



ARTICLE

Neural correlates of lexical-tone and vowel-quality processing in 6- and 9-month-old German-learning infants and adults

Antonia Götz^{1,2} , Claudia Männel^{3,4}, Gudrun Schwarzer⁵, Anna Krasotkina⁶ and Barbara Höhle² 

¹The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney, New South Wales, Australia, ²Linguistics Department, University of Potsdam, Potsdam, Germany, ³Department of Audiology and Phoniatrics, Charité–Universitätsmedizin Berlin, Berlin, Germany, ⁴Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, ⁵Department of Developmental Psychology Justus Liebig University Giessen, Giessen, Germany and ⁶McMaster University, Hamilton, Ontario, Canada

Corresponding author: Antonia Götz; Email: a.goetz@westernsydney.edu.au

(Received 18 April 2023; revised 16 December 2023; accepted 01 April 2024)

Abstract

We examined the neurophysiological underpinnings of lexical-tone and vowel-quality perception in learners of a non-tonal language. We tested 25 6- and 25 9-month-old German-learning infants, as well as 24 German adults and expected developmental differences for the two linguistic properties, as they are both carried by vowels, but have a different status in German. In adults, both lexical-tone and vowel-quality contrasts elicited mismatch negativities, with a stronger response to the vowel-quality contrast. Six-month-olds showed positive mismatch responses for lexical-tone and vowel-quality contrasts, with an emerging negative mismatch response for vowel-quality only. The negative mismatch responses became more pronounced for the vowel-quality contrast at 9 months, while the lexical-tone contrast elicited mainly positive mismatch responses. Our data reveal differential developmental changes in the processing of vowel properties that differ in their lexical relevance in the ambient language.

Keywords: perceptual reorganisation; non-native lexical tone perception; vowel perception; MMN; infants

Introduction

Infants' first year of life is characterised by extensive developmental changes, one major change being that infants' perceptual system attunes to the sound properties of their native language (for a review, see Werker & Gervain, 2013). Consequently, infants' sensitivity in their discrimination of phonologically and lexically relevant sound contrasts increases, while their discrimination of sound contrasts not relevant for the native linguistic system often decreases (e.g., Kuhl et al., 2006). This change from universal to

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

language-specific speech perception is referred to as perceptual reorganisation, which has been reported for consonants (e.g., Rivera-Gaxiola *et al.*, 2005; Werker & Tees, 1984), vowels (e.g., Polka & Bohn, 2011; Tsuji & Cristia, 2014), word stress (e.g., Bijeljac-Babic *et al.*, 2012; Höhle *et al.*, 2009), and for lexical tones (e.g., Götz *et al.*, 2018; Liu & Kager, 2014; Mattock *et al.*, 2008; Yeung *et al.*, 2013). However, developmental changes in infants' perception are not only related to the phonological system of infants' native language but also acoustic properties, such that for perceptually salient sound contrasts discriminability can be maintained throughout development (e.g., Chládková & Paillereau, 2020; Narayan, 2019, 2020). The interplay of lexically relevant and less relevant properties in speech-sound perception is, however, not yet fully understood. Our study aims to provide insights into this interplay by investigating the neural underpinnings of developmental changes in the perception of vowel quality and lexical tone in German-learning 6- and 9-month-olds and comparing infant processing with adult native German speakers. In our study, vowel quality refers to changes in vowel height. It is a sound property that is lexically relevant in German, while lexical tone is not – yet the same speech segment carries the acoustic properties that determine vowel quality and lexical tone. Although vowel length also holds lexical relevance in German, it potentially interacts with lexical tone contrasts and was thus not chosen as a contrast of interest. For example, syllables with a higher fundamental frequency (f_0) may be perceived as longer than those with a lower f_0 (e.g., Yu, 2010). We chose the broader term “vowel quality” because we believe our results would be generalisable across various aspects of vowel quality, including, but not limited to, vowel height, tongue position (front, central, or back), or lip rounding. This provides an opportunity to investigate developmental changes in the neural responses to acoustic changes within the same speech segment that are either relevant or irrelevant in the linguistic system of a given language.

Behavioural studies on the perceptual reorganisation of vowel quality and lexical tones

Previous research has examined infants' perception of vowels across various languages. A meta-analysis by Tsuji and Cristia (2014) on 22 studies showed that between 6 and 10 months of age, the effect sizes for native and non-native vowel discrimination begin to diverge with an increasing discrimination performance for native vowels. However, no decline in non-native vowel discrimination was found in the meta-analysis nor in two additional studies not covered in the analysis (de Klerk *et al.*, 2019; Mazuka *et al.*, 2014). Thus, perceptual attunement for vowels seems to be characterised by enhanced perceptual sensitivity to native vowel differences, with no clear evidence of a decline in non-native vowel discrimination.

Tonal languages, such as Mandarin, use pitch variations called lexical tones, mainly carried by the vocalic segments, to differentiate word meaning. Research on infants' perception of lexical tones presents mixed findings. For instance, Mandarin-learning infants demonstrated improved discrimination of a native, acoustically salient tone contrast from 6 to 13 months, while no changes were observed in their discrimination abilities for less acoustically salient tone contrasts (e.g., Shi *et al.*, 2017). Mixed findings have also been reported for studies with infants learning a non-tone language: some studies reported a decline in discrimination abilities at 9 months (Götz *et al.*, 2018; Mattock *et al.*, 2008; Yeung *et al.*, 2013), while others found an increase in perceptual sensitivity between 4 and 12 months (Chen & Kager, 2016; Chen *et al.*, 2017). Other studies found no evidence of a change in perceptual sensitivity for lexical tones (Ramachers *et al.*, 2018; Shi *et al.*, 2017). A few studies have tested infants beyond the

first year of life and found a U-shaped development with a decline in tone discrimination between 6 and 9 months of age and a regain in discrimination during the second year of life (Götz et al., 2018; Liu & Kager, 2014). The discrepancies among these studies might stem from differences in the contrasts tested, infants' native languages, and experimental methods (see Götz et al., 2018, for a discussion).

The current study has two aims. First, to compare the neural underpinnings of the perception of speech contrasts that are either lexically irrelevant (lexical tones) or lexically relevant (vowel quality) in the infants' native language. Second, to contribute to the understanding of the heterogeneous picture of tone perception in non-tonal language-learning infants. We examine whether the decrease in behavioural discrimination of a lexically irrelevant lexical-tone contrast in German-learning infants aged 6 to 9 months (Götz et al., 2018) would be reflected in developmental changes in the neurophysiological responses to identical sounds.

The auditory Mismatch Response

Neurophysiological measures offer the advantage of testing speech perception independently of potential restrictions from behavioural paradigms (e.g., infants' attention) and may thus be more sensitive to capture infants' speech discrimination and developmental changes. In adults, neurophysiological measures, such as the mismatch negativity (MMN) of the auditory event-related potentials (ERPs), have been used to assess neural speech discrimination (e.g., Näätänen et al., 2007). The MMN reflects differences between ERP responses to rare deviant stimuli and frequent standard stimuli, peaking at around 100-250 ms after acoustic divergence and is most prominent at frontocentral electrode positions. Following the MMN, a Late Discriminative Negativity (LDN) can occur at approximately 300-600 ms post-acoustic divergence. The occurrence of both components suggests a two-stage sequential process. In addition to auditory discriminability, the LDN is suggested to be more associated with complex auditory stimuli reflecting higher cognitive involvement (Čeponienė et al., 2004; Yu et al., 2018).

In contrast to adult listeners, infants show a Mismatch Response (MMR) with positive (pMMR) or negative (nMMR) polarities. This response is influenced by, for example, the infants' age (e.g., Leppänen et al., 2004), sex (e.g., Mueller et al., 2012), familial risk for dyslexia (e.g., Thiede et al., 2019), type and acoustic distance of tested speech contrasts (Cheng et al., 2015; Morr et al., 2002), and data pre-processing approaches (e.g., high-pass filtering; Weber et al., 2004). Infant MMRs often have a later onset and longer duration compared to adult MMNs (e.g., Friederici et al., 2002; Garcia-Sierra et al., 2016; Marklund et al., 2019; Shafer et al., 2011). They can occur in early (i.e., 150 to 350 ms) and late time windows (i.e., 350 to 600 ms), with pMMRs and nMMRs in either window, influenced by factors such as the infants' age (Yu et al., 2019), language experience (Garcia-Sierra et al., 2016; Marklund et al., 2019), and speech stimuli category (Cheng et al., 2015). The functional underpinnings of these temporal differences are still debated, with early effects possibly reflecting acoustic stimulus processing and later effects being related to native language experience (e.g., Garcia-Sierra et al., 2016). These findings suggest that infants show neural markers of speech discrimination with polarity and latency differences compared to adults, which are related to stimulus characteristics and language experience. This makes the MMR an ideal measure to study developmental changes in the perception of lexical tone and vowel quality as lexically irrelevant and relevant features in German, respectively.

ERP Studies on the perceptual reorganisation of vowel quality and lexical tones

ERP studies with infants have reported both pMMRs and nMMRs to vowel changes. For example, Yu *et al.* (2019) found that MMR amplitude/polarity and response time-window to the native vowel contrast /ɛ/ versus /i/ were associated with infant age. In the early window (160-360 ms), all age groups (3- to 47-month-olds) showed a pMMR, which became less positive with increasing age. In the late window (400-600 ms), infants up to 12 months displayed a pMMR, while older infants exhibited an nMMR. Similarly, Marklund *et al.* (2019) reported a pMMR for the native vowel contrast /e/ versus /i/ in 4- to 8-month-old Swedish-learning infants in the early time window (150-350 ms), with no MMR in the later one (350-550 ms). In a longitudinal study, Cheng *et al.* (2015) tested Mandarin-learning infants from birth to 6 months, finding that both acoustically similar (/da/ vs. /du/) and distinct (/da/ vs. /di/) vowel contrasts elicited pMMRs in newborns. At 6 months, infants were showing a pMMR for the similar vowel contrast in the late time window (250-400 ms) and an nMMR for the distinct contrast in the early time window (150-250 ms). This suggests that the polarity and timing of infant MMRs to vowel contrasts are influenced by age, the acoustic distance between vowels, and their status in the native language.

Few studies have investigated the neural processing of lexical tones in infants. Cheng *et al.* (2013) tested Mandarin-learning infants longitudinally from birth to 6 months with native-tone contrasts with a large or a small acoustic distance. The large tone contrast elicited a pMMR both at birth and at 6 months in the late time window (300-400 ms), with an additional nMMR in an early time window (150-250 ms) at 6 months only. In contrast, the small tone contrast elicited no MMR at birth, but a pMMR at 6 months in the late time window. Again, as for the studies on vowel discrimination, the results on lexical-tone discrimination demonstrate the influence of age and acoustic stimulus properties on the polarity and time window of the MMR.

Even fewer studies investigated how infants learning non-tonal languages process lexical tones on the neural level. Liu *et al.* (2019) investigated a Mandarin tone-contrast in English-learning infants aged 5-6 months and 11-12 months. Their ERP results revealed pMMRs between 100-400 ms for 5- to 6-month-olds, but no MMR for the 11- to 12-month-old infants. The absence of an MMR in older infants may indicate an attenuated neural response to the non-native tone contrast – which resembles the behavioural findings (Liu & Kager, 2014). Alternatively, the absence of an MMR in the older age group may have resulted from individual differences in a potential shift from a pMMR to an adult-like MMN, resulting in overall null effects in the ERPs.

In adults, neurophysiological evidence shows that an MMN can, in principle, be evoked by non-native tone contrasts, but is influenced by several factors (Chen *et al.*, 2018; Kaan *et al.*, 2008; Politzer-Ahles *et al.*, 2016). First, non-native speakers do not show an MMN for all contrasts that evoke an MMN in native speakers (Kaan *et al.*, 2008), and if present, the response may differ in latency and amplitude from that of native speakers (Chen *et al.*, 2018). Second, high acoustic variability in the stimuli and the duration of the interstimulus interval can influence the MMN response to non-native tone contrasts (Politzer-Ahles *et al.*, 2016; Yu *et al.*, 2017). Third, an MMN to non-native tones may only emerge throughout the experiment, suggesting that non-native speakers may need additional exposure for successful discrimination (Liu *et al.*, 2018).

The current study

So far, no study has simultaneously addressed the question of how infants learning a non-tonal language perceive lexical tones within the first year of life and how this compares to

their perception of vowel quality. The present study used ERPs to investigate the neurophysiological underpinnings of developmental changes in lexical-tone and vowel-quality perception in 6- and 9-month-old infants learning German (a non-tonal language). Presenting both lexical-tone and vowel-quality changes within one paradigm allowed us to investigate acoustic properties that are lexically irrelevant (lexical tone) or relevant (vowel quality) in German. Moreover, using identical stimuli as in Götz et al. (2018), we aimed to examine whether neurophysiological measures would reveal similar results as the behavioural measurements that revealed a developmental decline in German-learning infants' sensitivity to lexical tones.

We used a double-deviant oddball paradigm testing German adults' and German-learning infants' processing of the Cantonese mid-level (T33) versus high-rising (T25) tone and the vowel contrast /ɛ/ versus /i/. The stimuli were produced by a Cantonese speaker. The vowels differed in vowel height and corresponded to German vowel categories (see stimuli section). Deviant stimuli either differed from the standard by changing the vowel quality and keeping the tone constant (i.e., /sɛ/ as standard and /si/ as deviant, both with the mid-level tone) or by changing the tone and keeping the vowel quality constant (i.e., /sɛ/ with the mid-level tone as standard and /sɛ/ with the high-rising tone as deviant).

We hypothesize that as infants gain more exposure to sound properties of their native language, there will be corresponding changes in their neural speech discrimination. We propose two transitions for the vowel contrast to occur. Firstly, we suggest transitioning from pMMRs to nMMRs as age increases and exposure to the native language expands. Secondly, we propose a transition from an early to a late MMR as the early effects may predominantly show acoustic stimulus processing, while later effects are proposed to be associated with native language experience (e.g., Garcia-Sierra et al., 2016). Based on these developmental patterns, we expected infants at 6 months to show comparable MMRs for both types of contrasts. By 9 months, however, we expected differences in their responses to the lexical-tone and vowel-quality contrasts. If extended native-language experience affects the polarity of the discrimination effect, an initial pMMR should shift to an nMMR with increasing age, with more pronounced changes for the vowel-quality than the lexical-tone contrasts. We additionally tested adults to compare the properties of the MMNs in response to the tested lexical-tone and vowel-quality changes in the mature adult brain of native speakers of German.

Experiment 1: Neural correlates of lexical-tone and vowel-quality processing in German-speaking Adults

Methods

Participants

Twenty-four native German-speaking adults (aged 18-31 years, 14 females) were included in the final sample of this study. They reported not having learned any tone or pitch-accented language. All participants were right-handed (Edinburgh handedness inventory, Oldfield, 1971), had no self-reported hearing deficits, had normal or corrected-to-normal vision, and reported no history of neurological or psychological disorders. All participants received course credits as compensation for participating in the experiment and were recruited from the local student participant pool. Five additional participants were tested but excluded from the final data analysis because they contributed less than 50 artifact-free trials per condition. Each participant provided written informed consent according to the Declaration of Helsinki.

Stimuli

Several exemplars of the syllables /sɛ/ and /si/ were recorded with either the high-rising (T25) or the mid-level (T33) tone by a female native-speaker of Cantonese in a sound-attenuated booth. All recordings were digitalised with a sampling rate of 44.1 kHz. We selected two different tokens for each syllable. The syllable duration was similar across the different tokens (578–586 ms) and the vowels started 97–101 ms after stimulus onset. The results of acoustic analyses are given in Table S1 and the pitch contours of the stimuli are displayed in Figure S1 in the supporting information. All stimuli were normalised in intensity (using the scale intensity function in Praat – Boersma & Weenink, 2016) at 60 dB SPL. The syllable /sɛ33/ (i.e., /sɛ/ with a T33 tone contour) was always presented as standard, the syllable /sɛ25/ (i.e., /sɛ/ with a T25 tone contour) as tone deviant and /si33/ (i.e., /si/ with a T33 tone contour) as vowel deviant. To verify the assimilation to native vowel categories, we performed a perceptual test with 15 German-speaking adults who did not participate in the EEG study. The test consisted of a categorisation task and a perceptual goodness rating. In the categorisation task, two tokens of each syllable (/si33/, /si25/, /sɛ33/ and /sɛ25/) were presented three times (24 trials) and participants were asked to assign the vowels to one of the German categories: /i:/, /i/, /ɛ:/, /e:/. The vowel categories were displayed on the screen with the grapheme equivalents of <ie>, <i>, <äh>, and <eh>. Following the categorisation task, participants were asked to indicate the goodness of the category fit (1 = Poor and 7 = Perfect). The participants categorised the /ɛ/ vowel to 100% to the <äh> (the German /ɛ:/ vowel) and /i/ to 100% to the <i> (the German /i:/ vowel). Both vowels did not statistically differ in their category goodness fit (/ɛ:/ mean = 6.00, SD = 0.79; /i:/ mean = 6.18, SD = 0.82).

Procedure

Participants were seated approximately 1.5 m from a computer screen and listened to the auditory stimuli via earphones (E-A-RTONE 3A Insert Earphones, Aearo Technologies Auditory Systems). During the stimulus presentation (Presentation Software, Version 18.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com), the participants watched a silent movie (Baby Einstein). Stimuli were presented with varying interstimulus intervals (ISI) ranging between 800–900 ms in steps of 50 ms to prevent participants from perceiving regular rhythmic patterns. Overall, the experiment contained 800 stimuli: 640 standards, 80 lexical-tone deviants and 80 vowel-quality deviants, which were distributed to four blocks of 200 stimuli each. Each block started with eight standards. Deviants were presented pseudo-randomly with 3 to 8 standards presented between two deviants. The first eight standards and standards directly following deviants were excluded from further analysis. Five participant datasets were removed from analysis due to less than 50 artifact-free deviants per condition, as predetermined. Adults had an average of 413 (*SD* = 42.3) artifact-free trials for standards, 71 (*SD* = 7.6) for vowel deviants and 71 (*SD* = 7.1) for tone deviants.

ERP Recording and Analysis

The EEG was continuously recorded from 30 cap-mounted active Ag/AgCl electrodes (Brain Products, Gilching, Germany) at a sampling rate of 1000 Hz. Electrodes (F3, F7, F9, F4, F8, F10, FC1, FC5, C3, FC2, FC6, C4, CP1, CP5, P3, P7, CP2, CP6, P4, P8, FCz, Fz, Cz, CPz, Pz, O1, O2) were positioned following the 10–20 system convention. The

electrooculogram was recorded from electrodes placed below and above the right eye. Impedances were kept below 25 k Ω . The ground electrode was placed at the FP1 position. The EEG data were analysed using Brain Vision Analyzer (version 2.01; Brain Products, Gilching, Germany). The EEG recording was referenced online to the left mastoid and re-referenced offline to the linked mastoids. The signal was filtered with a 0.5-30 Hz bandpass filter (zero-phase IIR Butterworth filters of order 2, -12, dB/oct roll-off). Data were segmented in epochs of 1000 ms and baseline-corrected 100 ms before stimulus onset. Eye blinks and eye movements in the segments were corrected by an algorithm (Gratton et al., 1983). All other artifacts were detected automatically (exceeding $\pm 100 \mu\text{V}$) and excluded from further analysis.

The MMN is expected to occur in the time window of 100-250 ms after the point of acoustic divergence between stimuli (Näätänen et al., 2007). For the vowel-quality contrast, this time window was 200-350 ms after syllable onset, and for the lexical-tone contrast 400-550 ms after the syllable onset (due to the earlier point of divergences for the vowel-quality compared to the tone contrast, see Table S1 and Figure S1 in the supporting information). The LDN was analysed in a later time window of 350-600 ms after the point of divergence, which was at 450-700 ms for the vowel-quality contrast and at 650-900 ms for the lexical-tone contrast. While the MMN is typically most prominent at frontocentral regions, our objective was to incorporate a broad range of data points to enhance our comprehension of the developmental processes related to lexical tone and vowel-quality processing in infants. This approach also considers the possibility of observing an inverse negative response at posterior sites (see Peter et al., 2016). Hence, electrodes were clustered into two regions: frontocentral (F3, F4, F7, F8, FC5, FC6, Fz, C3, C4, Cz) and posterior (O1, O2, P3, P4, P7, P8, Pz, CP5, CP6, CPz).

Data analysis

We calculated the MMNs and LDNs for the lexical-tone and vowel-quality contrast by subtracting the ERP amplitude of the standard from the ERP amplitude of the tone or vowel deviant, respectively. For statistical analysis, we used the average amplitude of the difference wave (deviant minus standard) in the respective time window as the dependent variable. We conducted the analyses with the statistical software R, version 4.0.4 (R Core Team, 2021) and the lme4 package (Bates et al., 2015). Plots were created using ggplot2 (Wickham, 2016) and post-hoc tests were performed using emmeans (Lenth et al., 2018). P-values were calculated with lmerTest (Kuznetsova et al., 2017), which uses Satterthwaite approximations to degrees of freedom.

We computed separate linear mixed effects models for the MMN and LDN. The models included the effect of Deviant type (lexical tone vs. vowel quality, coded as +0.5 and -0.5) and Region (frontocentral vs. posterior, coded as +0.5 and -0.5) and their interaction. The models included a random intercept by Subject and a random by-subject slope of Deviant type. Region was not included as random slope as this led to singular model fits, showing overfitting. All contrast-codings were performed by using the general inverse (Schad et al., 2020). For the first step, the intercept was set to zero (*lmer(amplitude~-1+Deviant type*Region+ (1+Deviant type|Subject)*) to compare the difference wave against zero and verify an MMN and an LDN. For the second step, the grand-mean data were used as intercept (*lmer(amplitude~Deviant type*Region+ (1+Deviant type|Subject)*). For both MMN and LDN results, we first present the results of the difference wave comparison against zero, and then the model output, with the comparison of lexical-tone and vowel-quality processing.

Results

Figures 1 and 2 depict the grand-average ERPs for standards, lexical-tone contrasts, vowel-quality contrasts, and the corresponding ERP difference wave obtained from frontocentral and posterior electrode regions.

Mismatch Negativity (MMN)

The results for the MMN time-window revealed that both lexical-tone and vowel-quality changes elicited MMNs at frontocentral and posterior regions, as the respective difference between standards and deviants differed significantly from zero (lexical tones: frontocentral β (SE) = -0.852 (0.281), t = -3.030, p = 0.011, posterior β (SE) = -0.648 (0.236),

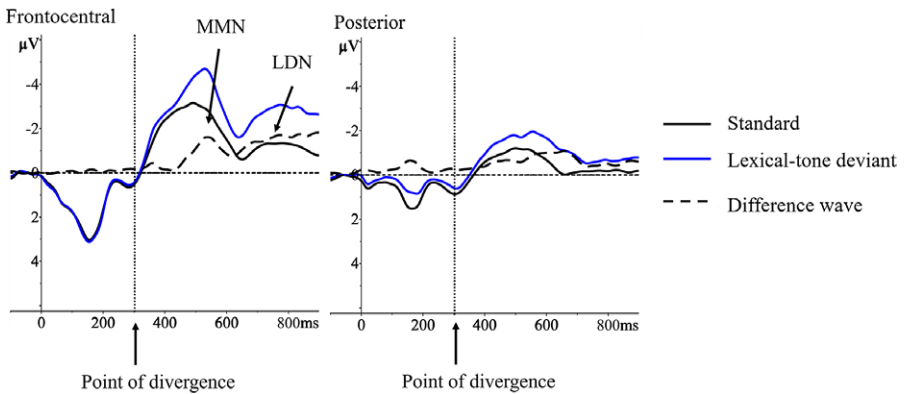


Figure 1. Grand-average ERPs of the lexical-tone contrast for German adults
 Note. Grand-average ERPs for German-speaking adults for standards (black line), lexical-tone deviants (blue line), and the corresponding difference waves (deviant–standard, dashed line) at frontocentral and posterior electrode regions.

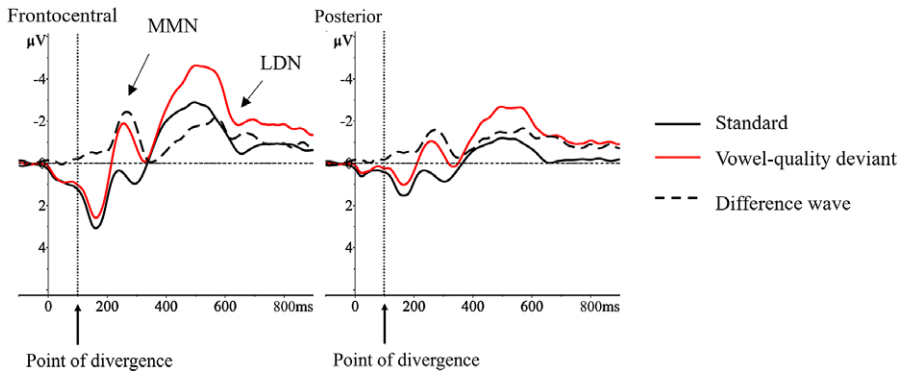


Figure 2. Grand-average ERPs of the vowel-quality contrast for German adults
 Note. Grand-average ERPs for German-speaking adults for standards (black line), vowel-quality deviants (red line), and the corresponding difference waves (deviant–standard, dashed line) at frontocentral and posterior electrode regions.

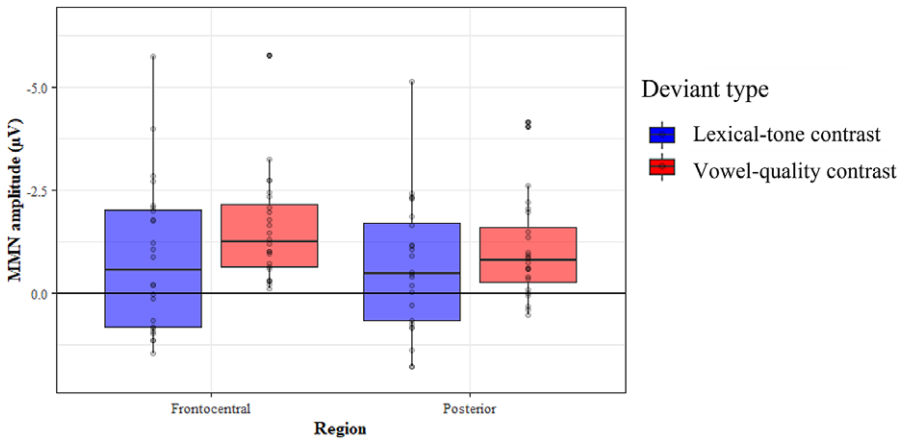


Figure 3. MMN amplitude in German-speaking adults

Note. MMN amplitude (difference wave) in German-speaking adults for lexical-tone and vowel-quality changes at frontocentral and posterior electrode regions. Each dot represents one participant. The horizontal lines within boxes show medians, boxes show the interquartile (IQ) range, whiskers the 1.5 * IQ range, dots are potential outliers.

$t = -2.748, p = 0.021$; vowel quality: frontocentral β (SE) = -1.539 (0.281), $t = -5.472, p < 0.001$, posterior β (SE) = -1.065 (0.236), $t = -4.515, p < 0.001$; all p -values are Bonferroni-corrected for multiple comparisons).

The full output of the linear mixed model with the MMN difference-waves of both lexical-tone and vowel-quality contrasts is given in the supporting information S2. The results revealed an effect of Deviant type (β (SE) = 0.276 (0.057), $t = 4.874, p < 0.001$) which indicates that the MMN differs between lexical-tone and vowel-quality changes (independent of electrode region), with a more pronounced MMN amplitude for vowels than tones. In addition, the effect of region shows that the MMN is generally larger at frontocentral than posterior regions (β (SE) = -0.169 (0.057), $t = -2.993, p < 0.01$), see Figure 3.

Late Discriminate Negativity (LDN)

The results for the LDN time-window revealed that both lexical-tone and vowel-quality deviants elicited a significant LDN, as the difference between standards and deviants differed significantly from zero (lexical tones: frontocentral β (SE) = -1.522 (0.354), $t = -4.301, p < 0.01$, posterior β (SE) = -0.687 (0.244), $t = -2.811, p = 0.018$; vowel quality: frontocentral β (SE) = -1.833 (0.354), $t = -5.208, p < 0.001$, posterior β (SE) = -1.460 (0.244), $t = -5.977, p < 0.001$; all p -values are Bonferroni-corrected for multiple comparisons).

The full output of the linear mixed model with the LDN difference-waves of both lexical-tone and vowel-quality contrasts is given in the supporting information S3. The effect of Region shows that the LDN is overall larger at frontocentral than posterior regions (β (SE) = -0.169 (0.057), $t = -2.993, p < 0.01$), see Figure 4. The interaction of Deviant type and Region reveals that the difference between lexical tones and vowel quality is larger at the posterior than the frontocentral region (β (SE) = -0.113 (0.056), $t = -2.019, p = 0.044$), as the LDN had a frontocentral focus in response to lexical-tone changes, but was similarly pronounced across regions for vowel-quality changes.

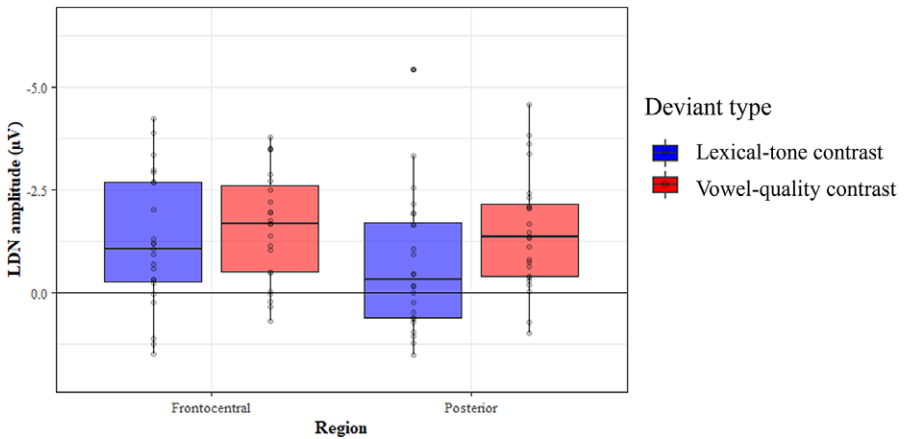


Figure 4. LDN amplitude in German-speaking adults

Note. LDN amplitude (difference wave) in German-speaking adults for lexical-tone and vowel-quality changes at frontocentral and posterior electrode regions. Each dot represents one participant. The horizontal lines within boxes show medians, boxes show the interquartile (IQ) range, whiskers the 1.5 * IQ range, dots are potential outliers.

Discussion

Experiment 1 yielded two main findings for German adults. First, both lexical-tone and vowel-quality deviants elicited MMNs and LDNs at frontocentral and posterior regions. Second, the MMN was generally more pronounced for vowel quality than lexical tone and the LDN was also more pronounced for vowel quality than lexical tone but only at the posterior region not at the frontocentral region. Thus, German-speaking adults showed neural responses indicating discrimination between the high-rising and mid-level lexical tones and the vowel-quality change from / ϵ / to / i /. Our MMN findings of adults' lexical-tone discrimination are in line with several behavioural and neurophysiological studies that found evidence for tone discrimination in adult speakers of non-tonal languages (Chen *et al.*, 2018; Kaan *et al.*, 2008; Politzer-Ahles *et al.*, 2016). Regarding vowel discrimination, our MMN results confirm the hypothesis that the tested changes in vowel quality, being acoustically close to participants' native language, should evoke stronger neural discrimination responses than the lexical-tone changes (see Yu *et al.*, 2018). Adult participants additionally showed an LDN following the MMN for both lexical-tone and vowel-quality contrasts. The functional significance of the LDN component has been controversially discussed (e.g., Čeponienė *et al.*, 2004), and might reflect high-order cognitive processing, such as integrating the speech stimuli into the native phonology (e.g., Barry *et al.*, 2009), reorienting of attention (Mueller *et al.*, 2008), or representations of long-term memory in deviancy detection (Zachau *et al.*, 2005). Moreover, the emergence of both components (MMN and LDN) suggests a two-stage sequential process in which the occurrence of both components suggests a higher need for resources to encode the speech signal (Yu *et al.*, 2018).

Our finding that vowel-quality changes elicited a stronger MMN and LDN than lexical-tone changes likely stems from the different relation of these sound classes to the participants' native phonological system. Given that German adults had assigned both vowels to German vowel categories, with no difference in category goodness fit, we suggest that listeners likely mapped the Cantonese vowels onto the German vowel system

and that the stronger neural responses to the vowels (compared to the tones) reflect their phonological processing. This explanation is in line with other studies showing stronger ERP responses to native compared to non-native sounds (e.g., Rivera-Gaxiola et al., 2005; Yu et al., 2018). However, acoustic differences in the realisation of the lexical-tone and the vowel-quality contrasts could also have induced differences in the responses. Specifically, the acoustic difference between the vowels /ɛ/ and /i/ is characterised by a fast formant transition while the differences in the pitch contour between the tones develop slowly over the length of the vowel. This faster acoustic transition from one vowel to the other vowel might have led to greater acoustic salience of the contrast in comparison to the slower trajectory of pitch in lexical tones. However, Yu et al. (2022) have demonstrated that native tone language speakers exhibit similar neural responses to vowel-quality and lexical-tone contrasts. These results may speak against a pure acoustic explanation for the difference in the German speakers' responses to the vowel-quality and lexical-tone contrasts, but support our assumption that these differences reflect the status of the contrast in the phonological system. However, our results confirm that German-speaking adults show a neural response to both types of contrasts – hence, we used the stimuli to investigate neural discrimination in German-learning infants.

Experiment 2: Neural correlates of lexical-tone and vowel-quality processing in German-learning 6- and 9-month-old Infants

Methods

Participants

The final participant sample included data from 50 German-learning infants: 25 6-month-olds ($M_{age} = 185$ days, $range = 165$ to 210 days, 12 females) and 25 9-month-olds ($M_{age} = 274$, $range = 262$ to 294 days, 11 females). Infants who had participated in the study at 6 months, did not participate at 9 months. An additional 36 infants were tested, but their data were excluded from further analysis for the following reasons: less than 35 artifact-free trials ($n = 14$ 6-month-olds, $n = 20$ 9-month-olds) or non-compliance ($n = 1$ 6-month-old, $n = 1$ 9-month-old). The infants were recruited from the participant pool of the local BabyLAB. According to the parental report, all infants were born full-term, did not suffer from repeated or acute ear infections, showed no indications of atypical development, and did not have exposure to a tone or pitch-accent language. Data collection took place from April 2018 to October 2019. This study was approved by the Ethics Committee of the local University. Parents gave written informed consent following the Declaration of Helsinki.

Stimuli and procedure

The stimuli and the double-deviant oddball task were the same as in Experiment 1. Before the experiment started, caretakers were informed about the procedure and signed or handed in the signed consent form. Infants were seated on their caretakers' lap approximately 1 m away from a computer screen. The acoustic stimuli were presented by two loudspeakers to the left and right of the screen. During stimulus presentation, infants watched a silent infant-friendly movie (Baby Einstein), or a second experimenter engaged the infant with silent toys. The experiment was terminated if the infant became fussy or if the maximum presentation time (20 min) was reached. A short break of five minutes was inserted after each 200 trials. Participants were excluded from the analysis if they

contributed fewer than 35 artifact-free trials per deviant condition (a priori criterion). On average, 6-month-old infants had 218 ($SD = 54$) artifact-free trials for standards, 37 ($SD = 9.1$) for lexical-tone deviants, and 38 ($SD = 9.6$) for vowel-quality deviants. The 9-month-olds had an average of 281 ($SD = 47.9$) artifact-free trials for standards, 49 ($SD = 8.4$) for lexical-tone deviants, and 49 ($SD = 8.6$) for vowel-quality deviants.

ERP recording and analysis

The EEG was continuously recorded from 32 cap-mounted active Ag/AgCl electrodes (Brain Products, Gilching, Germany) at a sampling rate of 1000 Hz. Electrodes (F3, F7, F9, F4, F8, F10, FC1, FC5, C3, FC2, FC6, C4, CP1, CP5, P3, P7, CP2, CP6, P4, P8, FCz, Fz, Cz, CPz, Pz, O1, O2) were positioned following the 10–20 system. The electro-oculogram was recorded from electrodes placed above the right and left eye. The ground electrode was placed at the AFz position. Impedances were kept below 25 k Ω . The procedure for EEG data preprocessing and artifact rejection was the same as in Experiment 1. In line with developmental MMR studies (e.g., Chládková *et al.*, 2021; Marklund *et al.*, 2019; Yu *et al.*, 2019), we analysed the EEG data within two a priori selected time windows, confirmed by visual inspection of the difference wave. We analysed the MMR responses in an early window from 150–350 ms (henceforth early MMR), and a later time window from 350–550 ms (henceforth late MMR) after the point of acoustic divergence of the stimuli. For the vowel-quality contrast, the early MMR was at 250–450 ms and the late MMR was at 450–650 ms after stimulus onset. For the lexical-tone contrast, the early MMR was at 450–650 ms, and the late MMR was at 650–850 ms after stimulus onset. Electrodes were clustered into the same frontocentral and posterior regions as in Experiment 1.

Data analysis

As in Experiment 1, we calculated the difference waves of the early and late MMRs and used the values of the amplitude of the difference wave as the dependent variable; the same statistical software and packages were used as in Experiment 1.

We computed separate linear mixed effect models for the early and late MMRs. The same contrast codings as in Experiment 1 were applied with the additional factor of Age (6 months vs. 9 months, coded as +0.5 and -0.5). The models included a random intercept by Subject and a random by-subject slope of Deviant type. Region was not included as random slope as this led to singular model fits, showing overfitting. As a first step, the intercept was set to zero to compare the difference wave against zero to verify the MMRs. For the second step, we fitted the model in such a way that the intercept reflected the grand mean data of the predictors. For both the early and late MMR results, we first present the results of the difference-wave comparison against zero, and then the output of the model with the comparison of lexical-tone versus vowel-quality changes in the age groups of 6 and 9 months.

Results

Figures 5 and 6 depict the grand-average ERPs for standards, lexical-tone deviants, vowel-quality deviants, and the corresponding difference waves at frontocentral and posterior regions at 6 and 9 months.

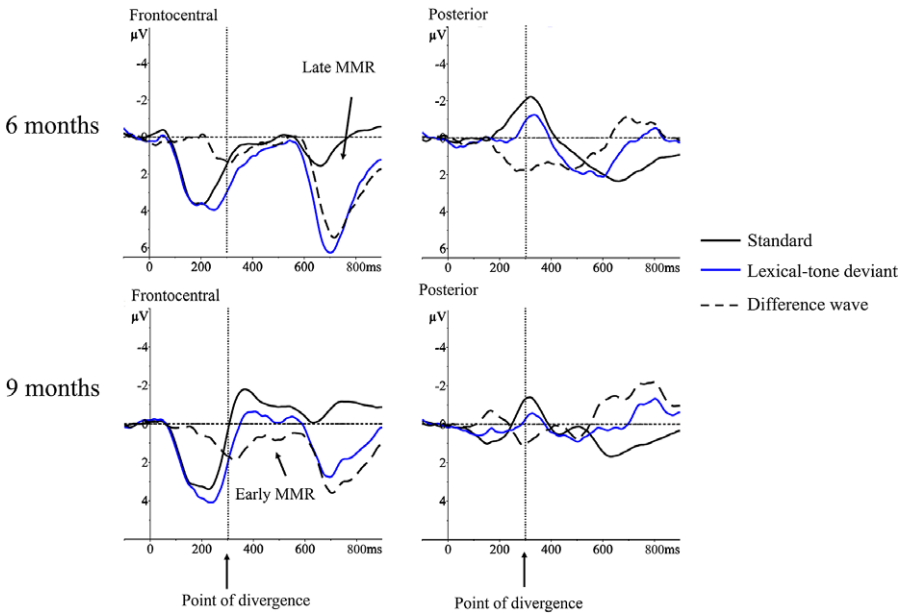


Figure 5. Grand-average ERPs of the lexical-tone contrast in German-learning infants
 Note. Grand-average ERPs at 6 and 9 months for standards (black line), lexical-tone deviants (blue line), and the corresponding difference waves (deviant–standard, dashed line) at frontocentral and posterior regions.

Mismatch Response in the early time window

The results of comparing the ERP difference waves against zero (see Figure 7 and supporting information Table S4, see also Table 1, all *p*-values Bonferroni-corrected) revealed that the lexical-tone deviant elicited a significant pMMR at the frontocentral region (β (SE) = 0.8899 (0.321), $t = 2.775$, $p = 0.032$) but not at the posterior region (β (SE) = -0.039 (0.413), $t = -0.094$, $p = 1$) in 9-month-olds. In contrast, there were no significant effects of the lexical tone contrast at 6 months (frontocentral: β (SE) = 0.312 (0.447), $t = 0.698$, $p = 1$; posterior: β (SE) = 0.875 (0.440), $t = 1.987$, $p = 0.23$). The vowel-quality deviant elicited a significant pMMR at the frontocentral region (β (SE) = 2.349 (0.440), $t = 5.261$, $p < 0.001$) but not at the posterior region (β (SE) = -0.041 (0.440), $t = -0.094$, $p = 1$) at 6 months while there were no significant effects at 9 months (frontocentral: β (SE) = 0.268 (0.321), $t = 0.837$, $p = 1$; posterior: β (SE) = -0.825 (0.417), $t = 1.978$, $p = 0.21$).

The linear mixed model for the early time window (see supporting information Table S5), with the grand mean of ERP difference waves as intercept, revealed a significant three-way interaction between Deviant type (lexical tone vs. vowel quality), Age (6 vs. 9 months), and Region (frontocentral vs. posterior) (β (SE) = 0.349 (0.105), $t = 3.33$, $p < 0.001$). Post-hoc analyses showed no significant changes in neural responses between 6 to 9 months at either region for the lexical-tone contrast (frontocentral: β (SE) = 0.578 (0.654), $t = 0.884$, $p = 0.813$; posterior: β (SE) = -0.914 (0.654), $t = -1.396$, $p = 0.505$) – despite the test against zero showing an early frontocentral pMMR at 9 months but not at 6 months (Figure 7 and Table S4). In contrast, vowel-quality changes evoked significant differences between 6 and 9 months at the frontocentral region (β (SE) = 2.081 (0.664), $t = 3.230$, $p = 0.009$), but not at the posterior region (β (SE) = 0.783 (0.664), $t = 1.216$,

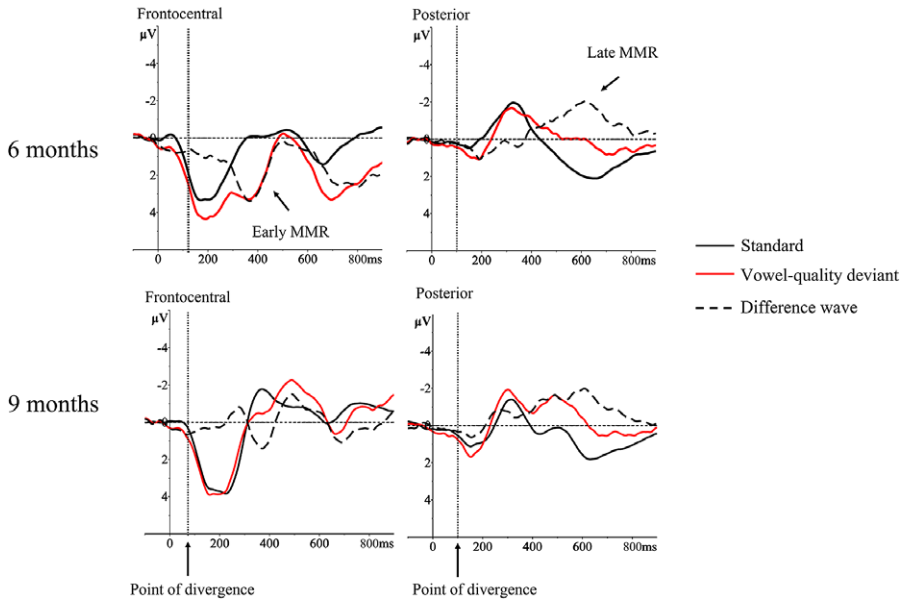


Figure 6. Grand-average ERPs of the vowel-quality contrast in German-learning infants
 Note. Grand-average ERPs at 6 and 9 months for standards (black line), vowel-quality deviants (red line), and the corresponding difference waves (deviant–standard, dashed line) at frontocentral and posterior regions.

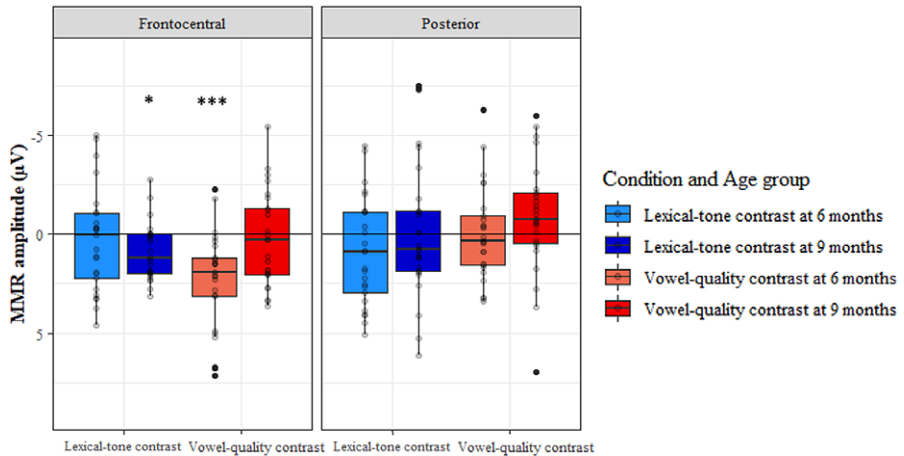


Figure 7. Distribution of the infant MMRs in the early time window
 Note. Distribution of the MMRs (difference waves) in the early time window for lexical tone (blue) and vowel quality (red) at 6 months (light colour) and 9 months (dark colour) at frontocentral and posterior regions. Each dot represents one participant. The horizontal lines within boxes show medians, boxes show the interquartile (IQ) range, whiskers the 1.5 * IQ range, dots are potential outliers. An MMR being significantly different from zero is indicated by * ($p < 0.05$), ** ($p < 0.01$), or *** ($p < 0.001$).

Table 1. MMRs in 6- and 9-month-olds

Time Window	Region	Lexical tone		Vowel quality	
		6 months	9 months	6 months	9 months
Early	Frontocentral	–	<i>pos</i>	<i>pos</i>	–
	Posterior	–	–	–	–
Late	Frontocentral	<i>pos</i>	<i>pos</i>	–	<i>neg</i>
	Posterior	–	<i>neg</i>	<i>neg</i>	<i>neg</i>

Note. MMRs in 6- and 9-month-olds for the lexical-tone and vowel-quality contrasts at frontocentral and posterior regions in the early and late time windows. Pos. refers to a statistically significant pMMR, whereas neg. refers to a statistically significant nMMR (from statistical tests against zero). Italics indicate effects that were not replicated in post-hoc analyses of the linear mixed model with direct comparisons across age and sound contrasts.

$p = 0.619$). These results coincide with the test against zero, which showed a frontocentral pMMR at 6 months but not at 9 months (see Figure 7). At 6 months, neural responses to lexical-tone and vowel-quality changes were significantly different, with the responses to vowel-quality changes being more positive than to lexical-tone changes at the frontocentral region, (β (SE) = 2.038 (0.668), $t = 3.052$, $p = 0.016$). This is supported by the test against zero showing a frontocentral pMMR for vowel quality only (see Figure 7). However, we observed no significant difference between lexical-tone and vowel-quality processing at 9 months at either region (frontocentral: β (SE) = 0.621 (0.668), $t = 0.931$, $p = 0.789$; posterior: β (SE) = - 0.786 (0.668), $t = 1.177$, $p = 0.643$) – although the test against zero revealed a frontocentral pMMR for the lexical tone (see Figure 7) but not for the vowel-quality contrast.

Mismatch Response in the late time window

Comparing the ERP difference waves against zero (see Figure 8 and supporting information Table S6, all p -values Bonferroni-corrected) for the late time window revealed that the lexical-tone deviant elicited a pMMR at the frontocentral (β (SE) = 4.408 (0.520), $t = 6.551$, $p < 0.001$, but not at the posterior region: β (SE) = -0.876 (0.534), $t = -1.625$, $p = 0.452$) at 6 months. At 9 months, we found a pMMR at the frontocentral (β (SE) = 2.497 (0.312), $t = 7.997$, $p < 0.001$) as well as an nMMR at the posterior region (β (SE) = -1.669 (0.444), $t = -3.758$, $p = 0.001$). The vowel-quality deviant elicited an nMMR at the posterior (β (SE) = -1.511 (0.539), $t = -2.804$, $p = 0.033$), but not at the frontocentral region (β (SE) = 0.518 (0.520), $t = 0.989$, $p = 1$) at 6 months and nMMRs at frontocentral (β (SE) = -1.415 (0.372), $t = -3.798$, $p = 0.002$) and posterior regions (β (SE) = -2.076 (0.444), $t = -4.675$, $p < 0.001$) at 9 months. The linear mixed model for the late analysis time-window (full model output in Table S7), with the grand mean of the ERP difference waves as intercept, revealed a significant three-way interaction between Deviant type, Age, and Region (β (SE) = 0.233 (0.112), $t = 2.072$, $p = 0.39$).

Post-hoc analyses revealed no significant change from 6 to 9 months for the lexical-tone contrast (frontocentral: β (SE) = -0.911 (0.670), $t = -1.359$, $p = 0.528$; posterior: β (SE) = -1.408 (0.670), $t = -2.102$, $p = 0.160$). This result showed that both age groups exhibit a frontocentral pMMR (Figure 8), whereas the posterior nMMR for the 9-month-olds, also observed in the test against zero, was not confirmed in the age comparison. Vowel-quality changes, however, elicited a more negative MMR at 9 than 6 months at the frontocentral

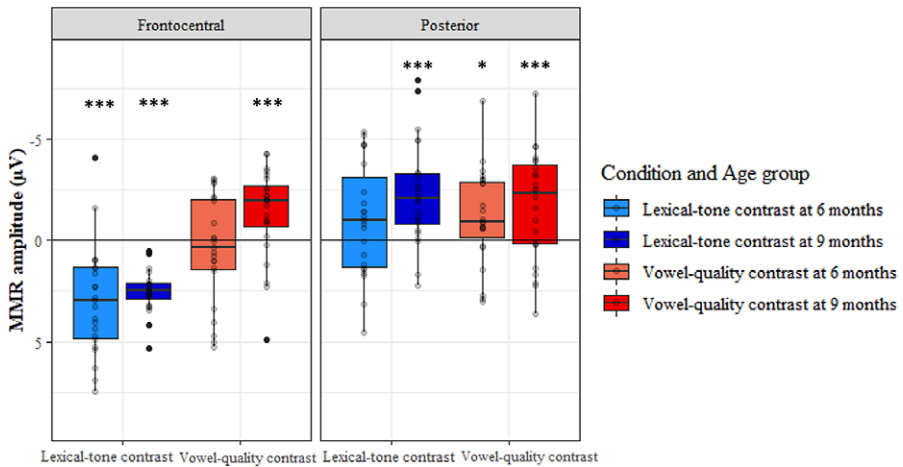


Figure 8. Distribution of the infant MMRs in the late time window

Note. Distribution of the MMRs (difference waves) in the late time window for lexical tone (blue) and vowel quality (red) at 6 months (light colour) and 9 months (dark colour) at frontocentral and posterior regions. Each dot represents one participant. The horizontal lines within boxes show medians, boxes show the interquartile (IQ) range, whiskers the 1.5 * IQ range, dots are potential outliers. A significantly different MMR from zero is indicated by * ($p < 0.05$), ** ($p < 0.01$), or *** ($p < 0.001$).

region (β (SE) = -1.929 (0.709), $t = -2.722$, $p = 0.039$) but not at the posterior region (β (SE) = -0.565 (0.709), $t = -0.798$, $p = 0.855$) – showing that only 9-month-olds exhibit a frontocentral nMMR, while both age groups alike showed a posterior nMMR (Figure 8). Moreover, we found significant differences between lexical-tone and vowel-quality contrasts in 6-month-olds at the frontocentral region (β (SE) = 2.893 (0.654), $t = 4.425$, $p < 0.001$), but not at the posterior region (β (SE) = 0.635 (0.654), $t = 0.972$, $p = 0.766$). These results are supported by the test against zero only showing a pMMR for the lexical-tone but no MMR for the vowel-quality contrast in this age group. In 9-month-olds, we found a significant difference between lexical-tone and vowel-quality responses at the frontocentral region (β (SE) = -3.912 (0.654), $t = -5.983$, $p < 0.001$), as infants showed a pMMR for lexical tones and an nMMR for vowel quality (see Figure 8). We found no difference in neural responses to the deviant types at the posterior region (β (SE) = -0.208 (0.654), $t = -0.318$, $p = 0.989$), as the 9-month-olds showed similar nMMRs for both sound contrasts (Figure 8).

Discussion

Experiment 2 investigated the neural correlates of lexical-tone and vowel-quality processing in 6- and 9-month-olds learning a non-tonal language. Our results (see Table 1) revealed first that the lexical-tone contrast only elicited a late pMMR at the frontocentral region in the 6-month-olds, while the 9-month-olds showed early and late pMMRs at the frontocentral and a late nMMR at the posterior region. Second, the vowel-quality contrast elicited an early pMMR at the frontocentral and a late nMMR at the posterior region in 6-month-olds. At 9 months, infants showed late nMMRs at both frontocentral and posterior regions for the vowel-quality contrast.

Thus, our data show that both types of sound contrasts – lexical-tone and vowel-quality contrasts – elicited MMRs indicating successful sound discrimination in 6- and 9-month-old infants despite learning a non-tonal language. However, we found differences in the polarity and the temporal course of these neural responses that will be discussed in detail below.

General discussion

The current study focused on the neural underpinnings of perceptual reorganisation in infancy with two research questions: 1) are there distinct neurophysiological responses to lexically relevant versus irrelevant acoustic changes even if the changes are carried by the same speech segment; and 2) are developmental changes in lexical-tone discrimination, reported in a behavioural study with German infants (Götz et al., 2018) manifested in respective neurophysiological responses? Using the same lexical-tone contrast as in Götz et al. (2018), we compared the neurophysiological responses to the lexical-tone contrast and a vowel-quality contrast in infants at 6 and 9 months of age and in adults. We hypothesised a transition from positive to negative MMRs with increasing age, being more pronounced for the vowel-quality than the lexical-tone contrast. Moreover, we expected the MMRs for the lexical-tone and vowel-quality contrasts to be similar for the 6-month-olds and, due to longer native-language exposure, different for the 9-month-olds.

Modulation of Mismatch Responses by age and sound contrasts

Learners of a non-tonal language perceived changes in vowel properties for both the lexically irrelevant lexical-tone and the lexically relevant vowel-quality contrast – yet with different neural responses. In the following, we will discuss developmental differences by evaluating the respective characteristics of the neural discrimination response.

Age

Data from adults indicating discrimination responses of the mature brain revealed for both speech contrasts negative responses (i.e., MMNs and LDNs). Infants, however, showed a combination of pMMRs and nMMRs at 6 and 9 months for both speech contrasts (see Table 1): the age differences for infants were statistically significant for the vowel-quality contrast, with a polarity shift from an early pMMR at 6 months to a late nMMR at 9 months. For the lexical-tone contrast, infants showed a late pMMR at both ages with an additional late nMMR for 9-month-old infants, yet only statistically confirmed in the test against zero and not when comparing the age groups directly. These patterns regarding the polarity of the discrimination responses comply with findings reporting nMMRs instead of pMMRs with increasing age, indicating an age-related maturation of the neural system (e.g., Leppänen et al., 2004; Yu et al., 2019). However, the coexistence of nMMRs and pMMRs in both infant groups provides further evidence that the polarity of the mismatch response is influenced not only by age but also by other factors like the relation of the stimuli to the listener's native language (Cheour et al., 1998). The age-related polarity shifts in the MMR, specifically for the lexically relevant vowel-quality contrast, are accompanied by a transition from an early time window at 6 months (pMMR) to a later time window at 9 months (nMMR). This

transition underscores the greater influence of native language experience asserted at 9 months than 6 months (Garcia-Sierra *et al.*, 2016). For the lexical-tone contrast, an MMR polarity change seems to emerge, as 9-month-olds showed a late nMMR. Yet, this effect could not be statistically confirmed in the age comparison, suggesting that the developmental change in MMR polarity is weaker for the lexically irrelevant lexical-tone contrast or reflects high individual differences in the timing of a developmental shift.

Sound contrast

In adults, both the lexical-tone and the vowel-quality contrast elicited MMNs and LDNs as discrimination responses, yet with stronger responses to the lexically relevant vowel-quality changes, which is in line with other studies showing stronger ERP responses to native compared to non-native speech contrasts (e.g., Näätänen *et al.*, 2007). Regarding infants' processing of tones and vowels, our study was the first one to examine potential developmental differences in the perceptual shaping from native-language experience within one experiment and realised in a single speech segment.

The occurrence of pMMRs for the vowel-quality change in 6- and the transition towards nMMRs in 9-month-olds corroborates findings from a recent longitudinal ERP study with German-learning infants (Werwach *et al.*, 2022). The authors measured neural discrimination responses to a native German vowel contrast in infants at 2, 6, and 10 months. Overall, infants in this study showed frontocentral pMMRs at all ages, yet with the largest positivity for the 6-month-olds that significantly decreased in amplitude towards 10 months. Other studies on vowel discrimination exhibit comparable results, wherein younger infants displayed pMMRs that transitioned towards nMMRs as they aged (Marklund *et al.*, 2019; Shafer *et al.*, 2011; Yu *et al.*, 2019). Regarding neurophysiological indicators of lexical-tone perception in learners of a non-tonal language, Liu *et al.* (2019) showed an early pMMR for 5- to 6-month-old English-learning infants, but no MMR for 11- to 12-month-olds. The absence of an MMR in the older age group may have resulted from infants' decreasing discrimination of a non-native contrast or reflected individual differences in the timing of a developmental shift from a pMMR to an nMMR, resulting in overall null effects. The latter interpretation is in line with our observation of 9-month-olds showing both pMMRs and an nMMR emerging for the lexical-tone contrast. Applying our experimental paradigm to infants older than 9 months, the nMMR should then be visible as the predominant discrimination response.

Interaction of age and native language in the discrimination of lexical-tone and vowel-quality contrasts: neurophysiological indications of perceptual reorganisation?

Our results confirm our hypotheses on postulated effects of perceptual reorganisation in infants' neural processing of lexical tone and vowel quality. Specifically, we observed developmental differences between 6 and 9 months, with effects of native-language influences becoming more apparent at 9 months of age. The maturation pattern for the vowel-quality contrast fits with the perceptual reorganisation, leading to efficient processing of native sound contrasts compared to non-native sound contrasts (Aslin & Pisoni, 1980; Tsuji & Cristia, 2014). The stronger neural response to vowel quality than to lexical tone in adults and the transition to a mature neural response for vowel quality in infancy can be explained by native-language experience, as German listeners are regularly exposed to vowels similar to the one used in the current study, but not to different lexical

tones. This aligns with the findings that the amount of exposure can modulate the polarity of infants' MMRs (Garcia-Sierra et al., 2016; Marklund et al., 2019; Shafer et al., 2011). Native phonological knowledge potentially contributed to the developmental shift towards more mature neural discrimination, particularly in the later time window of infancy (see also Garcia-Sierra et al., 2016). Nevertheless, we cannot rule out that the differences in neural responses to vowel-quality versus lexical-tone change may be confounded with differences in the sensory ERPs stemming from the acoustic differences between the stimuli (e.g., Scharinger et al., 2011). Acoustic distance plays a role in sound discrimination (Cheng & Lee, 2018; Cheng et al., 2013; Chládková et al., 2021; Schaadt et al., 2015), yet the acoustic properties relevant to the perceptual differences between tones and vowels are difficult to match (see Table S1 for a comparison of the just-noticeable difference between the stimuli). While vowels are distinguished by their formant frequencies, lexical tones are characterised by differences in their trajectory of f_0 within a syllable. However, evidence showing similar neural responses to vowel-quality and lexical-tone contrasts in native tonal speakers (Yu et al., 2022) disagrees with the assumption that acoustic differences are the driving force for the different neural patterns. Further research will need to assess the access of acoustic properties by the human perceptual system (especially in development) in more detail and consider acoustic distances in stimulus selection. We here investigated the neurophysiological underpinnings of the perceptual reorganisation of a tone contrast, for which a behavioural study had shown a decrease in perceptual sensitivity in infancy (Götz et al., 2018). Contrary to the behavioural study, the current ERP study revealed retained perceptual sensitivity to a non-native tone contrast in younger and older infants, even with different neurophysiological responses. These findings contradict previous results showing a pMMR in 5- to 6-month-old English-learning infants but no MMR in 11- to 12-month-olds for a Mandarin lexical-tone contrast (Liu et al., 2019). Instead, our study lines up with behavioural work reporting either no evidence for a developmental change (Ramachers et al., 2018; Shi et al., 2017) or enhanced perceptual sensitivity to lexical-tone contrasts across age by infants learning a non-tone language (Chen & Kager, 2016; Chen et al., 2017).

How can the lack of a developmental attenuation of the neural response of German-learning infants to lexical tones be reconciled with non-conforming ERP and behavioural evidence on lexical-tone perception (Götz et al., 2018; Mattock et al., 2008; Yeung et al., 2013)? First, listeners' native intonation system may influence tone perception (see discussions in Götz et al., 2018; Liu & Kager, 2014). German uses pitch changes for pragmatic functions (e.g., marking a question or declarative sentence). Thus, pitch is part of the linguistic system and is, therefore, a familiar acoustic feature for German-learning infants. The observed neural responses to these pitch differences may reflect automatic perceptual processes, while responses in the behavioural studies required additional attention- and memory-related processes. The 9-month-old German-learning infants may not have attended to the tonal sound properties and thus not responded behaviourally since these sounds are not relevant for infants' language development. Second, as our ERP data in adults indicate, further developmental changes in the neural responses to lexical tones are expected beyond 9 months of age. Thus, studying an older infant group is required to fully understand the neural underpinnings of developmental changes in lexical-tone perception, especially in light of behavioural evidence of recovering discrimination abilities within the second year of life (Götz et al., 2018; Liu & Kager, 2014).

Limitations of our study

One limitation of the current study is that differences in the neural processing of lexical tones and vowel quality cannot be clearly attributed to either their association with the German phonological system or the acoustic properties of the stimuli themselves. Further studies need to explore how sound perception and its neural basis change across development concerning the acoustic salience and sound category of the tested sound contrasts, as well as the amount of language exposure and the phonological system of the listeners' native language. In our case, this implies the necessity of conducting a study involving native Cantonese speakers, both adults and infants. Additionally, longitudinal studies would determine the reliability of the observed age effects, and future studies should be run with larger sample sizes that allow for considering inter-individual differences in MMR polarity.

Finally, keeping the experiment duration feasible for the age of the tested infants, we did not employ a balanced oddball design (counterbalancing the stimuli presented as standards and deviants, respectively), in which the identical standard and deviant stimuli are used to calculate the MMRs. The current design may have resulted in confounds between the MMRs and sensory ERPs (such as the N1) stemming from the acoustic differences between the stimuli (see Jacobsen & Schröger, 2003; Scharinger *et al.*, 2011). Future studies should use lexical-tone and vowel-quality stimuli as standards and deviants across different experimental blocks to disentangle phonetic processing and basic auditory processing. Moreover, sound perception is affected by asymmetries with better discrimination performance in one direction (e.g., from sound A to sound B) than in the other direction (e.g., from sound B to sound A). These asymmetries have been reported for vowels (Polka & Bohn, 2011) and for lexical tones (Politzer-Ahles *et al.*, 2016). Future studies should address this issue and take asymmetrical perception effects into account.

Conclusion

Our study shows differences between the neural processing of lexical-tone and vowel-quality contrasts already in early infancy and continue into adulthood. We found that the developmental trajectory of processing these two contrasts is qualitatively different, with vowel quality leading to an adult-like neurophysiological response earlier than lexical tones. We propose that these differences may stem from varying exposure levels to the native language, which is larger for lexically relevant vowel-quality than for lexically irrelevant lexical-tone contrasts.

Supplementary material. The supplementary material for this article can be found at <http://doi.org/10.1017/S030500092400014X>.

Data availability statement. All data reported here are available upon request to the corresponding author.

Acknowledgments. The authors wish to thank the Babylab Team and all the families that participated in the study. We would also like to thank Annika Werwach for discussing the statistical analysis.

Funding. The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: the research presented here was funded by the DFG (German Research Foundation) as part of the Research Unit Crossing the Borders (FOR 2253) with grants to GS (Schw 665/12-1/2) and BH (HO 1960/19-1/2).

References

- Aslin, R. N., & Pisoni, D. B. (1980). Some developmental processes in speech perception. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology, Vol. 2: Perception* (pp. 67–96). Academic Press. <https://ci.nii.ac.jp/naid/10018504644/>
- Barry, J. G., Hardiman, M. J., & Bishop, D. V. (2009). Mismatch response to polysyllabic nonwords: A neurophysiological signature of language learning capacity. *PLoS one*, 4(7).
- Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*.
- Bijeljac-Babic, R., Serres, J., Höhle, B., & Nazzi, T. (2012). Effect of Bilingualism on Lexical Stress Pattern Discrimination in French-Learning Infants. *PLoS ONE*, 7(2). <https://doi.org/10.1371/journal.pone.0030843>
- Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer* (6.0.14) [Computer software]. <http://www.praat.org/>
- Čeponienė, R., Lepistö, T., Soininen, M., Aronen, E., Alku, P., & Näätänen, R. (2004). Event-related potentials associated with sound discrimination versus novelty detection in children. *Psychophysiology*, 41(1), 130–141.
- Chen, A., & Kager, R. (2016). Discrimination of Lexical Tones in the First Year of Life. *Infant and Child Development*. <https://doi.org/10.1002/icd.1944>
- Chen, A., Peter, V., Wijnen, F., Schnack, H., & Burnham, D. (2018). Are lexical tones musical? Native language's influence on neural response to pitch in different domains. *Brain and Language*, 180–182, 31–41. <https://doi.org/10.1016/j.bandl.2018.04.006>
- Chen, A., Stevens, C. J., & Kager, R. (2017). Pitch Perception in the First Year of Life, a Comparison of Lexical Tones and Musical Pitch. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.00297>
- Cheng, Y.-Y., & Lee, C.-Y. (2018). The Development of Mismatch Responses to Mandarin Lexical Tone in 12- to 24-Month-Old Infants. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.00448>
- Cheng, Y.-Y., Wu, H.-C., Tzeng, Y.-L., Yang, M.-T., Zhao, L.-L., & Lee, C.-Y. (2013). The Development of Mismatch Responses to Mandarin Lexical Tones in Early Infancy. *Developmental Neuropsychology*, 38(5), 281–300. <https://doi.org/10.1080/87565641.2013.799672>
- Cheng, Y.-Y., Wu, H.-C., Tzeng, Y.-L., Yang, M.-T., Zhao, L.-L., & Lee, C.-Y. (2015). Feature-specific transition from positive mismatch response to mismatch negativity in early infancy: Mismatch responses to vowels and initial consonants. *International Journal of Psychophysiology*, 96(2), 84–94.
- Cheour, M., Alho, K., Čeponienė, R., Reinikainen, K., Sainio, K., Pohjavuori, M., Aaltonen, O., & Näätänen, R. (1998). Maturation of mismatch negativity in infants. *International Journal of Psychophysiology*, 29(2), 217–226. [https://doi.org/10.1016/S0167-8760\(98\)00017-8](https://doi.org/10.1016/S0167-8760(98)00017-8)
- Chládková, K., & Paillereau, N. (2020). The what and when of universal perception: A review of early speech sound acquisition. *Language Learning*, 70(4), 1136–1182.
- Chládková, K., Urbanec, J., Skálová, S., & Kremláček, J. (2021). Newborns' neural processing of native vowels reveals directional asymmetries. *Developmental Cognitive Neuroscience*, 101023.
- de Klerk, M., de Bree, E., Kerkhoff, A., & Wijnen, F. (2019). Lost and Found: Decline and Reemergence of Non-Native Vowel Discrimination in the First Year of Life. *Language Learning and Development*, 15(1), 14–31.
- Friederici, A. D. C., Friedrich, M., & Weber, C. (2002). Neural manifestation of cognitive and precognitive mismatch detection in early infancy. *Neuroreport*, 13(10), 1251–1254.
- García-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. K. (2016). Relationships between quantity of language input and brain responses in bilingual and monolingual infants. *International Journal of Psychophysiology*, 110, 1–17.
- Götz, A., Yeung, H. H., Krasotkina, A., Schwarzer, G., & Höhle, B. (2018). Perceptual Reorganisation of Lexical Tones: Effects of Age and Experimental Procedure. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.00477>
- Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and clinical neurophysiology*, 55(4), 468–484.
- 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 and Development*, 32(3), 262–274.

- Jacobsen, T., & Schröger, E. (2003). Measuring duration mismatch negativity. *Clinical Neurophysiology*, *114*(6), 1133–1143. [https://doi.org/10.1016/S1388-2457\(03\)00043-9](https://doi.org/10.1016/S1388-2457(03)00043-9)
- Kaan, E., Barkley, C. M., Bao, M., & Wayland, R. (2008). Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: An event-related potentials training study. *BMC neuroscience*, *9*(1), 1.
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*(2), F13–F21. <https://doi.org/10.1111/j.1467-7687.2006.00468.x>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. *R package version*, *1*(1), 3.
- Leppänen, P. H., Guttorm, T. K., Pihko, E., Takkinen, S., Eklund, K. M., & Lyytinen, H. (2004). Maturation effects on newborn ERPs measured in the mismatch negativity paradigm. *Experimental Neurology*, *190*, 91–101.
- Liu, L., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, *133*(2), 385–394.
- Liu, L., Ong, J. H., Tuninetti, A., & Escudero, P. (2018). One Way or Another: Evidence for Perceptual Asymmetry in Pre-attentive Learning of Non-native Contrasts. *Frontiers in Psychology*, *9*. <https://doi.org/10.3389/fpsyg.2018.00162>
- Liu, L., Peter, V., & Weidemann, G. (2019). A bilingual advantage in infant pitch processing. *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS2019)*, 5-9 August 2019, Melbourne, Australia, 1397–1401.
- Marklund, E., Schwarz, I.-C., & Lacerda, F. (2019). Amount of speech exposure predicts vowel perception in four- to eight-month-olds. *Developmental Cognitive Neuroscience*, *36*.
- Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, *106*(3), 1367–1381.
- Mazuka, R., Hasegawa, M., & Tsuji, S. (2014). Development of non-native vowel discrimination: Improvement without exposure. *Developmental Psychobiology*, *56*(2), 192–209. <https://doi.org/10.1002/dev.21193>
- Morr, M. L., Shafer, V. L., Kreuzer, J. A., & Kurtzberg, D. (2002). Maturation of Mismatch Negativity in Typically Developing Infants and Preschool Children. *Ear and Hearing*, *23*(2), 118–136.
- Mueller, J. L., Friederici, A. D., & Männel, C. (2012). Auditory perception at the root of language learning. *Proceedings of the National Academy of Sciences*, *109*(39), 15953–15958.
- Mueller, V., Brehmer, Y., von Oertzen, T., Li, S.-C., & Lindenberger, U. (2008). Electrophysiological correlates of selective attention: A lifespan comparison. *BMC Neuroscience*, *9*(1), 18. <https://doi.org/10.1186/1471-2202-9-18>
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*, 2544–2590.
- Narayan, C. R. (2019). An acoustic perspective on 45 years of infant speech perception, Part I: Consonants. *Language and Linguistics Compass*, *13*(10), e12352. <https://doi.org/10.1111/lnc3.12352>
- Narayan, C. R. (2020). An acoustic perspective on 45 years of infant speech perception. II. Vowels and suprasegmentals. *Language and Linguistics Compass*, *n/a*(n/a), e12369. <https://doi.org/10.1111/lnc3.12369>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.
- Peter, V., Kalashnikova, M., Santos, A., & Burnham, D. (2016). Mature neural responses to infant-directed speech but not adult-directed speech in pre-verbal infants. *Scientific reports*, *6*(1), 34273.
- Politzer-Ahles, S., Schluter, K., Wu, K., & Almeida, D. (2016). Asymmetries in the perception of Mandarin tones: Evidence from mismatch negativity. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(10), 1547.
- Polka, L., & Bohn, O.-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, *39*(4), 467–478. <https://doi.org/10.1016/j.wocn.2010.08.007>
- Ramachers, S., Brouwer, S., & Fikkert, P. (2018). No perceptual reorganisation for Limburgian tones? A cross-linguistic investigation with 6- to 12-month-old infants. *Journal of Child Language*, *45*(2), 290.
- R Core Team (2021). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL <https://www.R-project.org/>

- Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science*, *8*(2), 162–172. <https://doi.org/10.1111/j.1467-7687.2005.00403.x>
- Schaadt, G., Männel, C., van der Meer, E., Pannekamp, A., Oberecker, R., & Friederici, A. D. (2015). Present and past: Can writing abilities in school children be associated with their auditory discrimination capacities in infancy? *Research in Developmental Disabilities*, *47*, 318–333. <https://doi.org/10.1016/j.ridd.2015.10.002>
- Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038. <https://doi.org/10.1016/j.jml.2019.104038>
- Scharinger, M., Idsardi, W. J., & Poe, S. (2011). A comprehensive three-dimensional cortical map of vowel space. *Journal of cognitive neuroscience*, *23*(12), 3972–3982.
- Shafer, V. L., Yu, Y. H., & Datta, H. (2011). The development of English vowel perception in monolingual and bilingual infants: Neurophysiological correlates. *Journal of Phonetics*, *39*(4), 527–545. <https://doi.org/10.1016/j.wocn.2010.11.010>
- Shi, R., Santos, E., Gao, J., & Li, A. (2017). Perception of Similar and Dissimilar Lexical Tones by Non-Tone-Learning Infants. *Infancy*, *22*(6), 790–800.
- Thiede, A., Virtala, P., Ala-Kurikka, I., Partanen, E., Huotilainen, M., Mikkola, K., Leppänen, P. H. T., & Kujala, T. (2019). An extensive pattern of atypical neural speech-sound discrimination in newborns at risk of dyslexia. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *130*(5), 634–646. <https://doi.org/10.1016/j.clinph.2019.01.019>
- Tsujii, S., & Cristia, A. (2014). Perceptual attunement in vowels: A meta-analysis. *Developmental Psychology*, *56*(2), 179–191. <https://doi.org/10.1002/dev.21179>
- Weber, C., Hahne, A., Friedrich, M., & Friederici, A. D. (2004). Discrimination of word stress in early infant perception: Electrophysiological evidence. *Cognitive Brain Research*, *18*(2), 149–161.
- Werker, J. F., & Gervain, J. (2013). Speech perception in infancy: A foundation for language acquisition. *The Oxford Handbook of Developmental Psychology*, *1*, 909–925.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganisation during the first year of life. *Infant behavior and development*, *7*(1), 49–63. <https://doi.org/10.1121/1.390988>
- Werwach, A., Männel, C., Obrig, H., Friederici, A. D., & Schaadt, G. (2022). Longitudinal trajectories of electrophysiological mismatch responses in infant speech discrimination differ across speech features. *Developmental cognitive neuroscience*, *56*, 101127.
- Wickham, H. (2016). Programming with ggplot2. In H. Wickham (Hrsg.), *Ggplot2: Elegant Graphics for Data Analysis* (S. 241–253). Springer International Publishing. https://doi.org/10.1007/978-3-319-24277-4_12
- Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, *68*(2), 123–139.
- Yu, A. (2010). Tonal effects on perceived vowel duration. In C. Fougeron, B. Kühnert, M. D’Imperio & N. Vallée (Eds.), *Laboratory Phonology 10* (pp. 151–168). Berlin, New York: De Gruyter Mouton. <https://doi.org/10.1515/9783110224917.2.151>
- Yu, H. Y., Shafer, V. L., & Sussman, E. S. (2017). Neurophysiological and behavioral responses of mandarin lexical tone processing. *Frontiers in Neuroscience*, *11*, 95.
- Yu, H. Y., Shafer, V. L., & Sussman, E. S. (2018). The duration of auditory sensory memory for vowel processing: Neurophysiological and behavioral measures. *Frontiers in Psychology*, *9*, 335.
- Yu, H. Y., Tessel, C., Han, H., Campanelli, L., Vidal, N., Gerometta, J., Garrido-Nag, K., Datta, H., & Shafer, V. L. (2019). Neural Indices of Vowel Discrimination in Monolingual and Bilingual Infants and Children. *Ear and Hearing*, *40*(6), 1376–1390.
- Yu, K., Chen, Y., Wang, M., Wang, R., & Li, L. (2022). Distinct but integrated processing of lexical tones, vowels, and consonants in tonal language speech perception: Evidence from mismatch negativity. *Journal of Neurolinguistics*, *61*, 101039.
- Zachau, S., Rinker, T., Körner, B., Kohls, G., Maas, V., Hennighausen, K., & Schecker, M. (2005). Extracting rules: Early and late mismatch negativity to tone patterns. *Neuroreport*, *16*(18), 2015–2019.

Cite this article: Götz, A., Männel, C., Schwarzer, G., Krasotkina, A., & Höhle, B. (2024). Neural correlates of lexical-tone and vowel-quality processing in 6- and 9-month-old German-learning infants and adults. *Journal of Child Language* 1–23. <https://doi.org/10.1017/S030500092400014X>