancy Research Using the Infant-Directed-Speech Preference. *Advances in Methods and Practices in Psychological Science*, *3*(1), 24–52. https://doi.org/10.1177/2515245919900809

Frazier, L., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, *10*(6), 244–249. https://doi.org/10.1016/j.tics.2006.04.002

Frost, R. L. A., Dunn, K., Christiansen, M. H., Gómez, R. L., & Monaghan, P. (2020). Exploring the "anchor word" effect in infants: Segmentation and categorisation of speech with and without high frequency words. *PLoS ONE*, *15*(12 December), 1–28. https://doi.org/10.1371/journal.pone.0243436

Frost, R. L. A., Monaghan, P., & Christiansen, M. H. (2019). Mark My Words: High Frequency Marker Words Impact Early Stages of Language Learning. *Journal of Experimental Psychology: Learning Memory and Cognition*, *45*(10), 1883–1898. https://doi.org/10.1037/xlm0000683

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences*, *3*(11), 410–419. https://doi.org/10.1016/S1364-6613(99)01388-1

Gervain, J., & Mehler, J. (2010). Speech Perception and Language Acquisition in the First Year of Life. *Annual Review of Psychology*, *61*, 191–218. https://doi.org/10.1146/annurev.psych.093008.100408

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

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby)Talk to Me: The Social Context of Infant-Directed Speech and Its Effects on Early Language Acquisition. *Current Directions in Psychological Science*, *24*(5), 339–344. https://doi.org/10.1177/0963721415595345

Gomez-Ramirez, M., Kelly, S. P., Molholm, S., Sehatpour, P., Schwartz, T. H., & Foxe, J. J. (2011). Oscillatory sensory selection mechanisms during intersensory attention to rhythmic auditory and visual inputs: A human electrocorticographic investigation. *Journal of Neuroscience*, *31*(50), 18556–18567. https://doi.org/10.1523/JNEUROSCI.2164-11.2011

Gordon, R. L., Jacobs, M. S., Schuele, C. M., & McAuley, J. D. (2015). Perspectives on the rhythm-grammar link and its implications for typical and atypical language development. *Annals of the New York Academy of Sciences*, *1337*(1), 16–25. https://doi.org/10.1111/nyas.12683

Gordon, R. L., Shivers, C. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., & Devin Mcauley, J. (2015). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental Science*, *18*(4), 635–644. https://doi.org/10.1111/desc.12230

Goswami, U. (2002). In the beginning was the rhyme? A reflection on Hulme, Hatcher, Nation, Brown, Adams, and Stuart (2002). *Journal of Experimental Child Psychology*, *82*(1), 47–57. https://doi.org/10.1006/jecp.2002.2673

Goswami, U. (2019a). A neural oscillations perspective on phonological development and phonological processing in developmental dyslexia. *Language and Linguistics Compass*, *13*(5), e12328. https://doi.org/10.1111/lnc3.12328

Goswami, U. (2019b). Speech rhythm and language acquisition: an amplitude modulation phase hierarchy perspective. *Annals of the New York Academy of Sciences*, *1453*, 67–78. https://doi.org/10.1111/nyas.14137

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. *PLoS Biology*, *11*(12). https://doi.org/10.1371/journal.pbio.1001752

Grossi, G., Coch, D., Coffey-Corina, S., Holcomb, P. J., & Neville, H. J. (2001). Phonological processing in visual rhyming: A developmental ERP study. *Journal of Cognitive Neuroscience*, *13*(5), 610–625. https://doi.org/10.1162/089892901750363190

Gupta, P., Lipinski, J., & Aktunc, E. (2005). Reexamining the phonological similarity effect in immediate serial recall: The roles of type of similarity, category cuing, and item recall. *Memory & Cognition*, *33*(6), 1001–1016.

Haegens, S., & Zion Golumbic, E. (2018). Rhythmic facilitation of sensory processing: a critical review. *Neuroscience and Biobehavioral Reviews*, *86*, 150–165. https://doi.org/10.1016/j.neubiorev.2017.12.002

Hahn, L. E., Benders, T., Snijders, T. M., & Fikkert, P. (2018). Infants' sensitivity to rhyme in songs. *Infant Behavior and Development*, *52*, 130–139. https://doi.org/10.1016/j.infbeh.2018.07.002

Hahn, L. E., Benders, T., Snijders, T. M., & Fikkert, P. (2020). Six-month-old infants recognize phrases in song and speech. *Infancy*, 699–718. https://doi.org/10.1111/infa.12357

Hannon, E. E., Nave-Blodgett, J. E., & Nave, K. M. (2018). The Developmental Origins of the Perception and Production of Musical Rhythm. *Child Development Perspectives*, *0*(0), 1–5. https://doi.org/10.1111/cdep.12285

Hanson, K., & Kiparsky, P. (1996). A parametric theory of poetic meter. *Language*, *72*(2), 287–335. https://doi.org/10.2307/416652

Hawthorne, K., & Gerken, L. (2013). Younger versus older infants' use of prosody-like boundaries to locate musical phrases. *J Acoust Soc Am*, *134*(4105). https://doi.org/10.1121/1.4831067 [doi]

Hawthorne, K., & Gerken, L. (2014). From pauses to clauses: Prosody facilitates learning of syntactic constituency. *Cognition*, *133*(2), 420–428. https://doi.org/10.1016/j.cognition.2014.07.013

Hawthorne, K., Mazuka, R., & Gerken, L. (2015). The acoustic salience of prosody trumps infants' acquired knowledge of language-specific prosodic patterns. *Journal of Memory and Language*, *82*, 105–117. https://doi.org/10.1016/j.jml.2015.03.005

Hayes, B. (1989). The Prosodic Hierarchy in Meter. In P. Kiparsky & G. Youmans (Eds.), *Rhythm and Meter* (1st ed., pp. 201–260). Orlando: Academic Press.

Hayes, B. (2010). Metrics and phonological theory. *Linguistics: The Cambridge Survey*, (Evans 1974), 220–249. https://doi.org/10.1017/cbo9780511621055.013

Hayes, B., & Kaun, A. (1996). The role of phonological phrasing in sung and chanted verse. *Linguistic Review*, *13*(4), 243–304. https://doi.org/10.1515/tlir.1996.13.3-4.243

Hayes, D. S. (2001). Young children's phonological sensitivity after exposure to a rhyming or nonrhyming story. *The Journal of Genetic Psychology*, *162*(3), 253–259. https://doi.org/10.1080/00221320109597482

Hayes, R. A., Slater, A., & Brown, E. (2000). Infants' ability to categorise on the basis of rhyme. *Cognitive Development*, *15*(2000), 405–419.

Hayes, R. A., Slater, A. M., & Longmore, C. A. (2009). Rhyming abilities in 9-month-olds: The role of the vowel and coda explored. *Cognitive Development*, *24*(2), 106–112. https://doi.org/10.1016/j.cogdev.2008.11.002

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

Heffner, C. C., & Slevc, L. R. (2015). Prosodic structure as a parallel to musical structure. *Frontiers in Psychology*, *6*(1962), 1–14. https://doi.org/10.3389/fpsyg.2015.01962

Henrich, K., Alter, K., Wiese, R., & Domahs, U. (2014). The relevance of rhythmical alternation in language processing: An ERP study on English compounds. *Brain and Language*, *136*, 19–30. https://doi.org/10.1016/j.bandl.2014.07.003

Hochmann, J. R., Langus, A., & Mehler, J. (2016). An Advantage for Perceptual Edges in Young Infants' Memory for Speech. *Language Learning*, *66*(September), 13–28. https://doi.org/10.1111/lang.12202

Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: evidence from German and French infants. *Infant Behavior & Development*, *32*(3), 262–274. https://doi.org/10.1016/j.infbeh.2009.03.004

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, *5*, 69–95.

Huron, D. (2006). *Sweet Anticipation*. Cambridge, MA: MIT Press.

Hurschler, M. A., Liem, F., Jäncke, L., & Meyer, M. (2013). Right and left perisylvian cortex and left inferior frontal cortex mediate sentence-level rhyme detection in spoken language as revealed by sparse fMRI. *Human Brain Mapping*, *34*(12), 3182–3192. https://doi.org/10.1002/hbm.22134

Hurschler, M. A., Liem, F., Oechslin, M., Stümpfli, P., & Meyer, M. (2015). fMRI reveals lateralized pattern of brain activity modulated by the metrics of stimuli during auditory rhyme processing. *Brain and Language*, *147*, 41–50. https://doi.org/10.1016/j.bandl.2015.05.004

Ilari, B. (2005). On musical parenting of young children: musical beliefs and behaviors of mothers and infants. *Early Child Development and Care*, *175*(7–8), 647–660. https://doi.org/10.1080/0300443042000302573

Ilari, B., Moura, A., & Bourscheidt, L. (2011). Between interactions and commodities: musical parenting of infants and toddlers in Brazil. *Music Education Research*, *13*(1), 51–67. https://doi.org/10.1080/14613808.2011.553277

International Phonetic Association. (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University Press.

Jadoul, Y., Ravignani, A., Thompson, B., Filippi, P., & de Boer, B. (2016). Seeking temporal predictability in speech: Comparing statistical approaches on 18 world languages. *Frontiers in Human Neuroscience*, *10*(DEC2016). https://doi.org/10.3389/fnhum.2016.00586

Jakobson, R. (1960). Linguistics and Poetics. In T. Sebeok (Ed.), *Style in Language*. Cambridge, MA. https://doi.org/10.1515/jlse.1979.8.1.3

Jakobson, R., & Waugh, L. R. (1979). *The Sound Shape of Language*. Berlin, Boston: Mouton de Gruyter.

Janssen, C., Segers, E., McQueen, J. M., & Verhoeven, L. (2017). Transfer from implicit to explicit phonological abilities in first and second language learners. *Bilingualism*, *20*(4), 795–812. https://doi.org/10.1017/S1366728916000523

Jessen, S., Fiedler, L., Münte, T. F., & Obleser, J. (2019). Quantifying the individual auditory and visual brain response in 7-month-old infants watching a brief cartoon movie. *NeuroImage*, *202*(July), 116060. https://doi.org/10.1016/j.neuroimage.2019.116060

Johnson, E. K. (2016). Constructing a Proto-Lexicon: An Integrative View of Infant Language Development. *Annual Review of Linguistics*, *2*(1), 391–412. https://doi.org/10.1146/annurev-linguistics-011415-040616

Johnson, E. K., & Jusczyk, P. W. (2001). Word Segmentation by 8-Month-Olds: When Speech Cues Count More Than Statistics. *Journal of Memory and Language*, *44*(4), 548–567. https://doi.org/10.1006/jmla.2000.2755

Johnson, E. K., & Seidl, A. (2008). Clause segmentation by 6-month-old infants: A crosslinguistic perspective. *Infancy*, *13*(5), 440–455. https://doi.org/10.1080/15250000802329321

Johnson, E. K., Seidl, A., & Tyler, M. D. (2014). The edge factor in early word segmentation: utterance-level prosody enables word form extraction by 6-month-olds. *PloS One*, *9*(1), e83546. https://doi.org/10.1371/journal.pone.0083546

Junge, C., & Cutler, A. (2014). Early word recognition and later language skills. *Brain Sciences*, *4*(4), 532–559. https://doi.org/10.3390/brainsci4040532

Junge, C., Cutler, A., & Hagoort, P. (2014). Successful Word Recognition by 10-Month-Olds Given Continuous Speech Both at Initial Exposure and Test. *Infancy*, *19*(2), 179–193. https://doi.org/10.1111/infa.12040

Junge, C., Everaert, E., Porto, L., Fikkert, P., de Klerk, M., Keij, B., & Benders, T. (2020). Contrasting behavioral looking procedures: a case study on infant speech segmentation. *Infant Behavior and Development*, *60*(June), 101448. https://doi.org/10.1016/j.infbeh.2020.101448

Junge, C., Kooijman, V. M., Hagoort, P., & Cutler, A. (2012). Rapid recognition at 10 months as a predictor of language development. *Developmental Science*, *15*(4), 463–473. https://doi.org/10.1111/j.1467-7687.2012.1144.x

Jusczyk, P. W. (2003). American infants' perception of cues to grammatical units in non-native languages and music: evidence from Polish and Japanese. In *Jusczyk lab final report*. Retrieved from http://hincapie.psych.purdue.edu/Jusczyk

Jusczyk, P. W., Goodman, M., & Baumann, A. (1999). Nine-month-olds' attention to sound similarities in syllables. *Journal of Memory and Language*, *82*, 62–82. Retrieved from http://www.sciencedirect.com/science/article/pii/S0749596X98926055

Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D. G., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, *24*(2), 252–293. https://doi.org/10.1016/0010-0285(92)90009-Q

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in english-learning infants. *Cognitive Psychology*, *39*(3–4), 159–207. https://doi.org/10.1006/cogp.1999.0716

Jusczyk, P. W., & Krumhansl, C. L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology. Human Perception and Performance*, *19*(3), 627–640. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8331317

Kalashnikova, M., Peter, V., Di Liberto, G. M., Lalor, E. C., & 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

Kayser, S. J., Ince, R. A. A., Gross, J., & Kayser, C. (2015). Irregular Speech Rate Dissociates Auditory Cortical Entrainment, Evoked Responses, and Frontal Alpha. *The Journal of Neuroscience*, *35*(44), 14691–14701. https://doi.org/10.1523/JNEUROSCI.2243-15.2015

Keitel, A., Gross, J., & Kayser, C. (2017). Speech tracking in auditory and motor regions reflects distinct linguistic features. *PLoS Biology*, 1–19. https://doi.org/10.1101/195941

Kelly, M. H., & Rubin, D. C. (1988). Natural rhythmic patterns in English verse: Evidence from child counting-out rhymes. *Journal of Memory and Language*, *27*(6), 718–740. https://doi.org/10.1016/0749-596X(88)90017-4

Kemler Nelson, D. G., Jusczyk, P. W., & Mandel, D. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior & Development*, *16*, 111–116. Retrieved from http://www.sciencedirect.com/science/article/pii/0163638395900128

Khateb, A., Pegna, A. J., Landis, T., Michel, C. M., Brunet, D., Seghier, M. L., & Annoni, J.-M. (2007). Rhyme processing in the brain: an ERP mapping study. *International Journal of Psychophysiology*, *63*(3), 240–250. https://doi.org/10.1016/j.ijpsycho.2006.11.001

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks Effect in Infant Auditory Attention. *Child Development*, *85*(5), 1795–1804. https://doi.org/10.1111/cdev.12263

Kidd, E., Junge, C., Spokes, T., Morrison, L., & Cutler, A. (2018). Individual Differences in Infant Speech Segmentation: Achieving the Lexical Shift. *Infancy*, *23*(6), 770–794. https://doi.org/10.1111/infa.12256

Kiparsky, P. (1973). The Role of Linguistics in a Theory of Poetry. In *Deadalus, Language as a Human Problem* (Vol. 102, pp. 231–244).

Kluender, R., & Kutas, M. (1993a). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, *5*(2), 196–214. https://doi.org/10.1162/jocn.1993.5.2.196

Kluender, R., & Kutas, M. (1993b). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, *8*(4), 573–633. https://doi.org/10.1080/01690969308407588

Kooijman, V. M., Hagoort, P., & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research*, *24*(1), 109–116. https://doi.org/10.1016/j.cogbrainres.2004.12.009

Kooijman, V. M., Hagoort, P., & Cutler, A. (2009). Prosodic Structure in Early Word Segmentation: ERP Evidence From Dutch Ten-Month-Olds. *Infancy*, *14*(6), 591–612. https://doi.org/10.1080/15250000903263957

Kooijman, V. M., Johnson, E. K., & Cutler, A. (2008). Reflections on reflections of infant word recognition. In A. D. Friederici & G. Thierry (Eds.), *Early Language Development: Bridging Brain and Behaviour* (pp. 91–114). Amsterdam / Philadelphia: John Benhamins Publishing Company.

Kooijman, V. M., Junge, C., Johnson, E. K., Hagoort, P., & Cutler, A. (2013). Predictive brain signals of linguistic development. *Frontiers in Psychology*, *4*(February), 25. https://doi.org/10.3389/fpsyg.2013.00025

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: A timely subcortico-cortical framework. *Trends in Cognitive Sciences*, *14*(9), 392–399. https://doi.org/10.1016/j.tics.2010.06.005

Kragness, H. E., Johnson, E. K., & Cirelli, L. K. (2021). The song, not the singer: Infants prefer to listen to familiar songs, regardless of singer identity. *Developmental Science*, *e13147*(January). https://doi.org/10.1111/desc.13149

Krijnen, E., van Steensel, R., Meeuwisse, M., Jongerling, J., & Severiens, S. (2020). Exploring a refined model of home literacy activities and associations with children's emergent literacy skills. *Reading and Writing*, *33*(1), 207–238. https://doi.org/10.1007/s11145-019-09957-4

Kriukova, O., & Mani, N. (2016). Processing metrical information in silent reading: An ERP study. *Frontiers in Psychology*, *7*(SEP), 1–8. https://doi.org/10.3389/fpsyg.2016.01432

Krumhansl, C. L., & Jusczyk, P. W. (1990). Infants' perception of phrase structure in music. *Psychological Science*, *1*(1), 70–73. Retrieved from http://pss.sagepub.com/content/1/1/70.short

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., ... Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, *277*(5326), 684–686. https://doi.org/10.1126/science.277.5326.684

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2016). lmerTest: Tests in Linear Mixed Effects Models. *R Package Version*, p. https://cran.r-project.org/package=lmerTest. Retrieved from https://cran.r-project.org/package=lmerTest

Ladányi, E., Persici, V., Fiveash, A., Tillmann, B., & Gordon, R. L. (2020). Is atypical rhythm a risk factor for developmental speech and language disorders? *Wiley Interdisciplinary Reviews: Cognitive Science*, *11*(5), 1–32. https://doi.org/10.1002/wcs.1528

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, *320*(5872), 110–113. https://doi.org/10.1126/science.1154735

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology*, *94*(3), 1904–1911. https://doi.org/10.1152/jn.00263.2005

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 1–12. https://doi.org/10.3389/fpsyg.2013.00863

Lallier, M., Lizarazu, M., Molinaro, N., Bourguignon, M., Ríos-López, P., & Carreiras, M. (2018). From Auditory Rhythm Processing to Grapheme-to-Phoneme Conversion: How Neural Oscillations Can Shed Light on Developmental Dyslexia. In *Reading and Dyslexia - From Basic Functions to Higher Order Cognition* (Literacy S, pp. 147–163). https://doi.org/10.1007/978-3-319-90805-2_8

Lang, A., Ott, P., Giudice, R. Del, & Schabus, M. (2020). Memory traces formed in utero—newborns' autonomic and neuronal responses to prenatal stimuli and the maternal voice. *Brain Sciences*, *10*(11), 1–10. https://doi.org/10.3390/brainsci10110837

Langus, A., Marchetto, E., Bion, R. A. H., & Nespor, M. (2012). Can prosody be used to discover hierarchical structure in continuous speech? *Journal of Memory and Language*, *66*(1), 285–306. https://doi.org/10.1016/j.jml.2011.09.004

Langus, A., Mehler, J., & Nespor, M. (2017). Rhythm in language acquisition. *Neuroscience and Biobehavioral Reviews*, *81*, 158–166. https://doi.org/10.1016/j.neubiorev.2016.12.012

Large, E. W., & Jones, M. R. (1999). The Dynamics of Attending: How People Track Time-Varying Events. *Psychological Review*, *106*(1), 119–159. Retrieved from http://psycnet.apa.org/psycinfo/1999-10188-005

Lea, R. B., Rapp, D. N., Elfenbein, A., Mitchel, A. D., & Romine, R. S. (2008). Sweet silent thought: Alliteration and resonance in poetry comprehension. *Psychological Science*, *19*(7), 709–716. https://doi.org/10.1111/j.1467-9280.2008.02146.x

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 & Development*, *33*(4), 419–430. https://doi.org/10.1016/j.infbeh.2010.04.006

Lehrdahl, F., & Jackendoff, R. (1985). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. *Proceedings of the National Academy of Sciences*, *114*(50), 13290–13295. https://doi.org/10.1073/pnas.1702493114

Leong, V., & Goswami, U. (2015). Acoustic-emergent phonology in the amplitude envelope of child-directed speech. *PLoS ONE*, *10*(12), 1–37. https://doi.org/10.1371/journal.pone.0144411

Leong, V., & Goswami, U. (2016). Difficulties in auditory organization as a cause of reading backwardness? An auditory neuroscience perspective. *Developmental Science*, 1–13. https://doi.org/10.1111/desc.12457

Leong, V., Kalashnikova, M., Burnham, D., & Goswami, U. (2017). The Temporal Modulation Structure of Infant-Directed Speech. *Open Mind*, 1–13. https://doi.org/10.1162/OPMI_a_00008

Lerdahl, F. (2001). The sounds of poetry viewed as music. *Annals of the New York Academy of Sciences*, *930*, 337–354. https://doi.org/10.1111/j.1749-6632.2001.tb05743.x

Lindstromberg, S., & Boers, F. (2008). Phonemic repetition and the learning of lexical chunks: The power of assonance. *System*, *36*(3), 423–436. https://doi.org/10.1016/j.system.2008.01.002

Lomax, A. (1968). *Folk Song Style and Culture*. New Brunswick; London UK: Transaction Publishers.

Longhi, E. (2009). `Songese': maternal structuring of musical interaction with infants. *Psychology of Music*, *37*(2), 195–213. https://doi.org/10.1177/0305735608097042

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: evidence from a latent-variable longitudinal study. *Developmental Psychology*, *36*(5), 596–613. https://doi.org/10.1037/0012-1649.36.5.596

Luck, S. J., & Gaspelin, N. (2017). How to Get Statistically Significant Effects in Any ERP Experiment (and Why You Shouldn't). *Psychophysiology*, *54*(1), 146–157. https://doi.org/10.1111/psyp.12639

Luo, H., & Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron*, *54*(6), 1001–1010. https://doi.org/10.1016/j.neuron.2007.06.004

Luo, Y., & Zhou, X. (2010). ERP evidence for the online processing of rhythmic pattern during Chinese sentence reading. *NeuroImage*, *49*(3), 2836–2849. https://doi.org/10.1016/j.neuroimage.2009.10.008

Ma, W., Fiveash, A., Margulis, E. H., Behrend, D., & Thompson, W. F. (2020). Song and infant-directed speech facilitate word learning. *Quarterly Journal of Experimental Psychology*, *73*(7), 1036–1054. https://doi.org/10.1177/1747021819888982

MacLean, M., Bryant, P., & Bradley, L. (1987). Rhymes, nursery rhymes, and reading in early childhood. *Merrill-Palmer Quarterly*, *33*(3), 255–281. Retrieved from http://www.jstor.org/stable/10.2307/23086536

Maechler, M., Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M. Verbeke, T., … di Palma, A. M. (2021). *robustbase: Basic Robust Statistics*. Retrieved from http://robustbase.r-forge.r-project.org

Magne, C., Astésano, C., Aramaki, M., Ystad, S., Kronland-Martinet, R., & Besson, M. (2007). Influence of syllabic lengthening on semantic processing in spoken French: Behavioral and electrophysiological evidence. *Cerebral Cortex*, *17*(11), 2659–2668. https://doi.org/10.1093/cercor/bhl174

Magne, C., Jordan, D. K., & Gordon, R. L. (2016). Speech rhythm sensitivity and musical aptitude: ERPs and individual differences. *Brain and Language*, *153–154*, 13–19. https://doi.org/10.1016/j.bandl.2016.01.001

Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent Component Analyssis of Electroencephalographic Data. In D. Touretzky, M. Mozer, & M. Hasselmo (Eds.), *Advances in Neural Information Processing Systems* (pp. 145–151). Cambridge, MA: MIT Press. https://doi.org/10.1109/ICOSP.2002.1180091

Mandel, D. R., Jusczyk, P. W., & Kemler Nelson, D. G. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition*, *53*(2), 155–180. https://doi.org/10.1016/0010-0277(94)90069-8

Mann, V. A., & Foy, J. G. (2007). Speech development patterns and phonological awareness in preschool children. *Annals of Dyslexia*, *57*(1), 51–74. https://doi.org/10.1007/s11881-007-0002-1

Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience*, *23*(2), 294–305. https://doi.org/10.1162/jocn.2010.21413

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Markova, G. (2018). The games infants play: Social games during early mother-infant interactions and their relationship with oxytocin. *Frontiers in Psychology*, *9*(JUN), 1–12. https://doi.org/10.3389/fpsyg.2018.01041

Martensen, H., Maris, E., & Dijkstra, T. (2000). When does inconsistency hurt? On the relation between phonological consistency effects and the reliability of sublexical units. *Memory and Cognition*, *28*(4), 648–656. https://doi.org/10.3758/BF03201254

McCauley, S. M., Hestvik, A., & Vogel, I. (2013). Perception and Bias in the Processing of Compound versus Phrasal Stress: Evidence from Event-related Brain Potentials. *Language and Speech*, *56*(1), 23–44. https://doi.org/10.1177/0023830911434277

Mehr, S. A. (2014). Music in the home: New evidence for an intergenerational link. *Journal of Research in Music Education*, *62*(1), 78–88. https://doi.org/10.1177/0022429413520008

Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., … Glowacki, L. (2019). Universality and diversity in human song. *Science*, *366*(November, eaax0868), 1–17. https://doi.org/10.1126/science.aax0868

Mehr, S. A., Singh, M., York, H., Glowacki, L., & Krasnow, M. M. (2018). Form and Function in Human Song. *Current Biology*, *28*(3), 356-368.e5. https://doi.org/10.1016/j.cub.2017.12.042

Mehr, S. A., Song, L. A., & Spelke, E. S. (2016). For 5-Month-Old Infants, Melodies Are Social. *Psychological Science*, *27*(4), 486–501. https://doi.org/10.1177/0956797615626691

Meints, K., & Woodford, A. (2008). *Lincoln Infant Lab Package 1.0: A new programme package for IPL, preferential listening, habituation, and eye-tracking*. 1–44. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Lincoln+Infant+Lab+Package+1+.+0+:+A+new+programme+package+for+IPL+,+Preferential+Listening+,+Habituation#1

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*(2), 322–352. https://doi.org/10.1037/a0026744

Menn, K., Ward, E., Braukmann, R., van den Boomen, C., Kemner, C., Buitelaar, J., … Snijders, T. M. (2020). Relating neural entrainment to speech to later development of language and autism symptoms in infants with high likelihood of ASD. *Poster Presented at the Virtual International Congress of Infant Studies (VICIS 2020)*. Glasgow, UK.

Menninghaus, W., Bohrn, I. C., Altmann, U., Lubrich, O., & Jacobs, A. M. (2014). Sounds funny? Humor effects of phonological and prosodic figures of speech. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(1), 71–76. https://doi.org/10.1037/a0035309

Menninghaus, W., Bohrn, I. C., Knoop, C. A., Kotz, S. A., Schlotz, W., & Jacobs, A. M. (2015a). Rhetorical features facilitate prosodic processing while handicapping ease of semantic comprehension. *Cognition*, *143*, 48–60. https://doi.org/10.1016/j.cognition.2015.05.026

Menninghaus, W., Bohrn, I. C., Knoop, C. A., Kotz, S. A., Schlotz, W., & Jacobs, A. M. (2015b). Rhetorical features facilitate prosodic processing while handicapping ease of semantic comprehension. *Cognition*, *143*, 48–60. https://doi.org/10.1016/j.cognition.2015.05.026

Merrill, J., Sammler, D., Bangert, M., Goldhahn, D., Lohmann, G., Turner, R., & Friederici, A. D. (2012). Perception of words and pitch patterns in song and speech. *Frontiers in Psychology*, *3*(March), 76. https://doi.org/10.3389/fpsyg.2012.00076

Metsala, J. L., & Walley, A. C. (1998). Spoken Vocabulary Growth and the Segmental Restructuring of Lexical Representations: Precursors to Phonemic Awareness and Early Reading Ability. In J. L. Metsala & C. Ehri L (Eds.), *Word Recognition in Beginning Literacy* (pp. 89–120). Mahwah, NJ; London: Lawrence Erlbaum Associates.

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. *European Journal of Neuroscience*, *48*(7), 2609–2621. https://doi.org/10.1111/ejn.13748

Mills, D. L., Prat, C., Zangl, R., Stager, C. L., Neville, H. J., & Werker, J. F. (2004). Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month-olds. *Journal of Cognitive Neuroscience*, *16*(8), 1452–1464. https://doi.org/10.1162/0898929042304697

Mohan, R., & Weber, C. (2015). Neural systems mediating processing of sound units of language distinguish recovery versus persistence in stuttering. *Journal of Neurodevelopmental Disorders*, *7*(1), 1–13. https://doi.org/10.1186/s11689-015-9124-7

Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., & Carreiras, M. (2016). Out-of-synchrony speech entrainment in developmental dyslexia. *Human Brain Mapping*, *37*(8), 2767–2783. https://doi.org/10.1002/hbm.23206

Morgan, J., & Demuth, K. (1996). *Signal to syntax : Bootstrapping from speech to grammar in early acquisition*. Mahwah, NJ: L. Erlbaum Associates.

Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior & Development*, *27*(4), 455–464. https://doi.org/10.1016/j.infbeh.2004.03.002

Nakata, T., & Trehub, S. E. (2011). Expressive timing and dynamics in infant-directed and non-infant-directed singing. *Psychomusicology: Music, Mind and Brain*, *21*(1–2), 45–53. https://doi.org/10.1037/h0094003

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. https://doi.org/10.1037/0096-1523.24.3.756

Nazzi, T., Kemler Nelson, D. G., Jusczyk, P. W., & Jusczyk, A. M. (2000). Six-Month-Olds ’ Detection of Clauses Embedded in Continuous Speech : Effects of Prosodic Well-Formedness. *Infancy*, *1*(1), 123–147.

Nespor, M., & Vogel, I. (2007). *Prosodic phonology: with a new foreword*. Berlin, Germany: Mouton de Gruyter.

Newman, R. S., 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. https://doi.org/10.1037/0012-1649.42.4.643

Ngon, C., Martin, A., Dupoux, E., Cabrol, D., Dutat, M., & Peperkamp, S. (2013). (Non)words, (non) words, (non)words: Evidence for a protolexicon during the first year of life. *Developmental Science*, *16*(1), 24–34. https://doi.org/10.1111/j.1467-7687.2012.01189.x

Noordenbos, M. W., Segers, E., Wagensveld, B., & Verhoeven, L. (2013). Aberrant N400 responses to phonological overlap during rhyme judgements in children at risk for dyslexia. *Brain Research*, *1537*, 233–243. https://doi.org/10.1016/j.brainres.2013.09.018

Nordt, M., Hoehl, S., & Weigelt, S. (2016). The use of repetition suppression paradigms in developmental cognitive neuroscience. *Cortex*, *80*, 61–75. https://doi.org/10.1016/j.cortex.2016.04.002

Obermeier, C., Kotz, S. A., Jessen, S., Raettig, T., von Koppenfels, M., & Menninghaus, W. (2016). Aesthetic appreciation of poetry correlates with ease of processing in event-related potentials. *Cognitive, Affective, & Behavioral Neuroscience*, 362–373. https://doi.org/10.3758/s13415-015-0396-x

Obermeier, C., Menninghaus, W., von Koppenfels, M., Raettig, T., Schmidt-Kassow, M., Otterbein, S., & Kotz, S. A. (2013). Aesthetic and emotional effects of meter and rhyme in poetry. *Frontiers in Psychology*, *4*(January), 10. https://doi.org/10.3389/fpsyg.2013.00010

Obleser, J., & Kayser, C. (2019). Neural Entrainment and Attentional Selection in the Listening Brain. *Trends in Cognitive Sciences*, *23*(11), 913–926. https://doi.org/10.1016/j.tics.2019.08.004

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2011/156869

Opie, I., & Opie, P. (1959). *The Lore and Language of Schoolchildren*. Oxford, UK: Oxford University Press.

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

Oxford Learners Dictionairy. (n.d.). nursery rhyme. Retrieved from 2021 website: https://www.oxfordlearnersdictionaries.com/definition/english/nursery-rhyme?q=nursery+rhyme

Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, *3*(SEP), 1–17. https://doi.org/10.3389/fpsyg.2012.00320

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral Cortex*, *23*(6), 1378–1387. https://doi.org/10.1093/cercor/bhs118

Perrin, F., & García-Larrea, L. (2003). Modulation of the N400 potential during auditory phonological/semantic interaction. *Cognitive Brain Research*, *17*(1), 36–47. https://doi.org/10.1016/S0926-6410(03)00078-8

Peust, C. (2014). Towards a typology of poetic rhyme. In E. Grossman, M. Haspelmath, & T. S. Richter (Eds.), *Egyptian-Coptic Linguistics in Typological Perspective (EALT) 55*. Berlin, Munich, oston. https://doi.org/10.1515/9783110346510.341

Pinker, S. (1997). *How the Mind Works*. New York, NY: W.W. Norton & Company, Inc.

Politimou, N. (2018). *Relationships between musical and linguistic skills in early development : the role of informal musical experience in the home*.

Politimou, N., Dalla Bella, S., Farrugia, N., & Franco, F. (2019). Born to speak and sing: Musical predictors of language development in pre-schoolers. *Frontiers in Psychology*, *10*(APR). https://doi.org/10.3389/fpsyg.2019.00948

Port, R. F. (2003). Meter and speech. *Journal of Phonetics*, *31*(3–4), 599–611. https://doi.org/10.1016/j.wocn.2003.08.001

Power, A. J., Mead, N., Barnes, L., & Goswami, U. (2013). Neural entrainment to rhythmic speech in children with developmental dyslexia. *Frontiers in Human Neuroscience*, *7*(November), 777. https://doi.org/10.3389/fnhum.2013.00777

Praamstra, P., Meyer, A. S., & Levelt, W. J. M. (1994). Neurophysiological manifestations of phonological processing: Latency variation of a negative ERP component timelocked to phonological mismatch. *Journal of Cognitive Neuroscience*, *6*(3), 204–219. https://doi.org/10.1162/jocn.1994.6.3.204

Praamstra, P., & Stegeman, D. F. (1993). Phonological effects on the auditory N400 event-related brain potential. *Cognitive Brain Research*, *1*(2), 73–86. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8513242

Quené, H., & Port, R. F. (2005). Effects of timing regularity and metrical expectancy on spoken-word perception. *Phonetica*, *62*(1), 55–67. https://doi.org/10.1159/000087222

Quené, H., & van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: a tutorial. *Speech Communication*, *43*(1–2), 103–121. https://doi.org/10.1016/j.specom.2004.02.004

R Development Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://doi.org/10.1007/978-3-540-74686-7

Radeau, M., Besson, M., Fonteneau, E., & Castro, S. L. (1998). Semantic, repetition and rime priming between spoken words: behavioral and electrophysiological evidence. *Biological Psychology*, *48*(2), 183–204. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9700017

Ramus, F. (1999). Language identification with suprasegmental cues: a study based on speech resynthesis. *The Journal of the Acoustical Society of America*, *105*(1), 512–521. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9921674

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*(1), AD3–AD30. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10908711

Räsänen, O., Kakouros, S., & Soderstrom, M. (2018). Is infant-directed speech interesting because it is surprising? – Linking properties of IDS to statistical learning and attention at the prosodic level. *Cognition*, *178*(October 2017), 193–206. https://doi.org/10.1016/j.cognition.2018.05.015

Read, K. (2014). Clues cue the smooze: rhyme, pausing, and prediction help children learn new words from storybooks. *Frontiers in Psychology*, *5*(February), 1–10. https://doi.org/10.3389/fpsyg.2014.00149

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality & Social Psychology Review*, *8*(4), 364–382.

Richards, S., & Goswami, U. (2019). Impaired Recognition of Metrical and Syntactic Boundaries in Children with Developmental Language Disorders. *Brain Sciences*, *9*(2), 33. https://doi.org/10.3390/brainsci9020033

Riemann, H. (1912). *Vademecum der Phrasierung* (4th ed.). Berlin: Max Hesses illustrierte Handbücher.

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive Sensing of Periodic and Aperiodic Auditory Patterns. *Trends in Cognitive Sciences*, *22*(10), 870–882. https://doi.org/10.1016/j.tics.2018.08.003

Roncaglia-Denissen, M. P., Schmidt-Kassow, M., & Kotz, S. A. (2013). Speech Rhythm Facilitates Syntactic Ambiguity Resolution: ERP Evidence. *PLoS ONE*, *8*(2), 1–9.

Rosenberg, J. R., Amjad, A. M., Breeze, P., Brillinger, D. R., & Halliday, D. M. (1989). The Fourier approach to the identification of functional coupling between neuronal spike trains. *Progress in Biophysics and Molecular Biology*, *53*(1), 1–31. https://doi.org/10.1016/0079-6107(89)90004-7

Rösler, F., Pechmann, T., Streb, J., Röder, B., & Hennighausen, E. (1998). Parsing of Sentences in a Language with Varying Word Order: Word-by-Word Variations of Processing Demands Are Revealed by Event-Related Brain Potentials. *Journal of Memory and Language*, *38*(2), 150–176. https://doi.org/10.1006/jmla.1997.2551

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex*, *17*(5), 1147–1153. https://doi.org/10.1093/cercor/bhl024

Rothermich, K., & Kotz, S. A. (2013). Predictions in speech comprehension: fMRI evidence on the meter-semantic interface. *NeuroImage*, *70*, 89–100. https://doi.org/10.1016/j.neuroimage.2012.12.013

Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2012). Rhythm's gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia*, *50*(2), 232–244. https://doi.org/10.1016/j.neuropsychologia.2011.10.025

Rothermich, K., Schmidt-Kassow, M., Schwartze, M., & Kotz, S. A. (2010). Event-related potential responses to metric violations: rules versus meaning. *Neuroreport*, *21*(8), 580–584. https://doi.org/10.1097/WNR.0b013e32833a7da7

Rubin, D. C. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes.* New York, NY: Oxford University Press.

Rugg, M. D. (1984a). Event-related potentials and the phonological processing of words and non-words. *Neuropsychologia*, *22*(4), 435–443. https://doi.org/10.1016/0028-3932(84)90038-1

Rugg, M. D. (1984b). Event-related potentials in phonological matching tasks. *Brain and Language*, *23*(2), 225–240. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6518354

Rugg, M. D., & Barrett, S. E. (1987). Event-related potentials and the interaction between orthographic and phonological information in a rhyme-judgment task. *Brain and Language*, *32*(2), 336–361. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3690257

Rugg, M. D., Doyle, M. C., & Wells, T. (1995). Word and nonword repetition within- and across-modality: an event-related potential study. *Journal of Cognitive Neuroscience*, *7*(2), 209–227. https://doi.org/10.1162/jocn.1995.7.2.209

Sassenhagen, J., & Draschkow, D. (2019). Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location. *Psychophysiology*, (August 2018), 1–8. https://doi.org/10.1111/psyp.13335

Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). *Statistical universals reveal the structures and functions of human music*. (27). https://doi.org/10.1073/pnas.1414495112

Schmidt-Kassow, M., & Kotz, S. A. (2009). Event-related brain potentials suggest a late interaction of meter and syntax in the P600. *Journal of Cognitive Neuroscience*, *21*(9), 1693–1708. https://doi.org/10.1162/jocn.2008.21153

Schön, D., Gordon, R. L., Campagne, A., Magne, C., Astésano, C., Anton, J.-L., & Besson, M. (2010). Similar cerebral networks in language, music and song perception. *NeuroImage*, *51*(1), 450–461. https://doi.org/10.1016/j.neuroimage.2010.02.023

Schreuder, R., & Baayen, H. (1994). Prefix Stripping Re-Revisited. *Journal of Memory and Language*, Vol. 33, pp. 357–375. https://doi.org/10.1006/jmla.1994.1017

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. https://doi.org/10.1016/j.tins.2008.09.012

Segal, O., & Kishon-Rabin, L. (2012). Evidence for Language-Specific Influence on the Preference of Stress Patterns in Infants Learning an Iambic Language (Hebrew). *Journal of Speech Language and Hearing Research*, *55*(5), 1329. https://doi.org/10.1044/1092-4388(2012/11-0087)

Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language*, *57*(1), 24–48. https://doi.org/10.1016/j.jml.2006.10.004

Seidl, A., & Cristia, A. (2008). Developmental changes in the weighting of prosodic cues. *Developmental Science*, *4*, 596–606. https://doi.org/10.1111/j.1467-7687.2008.00704.x

Seidl, A., French, B., Wang, Y., & Cristia, A. (2014). Toward establishing continuity in linguistic skills within early infancy. *Language Learning*, *64*(SUPPL.2), 165–183. https://doi.org/10.1111/lang.12059

Seidl, A., & Johnson, E. K. (2006). Infant word segmentation revisited: edge alignment facilitates target extraction. *Developmental Science*, *9*(6), 565–573. https://doi.org/10.1111/j.1467-7687.2006.00534.x

Selkirk, E. (1980). The Role of Prosodic Categories in English Word Stress. *Linguistic Inquiry*, *11*(3), 563–605. https://doi.org/10.2307/4178179

Seuss. (1957). *The cat in the hat*. New York, NY: Random House.

Shanahan, D., & Huron, D. (2011). Interval size and phrase position: A comparison between German and Chinese folksongs. *Empirical Musicology Review*, *6*(4), 187–197.

Sheingold, K., & Foundas, A. (1978). Rhymes for some reasons: Effect of rhyme on children's memory for detail and sequence in simple narratives. *Psychological Reports*, *43*, 1231–1234. Retrieved from http://www.amsciepub.com/doi/pdf/10.2466/pr0.1978.43.3f.1231

Shenfield, T., Trehub, S. E., & Nakata, T. (2003). Maternal Singing Modulates Infant Arousal. *Psychology of Music*, *31*(4), 365–375. https://doi.org/10.1177/03057356030314002

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 of the United States of America*, *108*(15), 6038–6043. https://doi.org/10.1073/pnas.1017617108

Snijders, T. M. (2020). Getting the rhythm for infant language learning: infants' cortical tracking of speech rhythm relates to their word segmentation performance. *Talk Presented at the 45th Annual Boston University (Virtual) Conference on Language Development (BUCLD 45)*. Boston, MA, USA.

Snijders, T. M., Benders, T., & Fikkert, P. (2020). Infants segment words from songs - an EEG study. *Brain Sciences*, *10*(39), 1–25. https://doi.org/10.3390/brainsci10010039

Snijders, T. M., Kooijman, V. M., Cutler, A., & Hagoort, P. (2007). Neurophysiological evidence of delayed segmentation in a foreign language. *Brain Research*, *1178*(1), 106–113. https://doi.org/10.1016/j.brainres.2007.07.080

Soderstrom, M., Kemler Nelson, D. G., & Jusczyk, P. W. (2005). Six-month-olds recognize clauses embedded in different passages of fluent speech. *Infant Behavior & Development*, *28*, 87–94. https://doi.org/10.1016/j.infbeh.2004.07.001

Soderstrom, M., & Wittebolle, K. (2013). When do caregivers talk? The influences of activity and time of day on caregiver speech and child vocalizations in two childcare environments. *PLoS ONE*, *8*(11). https://doi.org/10.1371/journal.pone.0080646

Stadler, M. A., Watson, M., & Skahan, S. (2007). Rhyming and Vocabulary: Effects of Lexical Restructuring. *Communication Disorders Quarterly*, *28*(4), 197–205. https://doi.org/10.1177/1525740107311823

Stallworthy, J. (1996). Versification. In M. Ferguson, M. J. Salter, & J. Stallworthy (Eds.), *The Norton Anthropology of Poetry* (4th ed.). W W Norton & Company Incorporated.

Stern, D. N. (1974). The Goal and Structure of Mother-Infant Play. *Journal of the American Academy of Child Psychiatry*, Vol. 13, pp. 402–421. https://doi.org/10.1016/S0002-7138(09)61348-0

Stroo, M. (2020). *Investigating auditory processing of nursery rhymes: An EEG study of rhyme and metre and their relationship to musical aptitude*. M.Sc. Thesis at Radboud University Nijmegen, The Netherlands.

Sullivan, C. W. (1999). Songs, Poems, And Rhymes. In B. Sutton-Smith, J. Mechling, W. Johnson, Thomas, & R. McMahon, Felicia (Eds.), *Children's Folklore - A Source Book* (pp. 145–160). Logan, Utah: Utah State University Press. https://doi.org/10.2307/j.ctt46nskz.15

Suppanen, E., Huotilainen, M., & Ylinen, S. (2019). Rhythmic structure facilitates learning from auditory input in newborn infants. *Infant Behavior and Development*, *57*(July), 101346. https://doi.org/10.1016/j.infbeh.2019.101346

Swingley, D. (2005). 11 month olds' knowledge of how familiar words sound. *Developmental Science*, *5*, 432–443. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1467-7687.2005.00432.x/full

Swingley, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*, *60*(2), 252. https://doi.org/10.1016/j.jml.2008.11.003

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, *71*(2), 73–108. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10444905

Teixidó, M., François, C., Bosch, L., & Männel, C. (2018). The role of prosody in early speech segmentation and word-referent mapping: Electrophysiological evidence. In P. Prieto & N. Esteve Gibert (Eds.), *Prosodic Development in First Language Acquisition* (Trends in, pp. 79–100). Amsterdam: John Benjamins.

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*, 225–233. https://doi.org/10.1111/j.1749-6632.2009.04547.x

Tierney, A., Russo, F. a., & Patel, A. D. (2011). The motor origins of human and avian song structure. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(37), 15510–15515. https://doi.org/10.1073/pnas.1103882108

Tillmann, B., & Dowling, W. J. (2007). Memory decreases for prose, but not for poetry. *Memory & Cognition*, *35*(4), 628–639. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17848021

Torppa, R., Faulkner, A., Kujala, T., Huotilainen, M., & Lipsanen, J. (2018). Developmental links between speech perception in noise, singing, and cortical processing of music in children with cochlear implants. *Music Perception*, *36*(2), 156–174. https://doi.org/10.1525/MP.2018.36.2.156

Trainor, L. J. (1996). Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. *Infant Behavior & Development*, *19*(1), 83–92.

Trainor, L. J., & Adams, B. (2000). Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns. *Perception and Psychophysics*, *62*(2), 333–340.

Trainor, L. J., Clark, E. D., Huntley, A., & Adams, B. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior & Development*, *20*(3), 383–396.

Trehub, S. E. (2013). Communication, Music and Language in Infancy. In M. A. Arbib (Ed.), *Language, Music and the Brain - A Mysterious Relationship* (pp. 465–480). Cambridge, MA: MIT Press.

Trehub, S. E. (2019). Nurturing infants with music. *International Journal of Music in Early Childhood*, *14*(1), 9–15. https://doi.org/10.1386/ijmec.14.1.9_1

Trehub, S. E., Ghazban, N., & Corbeil, M. (2015). Musical affect regulation in infancy. *Annals of the New York Academy of Sciences*, *1337*(1), 186–192. https://doi.org/10.1111/nyas.12622

Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition*, *100*, 73–99. Retrieved from http://www.sciencedirect.com/science/article/pii/S0010027705002222

Trehub, S. E., & Nakata, T. (2002). Emotion and music in infancy. *Musicae Scientiae*, *5*(37). https://doi.org/10.1177/10298649020050S103

Trehub, S. E., & Trainor, L. J. (1998). Singing to infants: Lullabies and play songs. In Rovee-Collier, Lipsitt, & Hayne (Eds.), *Advances in Infancy Research. Vol. 12* (pp. 43–77). London, UK: Ablex Publishing.

Trehub, S. E., Unyk, A. M., Kamenetsky, S. B., Hill, D. S., Trainor, L. J., Henderson, J. L., & Saraza, M. (1997). Mothers' and fathers' singing to infants. *Developmental Psychology*, *33*(3), 500–507.

Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993a). Adults identify infant-directed music across cultures. *Infant Behavior & Development*, *16*(2), 193–211. https://doi.org/10.1016/0163-6383(93)80017-3

Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993b). Maternal singing in cross-cultural perspective. *Infant Behavior & Development*, *16*(3), 285–295. https://doi.org/10.1016/0163-6383(93)80036-8

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition*, *15*(1–3), 49–74. https://doi.org/10.1016/0010-0277(83)90033-1

Tsang, C. D., Falk, S., & Hessel, A. (2017). Infants Prefer Infant-Directed Song Over Speech. *Child Development*, *88*(4), 1207–1215. https://doi.org/10.1111/cdev.12647

Turk, A., & Shattuck-Hufnagel, S. (2013). What is speech rhythm? A commentary on Arvaniti and Rodriquez, Krivokapić, and Goswami and Leong. *Laboratory Phonology*, *4*(1), 93–118. https://doi.org/10.1515/lp-2013-0005

van Alphen, P., Brouwer, S. M., Davids, N., Dijkstra, E., & Fikkert, P. (2021). Word recognition and word prediction in preschoolers with (a suspicion of) a Developmental Language Disorder: Evidence from eye-tracking. *Journal of Speech Language and Hearing Research*, *64*, 2005–2021.

van den Bosch der Nederlanden, C. M., Joanisse, M., & Grahn, J. A. (2020). Music as a scaffold for listening to speech: Better neural phase-locking to song than speech. *NeuroImage*, *214*(May 2019), 116767. https://doi.org/10.1016/j.neuroimage.2020.116767

van der Velde, B., & Junge, C. (2020). Limiting data loss in infant EEG: putting hunches to the test. *Developmental Cognitive Neuroscience*, *45*(May), 100809. https://doi.org/10.1016/j.dcn.2020.100809

van Goch, M. M., McQueen, J. M., & Verhoeven, L. (2014). Learning Phonologically Specific New Words Fosters Rhyme Awareness in Dutch Preliterate Children. *Scientific Studies of Reading*, *18*(3), 155–172. https://doi.org/10.1080/10888438.2013.827199

Vaughan-Evans, A., Trefor, R., Jones, L., Lynch, P., Jones, M. W., & Thierry, G. (2016). Implicit detection of poetic harmony by the Naïve brain. *Frontiers in Psychology*, *7*(NOV), 1–7. https://doi.org/10.3389/fpsyg.2016.01859

Virtala, P., & Partanen, E. (2018). Can very early music interventions promote at-risk infants' development? *Annals of the New York Academy of Sciences*, *1423*(1), 92–101. https://doi.org/10.1111/nyas.13646

Vloedgraven, J. M. T., & Verhoeven, L. (2007). Screening of phonological awareness in the early elementary grades: An IRT approach. *Annals of Dyslexia*, *57*(1), 33–50. https://doi.org/10.1007/s11881-007-0001-2

Volkova, A., Trehub, S. E., & Schellenberg, E. G. (2006). Infants' memory for musical performances. *Developmental Science*, *9*(6), 583–589. https://doi.org/10.1111/j.1467-7687.2006.00536.x

Wagensveld, B., Segers, E., Alphen, P. Van, Hagoort, P., & Verhoeven, L. (2012). A neurocognitive perspective on rhyme awareness: the N450 rhyme effect. *Brain Research*, *1483*, 63–70. https://doi.org/10.1016/j.brainres.2012.09.018

Wagensveld, B., Segers, E., van Alphen, P., & Verhoeven, L. (2013). The role of lexical representations and phonological overlap in rhyme judgments of beginning, intermediate and advanced readers. *Learning and Individual Differences*, *23*, 64–71. https://doi.org/10.1016/j.lindif.2012.09.007

Wagensveld, B., van Alphen, P., Segers, E., Hagoort, P., & Verhoeven, L. (2013). The neural correlates of rhyme awareness in preliterate and literate children. *Clinical Neurophysiology*, *124*(7), 1336–1345. https://doi.org/10.1016/j.clinph.2013.01.022

Wagner, M., & Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*, *25*(7), 905–945. https://doi.org/10.1080/01690961003589492

Wallot, S., & Menninghaus, W. (2018). Ambiguity effects of rhyme and meter. *Journal of Experimental Psychology: Learning Memory and Cognition*, *44*(12), 1947–1954. https://doi.org/10.1037/xlm0000557

Wang, Y., Seidl, A., & Cristia, A. (2021). Infant speech perception and cognitive skills as predictors of later vocabulary. *Infant Behavior and Development*, *62*(February), 101524. https://doi.org/10.1016/j.infbeh.2020.101524

Wassiliwizky, E., Koelsch, S., Wagner, V., Jacobsen, T., & Menninghaus, W. (2017). The emotional power of poetry: Neural circuitry, psychophysiology and compositional principles. *Social Cognitive and Affective Neuroscience*, *12*(8), 1229–1240. https://doi.org/10.1093/scan/nsx069

Wellmann, C., Holzgrefe, J., Truckenbrodt, H., Wartenburger, I., & Höhle, B. (2012). How each prosodic boundary cue matters: evidence from german infants. *Frontiers in Psychology*, *3*(December), 580. https://doi.org/10.3389/fpsyg.2012.00580

Werker, J. F., Polka, L., & Pegg, J. (1997). The Conditioned Head Turn Procedure as a Method for Testing Infant Speech Perception. *Early Development and Parenting*, *6*(September 1996), 171–178. Retrieved from http://infantstudies.psych.ubc.ca/uploads/forms/1252960095WerkerPolkaPegg_1997.pdf

Wood, C. (2006). Metrical stress sensitivity in young children and its relationship to phonological awareness and reading. *Journal of Research in Reading*, *29*(3), 270–287. https://doi.org/10.1111/j.1467-9817.2006.00308.x

Wood, C., & Terrell, C. (1998). Preschool Phonological Awareness and Subsequent Literacy Development. *Educational Psychology*, *18*(March 2015), 253–274. https://doi.org/10.1080/0144341980180301

Woodruff Carr, K., White-Schwoch, T., Tierney, A., Strait, D. L., & Kraus, N. (2014). Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(40), 14559–14564. https://doi.org/10.1073/pnas.1406219111

Yamane, N., Sato, Y., Shimura, Y., & Mazuka, R. (2021). Developmental differences in the hemodynamic response to changes in lyrics and melodies by 4- and 12-month-old infants. *Cognition*, *213*(October 2020), 104711. https://doi.org/10.1016/j.cognition.2021.104711

Yoncheva, Y., Maurer, U., Zevin, J. D., & McCandliss, B. D. (2013). Effects of rhyme and spelling patterns on auditory word ERPs depend on selective attention to phonology. *Brain and Language*, *124*(3), 238–243. https://doi.org/10.1016/j.bandl.2012.12.013

Yoshida, K. A., Fennell, C. T., Swingley, D., & Werker, J. F. (2009). Fourteen-month-old infants learn similar-sounding words. *Developmental Science*, *12*(3), 412–418. https://doi.org/10.1111/j.1467-7687.2008.00789.x

Zamuner, T. S. (2009). The structure and nature of phonological neighbourhoods in children's early lexicons. *Journal of Child Language*, *36*(1), 3–21. https://doi.org/10.1017/S0305000908008829

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Zink, I., & Lejaegere, M. (2002). *N-CDIs: Lijsten voor communicatieve ontwikkeling.* Leeuven/Leusden: Acco.

Zion Golumbic, E. M., Poeppel, D., & Schroeder, C. E. (2012). Temporal context in speech processing and attentional stream selection: a behavioral and neural perspective. *Brain and Language*, *122*(3), 151–161. https://doi.org/10.1016/j.bandl.2011.12.010

Nederlandse samenvatting

Ouders en verzorgers op de hele wereld maken gebruik van taal op een speelse manier: ze rijmen en zingen voor hun kinderen. Deze intuïtieve interacties worden vaak rituelen en geven zo structuur aan het alledaags leven van een kind. Baby's bijvoorbeeld horen liedjes of rijmpjes in bad, voor het eten of bij het slapen gaan, maar ook tijdens feestelijke gelegenheden zoals verjaardagen. In zulke situaties worden liedjes en rijmpjes gebruikt om baby's te animeren of te kalmeren. Maar uiteraard is speels taalgebruik ook gewoon een uitdrukking van liefde voor een kindje.

Taal spelletjes met kleine kinderen zijn een heel bijzondere vorm van spraak. Ze bevatten poëtische kenmerken, zoals ritme, rijm en versregels. Ouders praten tijdens deze vorm van taal ook vaak langzamer, met meer herhaling en met een hogere liefdevollere stem dan tijdens andere omstandigheden. Speels taalgebruik bevat dus veel structuur, wat het aantrekkelijk en makkelijk te voorspellen maakt voor kleine kinderen. Tot nu toe was niet duidelijk of baby's eigenlijk linguïstische patronen in liedjes en rijmpjes waarnemen en of ze daarbij profijt hebben van de sterke akoestische structuur. De hoofdvraag van dit proefschrift was dus of baby's typische linguïstische klank patronen in speels taalgebruik waarnemen, namelijk rijm, ritme en de grenzen van frasen. Deze patronen komen ook in elk andere form van spraak voor, maar dan niet binnen de aantrekkelijke omgeving die liedjes en rijmpjes bieden.

In hoofdstuk 2 hebben we onderzocht of baby's kinderliedjes in akoestische eenheden segmenteren, namelijk frasen. Deze frasen komen overeen met regels van een vers in een liedje. In alledaagse spraak corresponderen deze frasen vaak met zinnen en vormen dan belangrijke structurele en semantische eenheden. Sprekers markeren de grenzen tussen deze eenheden door langzamer te praten, een kleine stilte te laten vallen en door veranderingen in de toonhoogte. Uit eerder onderzoek weten we dat baby's al vanaf zes maanden oud deze akoestische signalen combineren en zo frasen in een rij van zinnen herkennen. In ons onderzoek hebben we deze bevinding gerepliceerd en uitgebreid tot kinderliedjes. Hier markeren dezelfde akoestische signalen als in gewone spraak de grenzen tussen opeenvolgende frasen. In onze studie waren 6-maanden oude baby's in staat om frasen van een liedje te herkennen. Ze luisterden langer naar een liedje dat een rij woorden als gehele frasen bevat dan naar een liedje met dezelfde rij woorden maar dan niet als hele frase gezongen. Waarschijnlijk gebruiken baby's dus bij het segmenteren van liedjes dezelfde strategie als bij het segmenteren van spraak, ze combineren akoestische signalen en herkennen zo belangrijke structurele eenheden.

Hoofdstuk 3, 4 en 6 zijn een verzameling van bevindingen over het waarnemen van rijm in liedjes en gesproken rijmpjes bij baby`s. Rijm is een typisch taalpatroon in poëtische taal: op het einde van opeenvolgende versregels wordt het laatste deel van een lettergreep herhaalt, zoals in *beer* en *peer*. Het herkennen van deze rijmrelatie tussen woorden is één fundament voor de taalontwikkeling van kleuters. Zij gebruiken rijm om te leren lezen en schrijven. Wij vroegen ons af of rijm in liedjes überhaupt al waargenomen wordt door jonge baby's en of er een verband is tussen de algemene taalontwikkeling van een kind en deze vroege vorm van rijm waarneming. In onze studies luisterden baby's naar rijmende of niet-rijmende versjes. Deze waren of gezongen (hoofdstuk 3 en 4) of gesproken (hoofdstuk 6). In hoofdstuk 6 hebben we daarnaast ook het ritme van de versjes gevarieerd. De rijmwoorden zaten dan óf in een typisch versje met een regelmatig ritme, óf in een versje waar het ritme onregelmatig en minder goed te voorspellen was. Terwijl onze proefpersonen naar deze stimuli luisterden hebben we hun luistertijd (hoofdstuk 3) of hun hersenen activiteit (EEG) gemeten (hoofdstuk 4 en 6). In zowel luistertijd als eeg-resultaten vonden we verschillen tussen rijmende en niet-rijmende stimuli, echter waren deze verschillen niet altijd statistisch significant. Ook het veranderen van het ritme van een versje bleek geen invloed op het waarnemen van rijm te hebben. Deze bevindingen vragen dus voor een herhaling door toekomstig onderzoek. Daarnaast hebben we een verband gevonden tussen de gevoeligheid voor rijm in een liedje en de grote van de woordenschat enkele maanden later. Dit toont aan dat baby's er profijt van kunnen hebben in hun taalontwikkeling als hun ouders voor ze rijmen en zingen.

In hoofdstuk 5 hebben we rijmwaarneming van volwassenen onderzocht door wederom elektrische hersenactiviteit te meten. Ook zij luisterden naar rijmende en niet-rijmende versjes en weer was het ritme van elk versje óf regelmatig óf onregelmatig. Eerder onderzoek keek vooral naar rijmeffecten in losse woorden, maar nauwelijks naar poëtische spraak. In deze eerdere studies moesten proefpersonen rijm altijd actief verwerken, door bijvoorbeeld elke keer op een knop te drukken als iets rijmde. Ons onderzoek is een van weinige studies waar rijm in een natuurlijke taalomgeving voorkwam. De hersenactiviteit voor rijmende woorden was positiever dan voor niet-rijmende woorden en de vorm van dit effect leek erg op de effecten die in eerder onderzoek gevonden werden. Het herhalen van fonologisch materiaal tussen verschillende regels van een versje wordt dus onbewust waargenomen door volwassen luisteraars. We vonden ook een tweede rijm effect: op het moment dat er voor het eerst rijm in een versje voorkwam, zagen wij een negativiteit voor rijmende woorden. Dit effect begon ietsjes eerder in versjes die een onregelmatig ritme hadden. Waarschijnlijk is dit vroege rijmeffect een gevolg van gebruik van het werkgeheugen. Om het verband tussen rijmende regels te creëren moeten luisteraars namelijk het einde van de eerste regel van een vers onthouden en vergelijken met de volgende regel.

Naast de rijm effecten die hierboven samengevat zijn, hebben we in hoofdstuk 6 ook naar hersengolven van 10.5-maanden-oude baby's gekeken. Linguïstische eenheden zoals woorden en lettergrepen komen in spraak in verschillende ritmes voor. De frequentie waarmee deze eenheden optreden wordt door hersengolven gevolgd. Er bestaat dus een overeenkomst tussen spraakritme en hersenactiviteit. Voor kleine kinderen is deze overeenkomst nog nauwelijks onderzocht. Wij vroegen ons af of baby's bij het volgen van linguïstische eenheden profijt hebben van het sterke spraak ritme in een versje. We hebben daarom de spraak-hersenen relatie vergeleken voor ritmische en minder-ritmische versjes en vervolgens onderzocht of er verbanden zijn in de sterkte van deze relatie en de eeg-rijmeffecten en de woordenschat van een kind. Voor zowel ritmische én minder-ritmische versjes hebben we een duidelijke spraak-hersenen relatie gevonden, echter was deze sterker voor minder ritmische versjes. Dit verrassend resultaat heeft er wellicht mee te maken dat onze proefpersonen al bijna één jaar oud waren. Mogelijk is het profijt van een sterk taal ritme groter voor jongere baby´s, die meer moeite hebben met het volgen van alledaagse spraak. De relatie tussen spraak en hersenritmes was bijzonder duidelijk voor het ritme van lettergrepen. Voor de relatie van de hersenen met het ritme van deze eenheden hebben we dan ook verbanden met de woordenschat van een kind gevonden. Net als in hoofdstuk 4 hebben we dus laten zien dat de verwerking van taal patronen in speels taalgebruik (hier lettergrepen) mogelijk een voordeel heeft voor de taalontwikkeling van kleine kinderen.

De bevindingen in dit proefschrift tonen aan dat baby's gevoelig zijn voor het ritme, rijm en frasen van liedjes en rijmpjes. De waarneming van deze taalpatronen staat in verband met de woordenschat van baby's. Ook volwassen luisteraars nemen deze patronen in poëzie onbewust waar. Rijmen en zingen voor jonge baby`s is dus niet alleen een uiting van positieve emoties, het is ook een manier om ze spelenderwijs bekend te maken met belangrijke patronen van hun moedertaal.

English summary

Parents and caregivers all over the world make use of language play. They rhyme and sing songs for their children and engage them in bouncing games or finger plays. These intuitive interactions often become verbal rituals that structure children's everyday life. Babies, for example, are exposed to songs and nursery rhymes during diaper change or bath time, prior to meals or before being put to bed, but also during celebrations like birthdays. In situations like these songs and nursery rhymes are used to regulate infants' arousal state, for example to calm them prior to sleeping. But obviously they are also merely an expression of caregiver' love for their children.

From a linguistic perspective, language play is a very peculiar type of speech signal. It contains poetic features like rhyme, rhythm and verse lines. It is also often expressed much more slowly and with a higher and more loving tone of voice and it is more repetitive than speech during other situations. The language expressed in language play is thus much more structured, more attractive and predictable than everyday speech. For babies who are still learning their native language, language play might thus be easier to process than speech. Yet to date it has been unclear whether these young children perceive linguistic information provided in language play and whether they benefit from their clear acoustic structure.

The main research question of this dissertation was whether infants recognize linguistic sound patterns in oral language play, namely rhyme, rhythm and phrases. These sound patterns are not only prototypical in language play, recognizing each of them is also important for general language development outside of playful contexts.

In chapter 2 we asked whether infants segment songs into structural units, namely phrases. In everyday speech, intonational phrases often correspond to sentences and function as important structural and semantic units. Speakers highlight the boundaries between phrases by slowing down, pausing and changing their pitch. Previous research proves that six-month-olds already combine these acoustic boundary cues to parse incoming speech into smaller grammatical constituents. Here we replicated this finding and extended it to songs. In song melodies, the same acoustic cues mark the boundaries between melodic phrases. Six-month-olds in our study recognized song phrases, by showing longer listening times for a song containing a word sequence within a phrase, compared to a song containing the same sequence interrupted halfway through. For infants, sensitivity to melodic phrases in songs might be another steppingstone into parsing of their native language. Interestingly, we found no difference in listening times between song and speech stimuli, indicating that for infants songs are not easier to segment than speech.

Chapters 3, 4 and 6 asked whether infants detect rhymes in child songs and spoken nursery rhymes. Being able to associate rhyming words with each other is an important prerequisite for the development of literacy. We wondered whether a precursor of this rhyme awareness could be found in infants' sensitivity to rhymes in language play. We also explored whether the predictable rhythm of nursery rhymes has a positive influence on infants' rhyme perception. These hypotheses were tested by exposing infants to rhyming as well as non-rhyming songs and nursery rhymes and measuring their listening times (chapter 3) and their electrical brain activity (chapters 4 and 6). We also manipulated whether rhymes were occurring in nursery rhymes that had a regular or an irregular rhythm of stressed and unstressed syllables (chapter 6). Infants in our studies showed shorter listening times for rhyming versus non-rhyming songs and more negative EEG-potentials for repeating rhymes in songs and nursery rhymes. Regular rhythm did not alter infants' cortical rhyme response (chapter 6). We also observed a link between 10.5-month-olds' perception of rhymes in child songs and their vocabulary size half a year later (chapter 4). Not all of our results are generalizable to children outside of the lab and therefore require corroboration by future studies. However, at this point they already suggest that infants recognize rhymes in language play and this might be beneficial for their later language development.

In chapter 5, we studied adults' rhyme perception in poetry, using the same paradigm and stimuli as for the infants (chapter 6). Our participants listened to four-line nursery rhymes with a rhyming or non-rhyming word the end of each line, while their electric brain potentials were recorded. Rhymes were either occurring in nursery rhymes that had a typical predictable rhythm (metered) or in a nursery rhyme where the rhythm was disrupted and less predictable (non-metered). Adults' recognition of rhymes at the end of each nursery rhyme line was reflected in more positive electric potentials for rhymes compared to non-rhymes starting around 400 ms after word onset. This effect largely resembles results from previous studies where rhymes were presented in word pairs and probably reflects the subconscious detection of repeating phonological material in speech. At the end of line 2 of the nursery rhyme stimuli, an additional anterior rhyme negativity was found, that started slightly earlier in the non-metered nursery rhymes. We interpreted this effect as a reflection of memory processes involved in matching rhyming verse lines with each other. So far, adults' rhyme perception was mainly studied with rhyming word pairs, which listeners had to judge on whether they are rhyming or not. Our study provided insights into rhyme processing in more naturalistic poetic contexts, wherein rhymes typically occur in rhythmic sentences. Moreover, our participants were not required to judge our stimuli, rendering our paradigm more simple and thus of potential interest for clinical or other developmental populations.

In chapter 6 we again recorded electric brain potentials from 10.5-month-olds listening to rhyming and non-rhyming nursery rhymes that either had a predictable rhythm of stressed syllables alternating with unstressed syllables (metered) or were less predictable (non-metered). In addition to the rhyme effects summarized above, we also analysed infants’ neural tracking of the acoustic speech envelop. Recent advances in cognitive neuroscience indicate that speech rhythm provided in the amplitude envelop is mirrored in neural activity, as phonological units such as phrases, stress, syllables and phonemes, each occurring at different time scales, are reflected in neural activity in specific oscillatory frequency bands. This relationship between speech and brain rhythms might be particularly strong for rhythmic infant-directed speech, such as nursery rhymes. In our study, infants’ speech-brain coherence (*SBC*) was employed as an EEG index for synchrony between infants’ brain oscillations and the speech amplitude envelope. Infants’ neural activity mirrored the stress and syllable rhythm of the nursery rhymes. Contrary to our hypothesis, however, the resemblance between speech and brain rhythms was higher for non-metered nursery rhymes, especially at the syllable rate. This surprising finding might be attributable to the fact that our participants were already around one year old. Younger infants might show the hypothesized processing benefit for stressed syllables occurring with a regular rhythm as in our metered nursery rhymes. Infants’ neural tracking of the syllable rhythm in nursery rhymes had a negative relationship with concurrent vocabulary (10.5 months), but was positively correlated with future vocabulary (18 months). A syllabic processing focus at the end of the first year of life might thus be associated with different effects on concurrent and future vocabulary size.

The studies reported in this dissertation provide evidence for the perception of rhyme, rhythm and phrases in oral language play in infant and adult listeners. Our experimental studies converge in the notion that the linguistic information transmitted in oral language play is not merely a sea of sound for infants (and adults). Instead, listeners are sensitive to the phonological, structural and rhythmic patterns transmitted in the acoustic shape of oral language play, which might have consequences for their language development.

Deutsche Zusammenfassung

Eltern auf der ganzen Welt singen und reimen für ihre Kinder. Ganz intuitiv benutzen sie diese Sprachspiele in der Interaktion mit ihren Kindern und entwickeln sie häufig zu festen Ritualen. Babys zum Beispiel hören Lieder und Reime während sie gewickelt werden, vor den Mahlzeiten oder dem Schlafen gehen, aber auch an Geburtstagen. Je nach Situation werden Sprachspiele dann genutzt um Kinder zu beruhigen oder zu animieren. Natürlich sind sie auch ein Mittel um elterliche Liebe und Fürsorge auszudrücken.

Sprachspiele wie Lieder und Reime stellen einen ganz besonderen sprachlichen Stimulus für Kinder dar. Sie enthalten poetische Komponenten wie Reim, Rhythmus und Verszeilen. Sie werden meist auch viel langsamer und mit einer hohen liebevollen Stimme ausgesprochen und enthalten viel mehr Wiederholungen als Sprache die Kinder in anderen Situationen hören. Die Sprache in Sprachspielen ist also viel strukturierter, attraktiver und leichter vorhersagbar als in anderen Situationen. Für Babys und Kleinkinder, die ihre Muttersprache ja noch erlernen müssen, könnte diese simple Struktur von Vorteil sein. Bis heute war jedoch unklar, ob Kinder in diesem jungen Alter überhaupt sprachliche Information in Liedern und Reimen wahrnehmen und ob sie einen Nutzen aus dem simplen Aufbau dieser Sprachspiele ziehen.

Die vorliegende Arbeit beschäftigt sich vor allem mit der Frage, ob Babys die folgenden sprachlichen Muster in Liedern und Reimen wahrnehmen: Reim, Rhythmus und Phrasen. Diese Muster sind nicht nur typisch für Sprachspiele. Die Wahrnehmung dieser sprachlichen Muster in Sprache außerhalb des spielerischen Kontext ist von großer Relevanz für den Verlauf der weiteren Sprachentwicklung von Kindern.

In Kapitel 2 haben wir erforscht, ob Babys in Kinderliedern strukturelle Einheiten, nämlich Phrasen, erkennen können. In gesprochener Sprache bilden Sätze sogenannte Intonationsphrasen. In ihnen befindet sich zusammengehörende semantische und syntaktische Information. Sprecher markieren die Grenzen zwischen diesen Phrasen in dem sie etwas langsamer werden, die Stimme senken und kurz pausieren, bevor sie eine neue Phrase beginnen. Aus früheren Studien war bereits bekannt, das 6-Monate alte Säuglinge diese akustischen Hinweisreize bemerken und so Sätze in gesprochener Sprache erkennen können. Wir haben diesen Befund repliziert und erweitert auf Kinderlieder. In Melodien werden Phrasen mit den gleichen akustischen Mitteln markiert wie in gesprochener Sprache. Die Säuglinge in unserer Studie zeigten eine Präferenz für Lieder die eine Wortreihe als vollständige Phrase enthielten im Vergleich zu Liedern in der die gleiche Wortreihe mittendrin unterbrochen wurde. Diese Präferenz war für Kinderlieder jedoch nicht stärker als für gesprochene Sprache. Für Kinder im ersten Lebensjahr sind Kinderlieder also nicht unbedingt leichter zu segmentieren als gesprochene Sprache.

Die Kapitel 3, 4 und 6 beschäftigten sich mit der Frage ob Säuglinge Reime in Liedern und gesprochenen Versen bemerken. Reime sind nicht nur typischer Bestandteil von Sprachspielen. Zu erkennen das sich einzelne Worte reimen, erleichtert älteren Kindern auch den Erwerb der Schriftsprache. Möglicherweise tragen reimende Sprachspiele bei zu einer frühen Form dieser Reim-Bewusstheit. Wir haben auch untersucht, ob ein regelmäßiger Sprachrhythmus, wie er typisch ist für Kinderlieder und Verse, von Vorteil ist für das Erkennen von Reimen. In unseren Studien hörten Säuglinge reimende und nicht-reimende Kinderlieder oder gesprochene Verse während wir ihre Hörpräferenz (Kapitel 3) und ihre Hirnströme (EEG) aufgezeichnet haben (Kapitel 4 und 6). In Kapitel 6 haben wir zusätzlich den Sprachrhythmus manipuliert. Reime traten dann entweder in einem Vers mit typischen regelmäßigen Wechsel aus betonten und unbetonten Silben auf, oder in Versen mit weniger regelmäßigem Rhythmus. Die Säuglinge in unseren Studien zeigten kürzere Hörzeiten (Kapitel 3) und negativere Hirnpotentiale in ihrem EEG (Kapitel 4 und 6) für reimende versus nicht-reimende Stimuli. Ein regelmäßiger Sprachrhythmus schien keinen Einfluss auf die Reimwahrnehmung zu haben. Darüber hinaus zeigte sich eine Korrelation zwischen der Wahrnehmung von Reimen in Kinderliedern im Alter von 10.5 Monaten und dem Wortschatz der Kinder ein halbes Jahr später. Nicht all unsere Ergebnisse waren statistisch signifikant. Sie lassen sich deshalb nicht verallgemeinern auf die Sprachverarbeitung von Säuglingen außerhalb unser Studie. Trotzdem weisen diese Befunde bereits jetzt darauf hin, dass Reime in Sprachspielen in jungem Alter wahrgenommen werden und dass diese Wahrnehmung einen positiven Einfluss auf die weitere Sprachentwicklung, insbesondere den Wortschatz, haben kann.

In Kapitel 5 haben wir eine ähnliche Studie mit Erwachsenen durchgeführt. Auch diese Probanden hörten reimende und nicht-reimende Verse mit einem regelmäßigen oder weniger regelmäßigen Rhythmus, während wir ihre Hirnaktivität mittels EEG aufgezeichnet haben. Die Wahrnehmung des Reims am Zeilenende der Verse zeigte sich in einem positiveren Hirnpotential für reimende versus nicht-reimende Worte. Dieser Effekt ähnelt Ergebnissen aus früheren Studien in denen Reim mit einzelnen Worten untersucht wurde und spiegelt die unbewusste Wahrnehmung von wiederholendem phonologischem Sprachmaterial wider. Das Auftreten des ersten Reims jedes Verses wurde begleitet von einem negativen Potential für Reime versus Nicht-Reime. Dieser Effekt setzte etwas früher in nicht-rhythmischen Versen ein. Effekte dieser Art sind typisch für ein aktives Arbeitsgedächtnis. Um den Reim zu bemerken, mussten unsere Probanden das Ende der ersten Verszeile im Gedächtnis behalten und diese mit dem Ende der zweiten Verszeile vergleichen. Bisher wurde Reimwahrnehmung von Erwachsenen vor allem mit aktiven Paradigmen untersucht, in denen einzelne Worte in Hinsicht auf Reim/Nicht-Reim beurteilt werden mussten. Unsere Studie dagegen verwendet als eine der wenigen natürliche poetische Sprache in einem simplen passiven Paradigma. Dieses könnte zum Beispiel zukünftigen Studien mit Kleinkindern oder Patienten zu Gute kommen.

In Kapitel 6 haben wir neben den bereits erwähnten Reimeffekten auch die Wahrnehmung von Sprachrhythmus im Säuglingsalter untersucht. Aktuelle Befunde aus der Kognitionswissenschaft deuten darauf hin, dass unser Gehirn akustischen Rhythmen folgt, indem es seine neuronalen Oszillationen auf die Frequenz dieser Rhythmen einstellt. Wir erfassen also sprachliche Einheiten wie Silben, Worte und Phrasen in dem sich die Hirnaktivität am Rhythmus dieser Einheiten ausrichtet. Dieser Zusammenhang zwischen sprachlichen und neuronalen Rhythmen könnte besonders stark sein für sprachliche Stimuli die einen sehr auffälligen Sprachrhythmus aufweisen, zum Beispiel Kinderreime. In diesen findet sich typischerweise ein regelmäßiger Wechsel zwischen betonten und unbetonten Silben. Wir haben untersucht, ob sich eine besonders starke Kohärenz finden lässt zwischen der neuronalen Aktivität von 10.5-Monate alten Säuglingen und dem Rhythmus von Kinderreimen. Darüber hinaus wollten wir wissen, ob sich Zusammenhänge zwischen dieser Sprach-Hirn-Kohärenz und dem Wortschatz oder der Reimwahrnehmung im EEG finden lassen. Tatsächlich spiegelte sich der Sprachrhythmus der Kinderreime in der neuronalen Aktivität der Säuglinge wider. Allerdings war diese Kohärenz stärker für die weniger rhythmischen Kinderreime, insbesondere auf dem Niveau des Silbenrhythmus. Dieser überraschende Befund könnte damit zusammenhängen, dass unsere Probanden bereits fast ein Jahr alt waren. Wahrscheinlich lässt sich ein Vorteil für die Sprachverarbeitung von rhythmischen Kinderreimen eher bei jüngeren Kindern nachweisen, denen die weniger rhythmische Sprache außerhalb von Sprachspielen mehr Mühe bereitet. Die Güte mit dem der Silbenrhythmus der Kinderreime verfolgt wurde, korrelierte mit dem Wortschatz unserer Probanden. Es zeigte sich allerdings kein Effekt auf das Wahrnehmen des Reims. Auch diese Studie deutet also darauf hin, dass Sprachspiele einen positiven Beitrag zum frühkindlichen Spracherwerb leisten können.

Die Befunde dieser Arbeit zeigen, dass Kinder bereits während des ersten Lebensjahrs Phrasen, Reime und Rhythmen in Kinderliedern und Versen bemerken. Die Wahrnehmung dieser sprachlichen Muster steht in Zusammenhang mit ihrem Wortschatz. Auch Erwachsene verfolgen unbewusst das auftreten dieser Muster in poetischer Sprache. Singen und Reimen mit Kindern ist also nicht nur ein Mittel um positive Emotionen mit Babys zu teilen oder ihren Gemütszustand zu beeinflussen. Sie ermöglichen, Babys in einem spielerischen Kontext bekannt zu machen mit wichtigen sprachlichen Mustern ihrer Muttersprache.

Acknowledgements, Dankwoord, Danksagung

op een dag was nijntje pluis
al heel vroeg opgestaan
zij waste zich van top tot teen
ziezo, dat was gedaan

zij zocht een heel mooi jurkje uit
het mooiste uit de kast
en weet je waarom zij dat deed
omdat zij jarig was

… envisioning my upcoming defence and the people that have helped me get this far, makes me feel like little birthday Nijntje.

Thank you to the many **families and children** that have participated in my studies! Thank you babies for remaining patient during long test sessions, thank you parents and caregivers for finding a way to fit in babylab visits into your busy days! Your effort and enthusiasm for my studies have made me feel proud to be a researcher.

Paula, Titia en Tineke, mijn dank kent geen grenzen. Ik had geen betere begeleiders kunnen verzinnen. Elk contact moment met jullie liet mij verder denken en nieuwe onderzoekspaden verkennen, maar leverde ook gewoon plezier en vertrouwen op. Your (many) emails and our shared memories will remain a treasure box of scientific and of more general wisdom for me. Thank you for your trust in me, for always being available, for granting me all the time I needed to finish (a lot, apparently), for never pushing me and for making the supervision trio run so smoothly. Lucky me, being able to enjoy the wisdom of all three of you. I hope our paths will keep on crossing each other for many years to come.

Tineke: Dank voor onze mooie dagen en lange gesprekken in Bremen en Lyon, dank dat je met mij en mijn dikke Mara-buik naar Yosemite bent gevlogen. Dank voor al het last minute feedback vlak voor mijn submissions en het uitgebreide uitleg dat je altijd nog op een van de manier in je overvolle werkdag wist te proppen. Dank dat je mij kennis liet maken met Nijntje, Annie MG Smidt, Jip en Janneke en jullie Sinterklaas tradities. You are a great mentor, not only for my infant EEG adventures, but for so much more.

Titia: Thank you for ALWAYS thinking along and for reading my manuscripts so thoroughly, despite of your crazy teaching load. Your knowledge and work ethics combined with your trust in my abilities were the combination I needed to reach my goals, especially during dark months of sleep deprivation after the girls were born. It was great to start the workday with an encouraging email from you. You taught me to think scientifically and how to come up with more precise arguments. Somehow magically, though far away, you were always close.

Paula: Thank you for keeping it all together over these long past years, for pushing the right administrative buttons (and people) so that I could flourish, for travelling together to Terschelling, Ghent, Den Haag, Amsterdam, Potsdam and Lyon and enjoying plenty of good food, long walks and conversations on our way (and coffee of course). Thank you for nourishing me with self-confidence when my spirits were low. I know I am repeating others here, but you really are a great role model. **Theo**, dank dat ik zo veel tijd met je lieve Paula mocht door brengen en dank dat jij mijn promotie wilt voorzitten!

Caroline Junge, Caroline Rowland, Natalie Boll-Avetisyan, Marc van Oostendorp and James McQueen, thank you for assessing my dissertation and for acting as opponents during my defence. This book contributed greatly from your research and/or from you personally. It is a an honour to now be able to discuss my work in person with you. **Christina Bergmann and Roel Willems**, thank you for joining the committee! I feel my corona now perfectly combines experts on every discipline related to my studies.

Christina Bergmann, thank you for being the cheerful and supportive person you are and for committing so much to open and meta science.

Natalie Boll-Avetisyan, wenn Du meine slides vom Master colloqium damals nicht mit Titia geteilt hättest, wäre dieses Buch nie zu Stande gekommen. Danke für Dein Vertrauen in mich und für die Türen die Du mir geöffnet hast.

Caroline Rowland, thank you for making the LADD at the MPI such a great place for child language research and for allowing us to join LADD group meetings. I really miss this club of brilliant minds.

Eva und Jil: es tut so gut euch Beide an meiner Seite zu wissen während meiner Verteidigung. Vielen Dank für Eure Bereitschaft, diesen Dienst zu übernehmen.

Dear **FLA-group** within the Centre for Language Studies: Thanks for having me! I learned a lot from you sharing your research trials and tribulations and from your success. **Stefanie, Antje, Sho, Helen, Christina, Hannah, Nadine, Nienke, Imme** and many others, it was great to grow into science with you being around.

Sabine Hunnius, Thank you for hosting the very constructive babylab meetings and for always asking about the underlying mechanisms behind my studies. I feel I still owe you an answer to that.

Angela Khadar and Patricia Manko, thank you for managing my research and the studies of so many others so smoothly. The two of you were so dear to me, I always kept lingering in your offices much longer than was necessary.

Jessica Askamp, Tildie Stijns and the whole DCCN-administration, thank you for allowing me in at the DCCN. It was amazing to experience how easy it can be to set up a study as a guest of your institute.

My dear interns and assistants, **Katharina, Francie, Maaike, Mareike, Demi, Fleur, Iris, Julia, Malou, Anna, Emma and Lisa**, thank you for your help with data collection and annotations, for being the most professional baby-entertainers ever and for brightening up those long lab hours.

Maartje Stroo: het was een plezier je scriptie begeleider te zijn. Ik ben heel trots op ons hoofdstuk 5!

Katharina Menn: auf Dich bin ich auch so stolz und ein kleines Bisschen erleichtert, dass Du nun doch nicht vor mir Frau Doktor geworden bist. Es ist eine Freude zu erleben, wie Du Deinen Weg gehst.

Thank you **CLS board** for granting me extra time and money to pursue the research reported here. **Kevin Lam, Dirkje van der Aa and Tanja Döller,** thank you for being so extremely supportive and for cheering me up during months of insecurity.

Thank you **Jo, Ricci and Irati** for our beautiful Nijmegen home. Ricci, Danke, dass Du mich Familie Humann-Becker vorgestellt hast! Irati, Danke für Deine treuen Besuche in Kleve und Strausberg. Ich hoffe sehr, dass wir uns bald zu Dir auf die Reise machen können.

Claudia Lüttke, Danke für Deine Freundschaft. Nach unseren verkopften Working-Mum-Gesprächen ist es nun höchste Zeit, dass wir endlich mit Jil Achterbahn fahren gehen!!!

Varun (and Amaia), thanks for always reaching out, so that we could enjoy another discussion about life and science, while strolling around park Brakkenstein.

Lieke en Alan, Fons en Marieke: Dank voor onze gezellige tijden in 6.23! Thank you for sharing the ups and downs of PhD life, while providing me with valuable insights into Dutch language and culture, silly jokes and home-made cake. (Maarten, Adriaan en Paul, dank ook voor jullie bijdrage!) Fons, thank you for the small talk and the not so small talk.

Dank aan **alle collega's binnen de afdeling Nederlandse Taal en Cultuur**, especially **Chara, Hongling, Afrooz and Susanne Brouwer.** Susanne, thank you for your everlasting positivity and for advertising yoga with Adrienne! These videos really helped me staying sane during the lockdown months.

Sabine Jentges, Yvonne Delhey, Eva Knopp und Jana Hermann und alle anderen in der **afdeling Duitse Taal en Cultuur**: es war ein Vergnügen bei euch zu arbeiten! Sabine und Eva, Danke für Eure Unterstützung und Euer Vertrauen während meiner ersten aufregenden Unterrichts-Erfahrungen; Jana, ich hab unseren PhD-chitchat über Daten, Analysen und das Leben außerhalb des Erasmusgebouw immer sehr genossen. Eva, wir waren erst Freunde und wurden dann Kollegen. Zum Glück hat Ersteres Bestand auch wenn uns nun so viele Kilometer trennen. Danke auch für Deine wunderbare Art mich einzuarbeiten.

Claudia Männel, Danke, für die Chance ein Teil Deiner Arbeitsgruppe an der Charité zu sein! Dein Vertrauen, Deine Inspiration und Dein unermüdliches Engagement für diese Gruppe haben mir enorm geholfen dieses Buch endlich abzuschließen. **Claudia, Evgenia, Nora, Dimitra und Christina**, Dank euch kann ich meine ersten vorsichtigen Post-Doc Schritte sehr genießen. Danke, dass wir so ein tolles Team geworden sind. **Sonja**, wat leuk om samen met jou eerste onderzoekservaring binnen een kliniek te maken. Dank voor het checken van mijn Nederlandse samenvatting!

Jil, Jonas, Ava und Jakob, ihr vier rockt! Wir hätten uns wohl nie getraut Berlin zu verlassen, wenn ihr nicht in Nijmegen auf uns gewartet hättet. Danke für all die Bottendaal/Kleve-Wochenenden und für Eure Freundschaft.

Florian Hintz, und Conni, Julius und Johannes (und Valentin!), Danke, dass ihr uns die Kleve-Zeit so versüßt habt mit gemeinsamen Unternehmungen. Mit euch war die Hürde der weit entfernten Großeltern plötzlich gar nicht mehr so groß, auf euch war einfach immer Verlass. Flo, Danke auch für unseren tollen fietstouren KLE-NMG and back!

Hanna: Danke für unsere Wochenendtrips mit den Kindern, das gemeinsame Frieren auf den winterlichen Klever-Spielplätzen, die Teestunden und das Mutmachen wenn es irgendwie nicht weiter ging.

Kristin, Ada und Fenna: Danke für die wunderbare Zeit die ihr Mara und Sina im Schneckenhaus bereitet habt und für unsere herrlichen Ausflüge. Wir vermissen euch sehr.

Jules, thank you for the carpool, fürs Zuhören, and for introducing me to Canarinho, Gazella, Faísca, Lobao and of course Mariposa. **Kristin**, vielen lieben Dank für den Kaffee!, Danke euch Beiden für die tollen Klamottenkisten für Mara und Sina.

Jule, Netti, Evi, Sarita, Sophia, Thesi und Luise, irgendwie wart ihr immer nah. Auch nach langer Funkstille ist mit euch sofort die alte Nähe da. Danke, dass ihr meine Orientierungslosigkeit ausgehalten habt. Wie gut euch nun wieder öfter zu sehen. (Evi, ich hoffe es klappt noch mit unserem Trip nach La Reunion).

Blair, my most remote friend, thank you for your company during a rather dark period. I hope we'll get to see each other soon.

Johannes, Danke für Deine Musik und Deine Kochkünste, Danke für unsere schöne gemeinsame Zeit!

Linni: Deine Sprach- und Farb-Künste, schmücken dieses Buch, Danke fürs gemeinsam Stolpern, Hinfallen und wieder Aufstehen, Danke für Deine Liebe und Deine Freundschaft.

Mama und Tanja, Danke für Eure regelmäßigen Fahrten zu uns an den Niederrhein. Unzählige Stunden habt ihr in Zügen und auf Bahnhöfen zugebracht, damit Alex und ich kurz verschnaufen können und Sina und Mara euch endlich um den Hals fallen können.

Tanja und Dete, Riki und Sandy, Ele und Paul (und Jakob!): was hab ich doch für ein Glück mit dieser Schwieger-Familie. Vielen vielen Dank für Euren Humor und Euren Zusammenhalt, der mich ganz selbstverständlich mit einschließt. Dete, Danke für die hervorragende Verpflegung!

Papa, Jana, Lotti, Mo und Lena: Danke, für unsere schönen Wochenenden und Eure Großzügigkeit. Danke, dass ihr stolz auf mich wart wenn ich es selbst gerade nicht sein konnte. **Jana**, Danke für unsere langen Spaziergänge.

Oma Gerda: Danke für die schönen Nachmittage zwischen Deinen Bildern und Büchern. Dein Sinn für das Schöne und für die Künste hat mich immer bestärkt, diese Dinge auch in meinem Alltag zu suchen. Ich hoffe sehr, dass wir uns, nun dieses Buch endlich fertig ist, wieder öfter sehen.

Andreas, Robert und Nicole, Mats und Jannis und Opa: wie gut, dass wir uns nun alle wieder in der Nähe haben. Ich freu mich auf ganz viel Familienzeit mit euch.

Mama: Wege entstehen indem wir sie gehen, hast Du mal gesagt. Dieses Buch ist der Beweis, dass Du wie immer Recht hattest. Danke, dass Du immer für mich da bist.

Alex, Mara und Sina, Ihr lasst mich jeden Tag lachen und genießen. Ich hoffe, dass wir noch viele Jahre abends laut singend durch die Küche tanzen.

Alex, der größte Dank gebührt Dir! Dieses Buch überstieg viel zu oft meine Kräfte, zum Glück warst Du dann da. Danke für die Freiheit die Du mir lässt, für Deine Treue und Dein Vertrauen und für unser schönes Zuhause. Ich freu mich auf viele weitere gemeinsame Jahre mit Dir.

toen moeder nijntje 's avonds laat
met beer naar bedje bracht
zei nijn: heel hartelijk bedankt
het was een fijne dag

from: “Nijntje is jarig” by Dick Bruna

Curriculum Vitae

Laura Hahn was born in 1988 in Rüdersdorf, Germany. She spent one year during high school with a host family in Rymarov, Czech Republic, an experience that sparked her interest in languages and linguistics. In 2010, she obtained her Bachelor’s degree in Linguistics and in 2013 her Master’s degree in Experimental and Clinical Linguistics, both from Potsdam University, Germany. While being a student, she worked as a research assistant for different psycho-linguistic research labs in Berlin and Potsdam and spent two months as an intern within the Neurobiology of Language Group at the Max Planck Institute for Psycholinguistics. Laura also gained some journalistic experience, while co-founding the Potsdam University radio FunkUP and during internships for Deutsche Presse Agentur and mephisto 97.6. In 2013, Laura began her PhD within the First Language Acquisition Group at the Centre for Language Studies at Radboud University. Upon finishing her dissertation, Laura worked as a lecturer in the study programme German Language and Culture at Radboud University. Currently, Laura is a Post-Doc in Prof. Claudia Männel’s Early Language Development group at Charité – Universitätsmedizin Berlin.

Author Publications

Hahn, L. E., Benders, T., Fikkert, P., & Snijders, T. M. (2021). Infants' implicit rhyme perception in child songs and its relationship with vocabulary. *Frontiers in Psychology*, *12*(680882), 1–8.

Hahn, L. E., Benders, T., Snijders, T. M., & Fikkert, P. (2020). Six-month-old infants recognize phrases in song and speech. *Infancy*, 699–718.

Frank, M. C., Alcock, K. J., Arias-Trejo, N., Aschersleben, G., Baldwin, D., Barbu, S., … **Hahn., L. E.,** … Soderstrom, M. (2020). Quantifying Sources of Variability in Infancy Research Using the Infant-Directed-Speech Preference. *Advances in Methods and Practices in Psychological Science*, *3*(1), 24–52.

Hahn, L. E., ten Buuren, M., Snijders, T. M., & Fikkert, P. (2020). Learning words in a second language while cycling and listening to children's songs: The Noplica Energy Center. *International Journal of Music in Early Childhood*, *15*(1), 95–108.

Hahn, L. E., Ten Buuren, M., De Nijs, M., Snijders, T. M., & Fikkert, P. (2019). Acquiring novel words in a second language through mutual play with child songs-The Noplica Energy Center. In *Counterpoints of the senses: Bodily experiences in musical learning A conference on Early Childhood Music Education (Meryc 2019)* (pp. 78-87). EuNet MERYC 2019.

Hahn, L. E., Benders, T., Snijders, T. M., & Fikkert, P. (2018). Infants' sensitivity to rhyme in songs. *Infant Behavior and Development*, *52*, 130–139.