



Article

The Neural Development of Chinese Lexical Tone Perception: A Mismatch Negativity Study Across Childhood, Adolescence, and Adulthood

Han Wu ¹, Yixiao Zhang ² , Yiru Liu ³, Shijun Zhang ¹, Linjun Zhang ^{4,*}, Hua Shu ⁵ and Yang Zhang ⁶

¹ Institute on Education Policy and Evaluation of International Students, Beijing Language and Culture University, Beijing 100083, China; wuh2017@blcu.edu.cn (H.W.); 202321198055@stu.blcu.edu.cn (S.Z.)

² School of Biomedical Engineering, Dalian University of Technology, Dalian 116038, China; yixiao.zhang@donders.ru.nl

³ International Division, The Second High School Attached to Beijing Normal University, Beijing 100088, China; 11112019053@bnu.edu.cn

⁴ School of Chinese as a Second Language, Peking University, No. 5 Yiheyuan Road, Haidian District, Beijing 100871, China

⁵ State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China; shuh@bnu.edu.cn

⁶ Department of Speech-Language-Hearing Sciences and Center for Neurobehavioral Development, University of Minnesota, Minneapolis, MN 55455, USA; zhang470@umn.edu

* Correspondence: zhanglinjun@pku.edu.cn

Abstract: Background/Objectives: In a tonal language like Chinese, phonologically contrasting tones signify word meanings at the syllable level. Although the development of lexical tone perception ability has been examined in many behavioral studies, its developmental trajectory from childhood to adulthood at the neural level remains unclear. This cross-sectional study aimed to examine the issue by measuring the mismatch negativity (MMN) response to a Chinese lexical tonal contrast in three groups. Methods: The MMN response to a flat-falling tonal contrast (Tone1 versus Tone4) were recorded from children (25 participants aged 6–8), adolescents (26 participants aged 12–14), and young adults (20 participants aged 18–24). Nonsense speech stimuli were also used by superimposing Tone1 and Tone4 on an English syllable. Results: All three groups demonstrated typical early MMN responses in both the meaningful and nonsense syllable conditions. However, the MMN amplitudes varied significantly across groups, with the child group showing smaller responses compared to the adolescent and adult groups, while the latter two groups had similar MMN amplitudes. Conclusions: Neural sensitivity to tonal contrasts is not fully mature in children and reaches a more adult-like level during adolescence, with no significant difference in sensitivity to meaningful versus nonsense syllables. These results provide new insights into the neural development of lexical tone perception in a tonal language, highlighting its maturation during adolescence in this process.

Keywords: mismatch negativity (MMN); lexical tone perception; development; childhood; adolescence



Academic Editor: Alberto Pisoni

Received: 11 December 2024

Revised: 14 January 2025

Accepted: 15 January 2025

Published: 19 January 2025

Citation: Wu, H.; Zhang, Y.; Liu, Y.; Zhang, S.; Zhang, L.; Shu, H.; Zhang, Y. The Neural Development of Chinese Lexical Tone Perception: A Mismatch Negativity Study Across Childhood, Adolescence, and Adulthood. *Brain Sci.* **2025**, *15*, 93. <https://doi.org/10.3390/brainsci15010093>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The study of speech perception provides an important entry point for understanding how humans acquire their native language. Research has shown that infants gradually improve their ability to recognize native language sounds and features, while their ability to discriminate nonnative language features declines [1–5]. This preference for the native

language sounds in infancy can, to some extent, predict the development of future reading skills [6,7]. However, speech perception does not fully mature during infancy. Instead, it continues to develop over time, influenced by their linguistic environment and the maturation of the auditory system [6,8–11]. Additionally, extensive reading and writing experiences can further enhance speech perception ability [12,13]. For example, even at age 12, children have not yet reached the level of categorical perception seen in adults and are not able to use fine-grained spectral cues as adults do [14]. The precise age at which speech perception fully matures remains unclear [15,16]. Moreover, most previous studies have focused on non-tonal languages, leaving limited evidence on the development of lexical tone perception.

Behavioral and electrophysiological measures are most commonly used in speech perception research. Despite the task's simplicity, it is often challenging to use behavioral methods to examine speech perception in children. Compared to adults, children are less likely to maintain focus on a task for an extended period of time and their motivation or patience may easily influence the task outcomes. Therefore, it is unclear whether behavioral improvements should be interpreted as an increase in the capacity to perceive speech sounds or as the ability to use strategies to complete a task. By contrast, the electrophysiological method is an objective technique for recording the brain's activity in response to the stimuli. Specifically, the mismatch negativity (MMN) response, a well-known component of event-related potentials (ERPs), is widely used to investigate speech perception in infants, children, and adults [17–23]. MMN is a difference waveform obtained by subtracting the standard stimulus from the deviation stimulus in an oddball paradigm. As the MMN component reflects automatic detection of acoustic and phonological changes during speech perception, it is particularly suitable for testing speech perception in infants and children who are not required to engage in an active listening task [24].

Some previous studies have utilized MMN to examine the developmental aspects of speech perception [25–31]. For example, Shafer et al. measured the development of mismatch responses (MMRs) to the English vowel contrast (/I/ in bit versus /e/ in bet) in children from 4 to 5 and 6 to 7 years old [32]. Their findings demonstrated that in the early time window (200–300 ms), most of the younger children showed positive MMRs (i.e., p-MMRs), whereas more than half of the older children showed negative responses. Moreover, there was no significant difference in the later time window (300–400 ms) between the two groups of children. Bishop et al. examined MMNs elicited by the change from /ba/ to /da/ in children (7–12 years), teenagers (13–16 years), and adults (35 to 56 years) [33]. Their findings showed that in the early time window (100–250 ms), the amplitude of MMN increased with age, while in the late time window (300–550 ms), the amplitude of the late discriminative negativity (LDN) decreased with age.

Although there are numerous studies on the development of consonant and vowel perception, only a small number of studies have examined the development of Chinese lexical tone perception [16,34]. For example, Ref. [16] found that a large tonal difference (Tone1 versus Tone3) elicited typical MMNs in the 150–300 ms time window in all the three groups of children (i.e., 4, 5, and 6 years old), whereas a small tonal difference (Tone2 versus Tone3) did not elicit the typical MMN response but only P-MMRs in the 5- and 6-year-old groups. Liu et al. also showed that a small tonal difference (Tone2 versus Tone3) did not elicit the typical MMN responses but only elicited LDN responses in preschool (3–4 years) and school-aged children (7–8 years) [34]. No conclusive results are available on the development of lexical tone perception from childhood to adolescence and adulthood as only children were included in the previous studies.

To distinguish acoustic processing from phonological processing during speech perception, comparisons could be made between native versus nonnative speech sounds

and between speech versus nonspeech sounds [23,35–38]. For example, Näätänen et al. compared distinct phonological versus acoustic processing by using a Finnish vowel contrast and a Finish–Estonian vowel contrast [36]. Vandermosten et al. created nonspeech sounds by rotating the frequencies of the second formant transition of consonants for comparing speech and nonspeech processing [37]. Various acoustic manipulations have also been applied to lexical tones, but the results showed clear evidence of transfer of learning regardless of whether the tones were carried by native, nonnative, and nonspeech sounds [23,38–40]. For example, Xu et al. demonstrated that when native Thai speakers discriminated Thai tones that were superimposed on a Chinese syllable (so-called tonal chimeras), the left planum temporale was activated, indicating that phonological processes were recruited [38]. It remains to be tested whether tonal contrasts in the context of tonal chimeras and naturally produced lexical tones would be similarly perceived in children. In a recent article by Stilp et al., the researchers discussed the importance of considering the faithfulness of nonspeech stimuli to speech acoustics, their role within everyday hearing, and their relationship to listening experience and recognition [41]. In this regard, using chimeric lexical tones in a nonnative speech context provides a middle-ground approach to the study of phonological vs. acoustic processing. This approach adds a different acoustic control method from traditional comparisons using native vs. non-native or speech vs. nonspeech as it allows researchers to isolate the perceptual processes involved in recognizing identical tonal patterns with and without the influence of native phonological context. It can be adopted to provide insights into the developmental trajectory and maturation of tonal perception, which can reveal how children and adults may use similar or different processing strategies for native and nonnative speech signals carrying the same lexical tone contrast.

The current study aimed to examine the development of lexical tone perception at the neural level, which is reflected by the electrophysiological index (i.e., the MMN response) from childhood to adulthood. We recorded MMN responses to a Chinese tonal contrast (Tone1 versus Tone4) in children (6–8 years), adolescents (12–14 years), and young adults (aged 18–24 years), with the two lexical tones in a natural Chinese syllabic (natural tones, NT) context and a chimeric English syllabic (chimeric tones, CT) context, respectively. We predicted that typical early MMN responses would be observed in all the three groups and that the NT and CT contrasts would elicit similar MMN responses in the adult group irrespective of the syllable context. However, we were unable to make specific predictions about the developmental trajectory of MMN responses from childhood to adolescence and adulthood, or whether the contrasts in the NT and CT contexts would elicit similar MMN responses in the child and adolescent groups.

2. Materials and Methods

2.1. Participants

Three groups were included: Group 1, 25 children aged 6–8 years (mean age = 92.4 months, SD = 8.2); Group 2, 26 adolescents aged 12–14 years (mean age = 156.0 months, SD = 8.1); and Group 3, 20 young adults aged 18–24 years (mean age = 254.4 months, SD = 17.8). We used G*Power (Version 3.1.9.7) to verify a priori sample size requirement for MANOVA with three subject groups and six within-subject measures (three electrodes and two syllable contexts) [42]. To achieve a relatively large effect size of 0.4 at the alpha level of 0.05 and the power of 0.8 with our experimental setup, the required total sample size would be 56. Our total sample size of 71 exceeded the required minimum.

All participants had normal hearing with no history of neurological diseases or speech/language disorders. This study was approved by the ethics review board at Beijing

Language and Culture University (approval code: #2019-11-01). Informed consent was obtained from all of the adult participants and parents of the child participants.

2.2. Stimuli

Two sets of stimuli were constructed for this experiment by superimposing two Chinese lexical tones (i.e., Tone1 and Tone4) on the Chinese syllable /tɕi/ and English syllable /vi/. The original stimuli were recorded at a sampling rate of 44.1 kHz from a male Chinese–English bilingual speaker. The two monosyllables were digitally edited with duration (200 ms) and intensity (70 dB SPL) equalized. Pitch tier transfer was then performed using the Praat software (<http://www.fon.hum.uva.nl/praat/>, accessed on 1 December 2024) to superimpose pitch patterns of the two lexical tones on the two syllables while keeping the rest of the acoustic features identical. This procedure generated two pairs of stimuli, i.e., /tɕi1/ versus /tɕi4/ and /vi1/ versus /vi4/ (Figure 1).

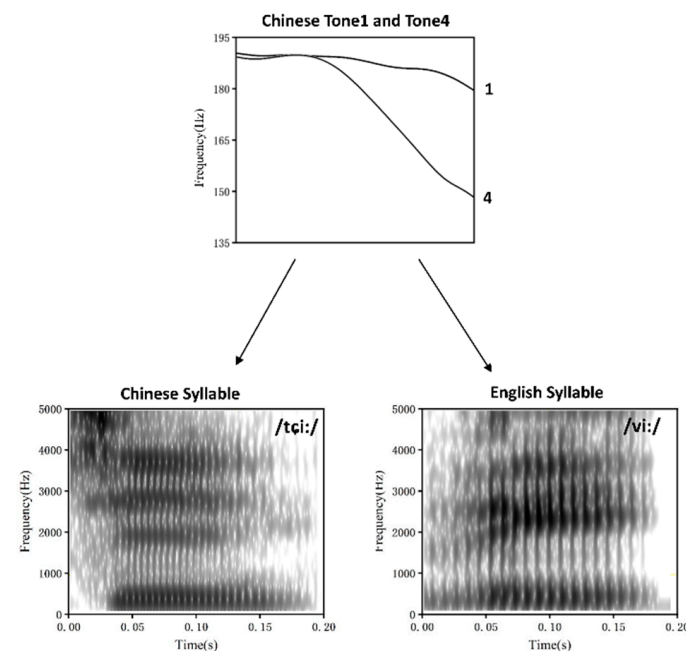


Figure 1. Schematic diagram illustrating how the two Chinese lexical tones (Tone1 and Tone4) are superimposed on the Chinese and English syllables (/tɕi/ and /vi/) to create Chinese and chimeric stimuli. The upper panel displays pitch contours of the two Chinese tones. The lower panel shows the broad-band spectrogram of the Chinese and English syllables.

2.3. ERP Procedure

Four oddball blocks were presented to each participant with each syllable condition consisting of two blocks in which the standard and deviant stimuli were counterbalanced. Each block began with 30 standard trials, followed by 723 trials comprising 87.5% standards and 12.5% deviants. The order of stimuli was pseudo-randomized with at least three successive standards between deviants. The interstimulus interval was 450 ms. The participants were seated on a comfortable chair in a shielded chamber. They were instructed to watch a selected animation movie but to ignore the sounds. The movie was presented in a DVD player, which was set approximately 60 cm in front of the seat. Participants were informed to complete a questionnaire about the movie after the experiment.

2.4. EEG Recording and Data Analysis

The 128-channel EEG signals were recorded using a Hydrocel Geodesic Sensor Net (Electrical Geodesics Inc., Eugene, OR, USA) referenced to Cz. Data were sampled at

500 Hz/channel with filters set at 0.01–100 Hz and calibrated technical zero baselines. Electrode impedances were maintained below 50 k Ω during the recording.

Off-line analysis of EEG signals was re-referenced to the average of all scalp channels. Signals were digitally filtered with a 0.3–30 Hz band-pass filter using Netstation software (Version 4.2). The filtered EEG data were first submitted to BESA for bad channel interpolation. Specifically, data from channels with excessive artifacts were spline-interpolated (BESA 5.1 software) and artifacts exceeding ± 100 μ V in any channel were automatically rejected. Independent component analysis (ICA) was then performed in EEGLAB to remove eye blinks and eye movements. Epochs were set at 600 ms in length, including a 100 ms pre-stimulus baseline.

Average ERPs were calculated separately for standard and deviant stimuli in the two syllable contexts for each participant, and difference waveforms were produced by subtracting the standard waveform from the deviant waveform for every participant. As the literature review in the Introduction shows, MMN quantification using subtracted ERP waveforms would face the polarity mixture problem, particularly for the child and adolescent groups. This issue arises because the presence of both MMN and P-MMR responses within these age groups can create variability in the polarity of the waveforms, making it difficult to accurately quantify and compare the data across individuals. There are several strategies to address this problem for statistical comparison between groups. One approach is to separate subgroups for analysis to ensure that each subgroup's data can be evaluated independently. Alternatively, researchers have attempted to rectify the polarity discrepancies among subjects using methods such as global field power measures [43,44] or time-frequency analysis [33,44,45], which involves analyzing spectral power or inter-trial phase coherence using phase-locking values. Another technique involves temporal principal component analysis (t-PCA) [46,47], which can realign the polarity of selected principal components for the target MMN/p-MMR activity to facilitate between-group comparison. By applying PCA to the time series data, researchers would have the opportunity to evaluate the components that maximally account for the variance in the ERP data and isolate the MMN response from overlapping brain activities to enhance the reliability of MMN amplitude and latency quantification based on its temporal and spatial patterns [48].

In our t-PCA analysis, a fourth-order tensor (electrodes $\times$ time points $\times$ stimuli $\times$ subjects) was constructed and imported into the analysis toolbox [49]. Here, “electrodes” refer to the EEG recording channels, “time points” correspond to the sampled intervals of the averaged ERPs, “stimuli” represent the two syllable contexts, and “subjects” encompass the children from the three experimental groups. The dataset was first filtered using the wavelet filter, and the selected mother wavelet is the reverse biorthogonal wavelet of the order of 6.8 (rbio6.8), which is ideal for decomposing conventionally averaged EEG recordings of children's MMN activity elicited by an uninterrupted sound using the oddball paradigm [50]. In order to extract the component of interest, t-PCA and Promax rotation were applied. Then, the data were projected to the electrode field to correct the indeterminacy of variance and polarity [49]. Finally, the amplitude of the MMN response for three electrode sites of interest (Fc3, Fc4, and Fcz) on the time window was extracted and the topographies were plotted. Based on the final grand average waveforms in this study, and the time window of MMN in the previous studies [51,52], we chose 30 ms wide windows placed 230–260 ms for the NT context and 250–280 ms for the CT context. A repeated measures MANOVA was performed to examine the difference in MMN responses, with group (group 1, 2, 3) as the between-subject factor and electrode (Fc3, Fc4, Fcz) and syllable context (NT and CT) as the within-subject factors.

3. Results

During the procedure of t-PCA, 16 components that explained 99% of the variance were retained. Based on the temporal and spatial properties of the MMN response and the similarity of spatial components among all the participants, the first and eleventh components, which, respectively, peaked at 274 ms and 236 ms with topographical distributions that characterized the MMN response, were selected. These components explained 51.42% and 1.14% of the variance, respectively (Figure 2). Although the polarity of the projected first and eleventh components was positive when viewed in the PCA space, they could be considered negative components because their original waveforms were negative-going peaks [53]. This happens because PCA maximizes variance and can rotate components in a way that changes their apparent polarity. By definition, variance is always non-negative. In PCA, the direction of the components can be reversed without affecting the variance they explain. As a result, components with negative peaks in the original signal can appear positive after PCA simply because of the reorientation of the components during the transformation process. However, when these components are back-projected into the electrode space, their spatial patterns align with their original negative peaks, revealing their true negative polarity. Thus, the polarity of the target components can be consistently realigned across subjects to resolve issues related to the polarity mixture problem in the mismatch responses.

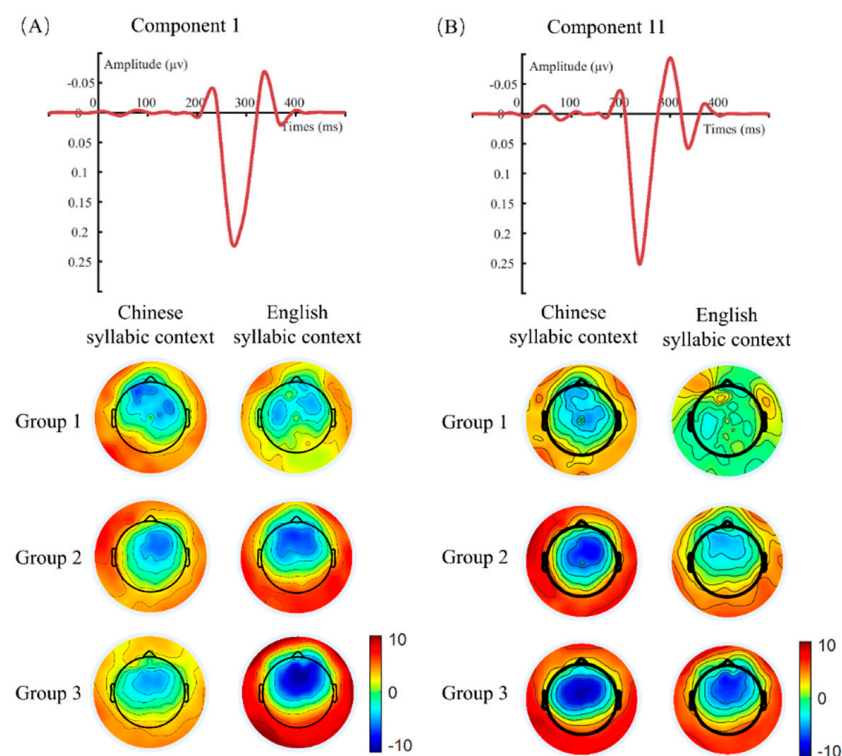


Figure 2. The waveforms and topographies of selected components. (A) The waveform and topography of component 1. Component 1 peaking at 274 ms explained 51.42% of the variance. (B) The waveform and topography of component 11. Component 11 peaking at 236 ms explained 1.14% of the variance. Each topography depicts the average scalp distribution of electrical potentials, with color gradients indicating neural activity magnitude and polarity. The map is characterized by a circular outline marking the reference plane, with the positions of the nose and ears providing a reference for the placement of electrodes.

The difference waveforms across the three age groups for the two syllable conditions and three electrodes, and the corresponding brain topographies, are shown in Figure 3. The results of the repeated measures MANOVA demonstrated that the main effect of the

group was significant ($F(2, 68) = 8.631, p < 0.001$), but the main effects of electrode ($F(2, 68) = 1.520, p = 0.222$) and syllable context ($F(1, 68) = 0.803, p = 0.373$) fell short of significance. Neither the two-way interactions nor the three-way interaction was significant (all $ps > 0.05$). Pairwise comparisons were applied as post hoc tests for the significant main effect of group. The results showed that the difference between the children and adolescent groups (Bonferroni corrected $p = 0.030$) and the difference between the children and adult groups (Bonferroni corrected $p < 0.001$) were significant.

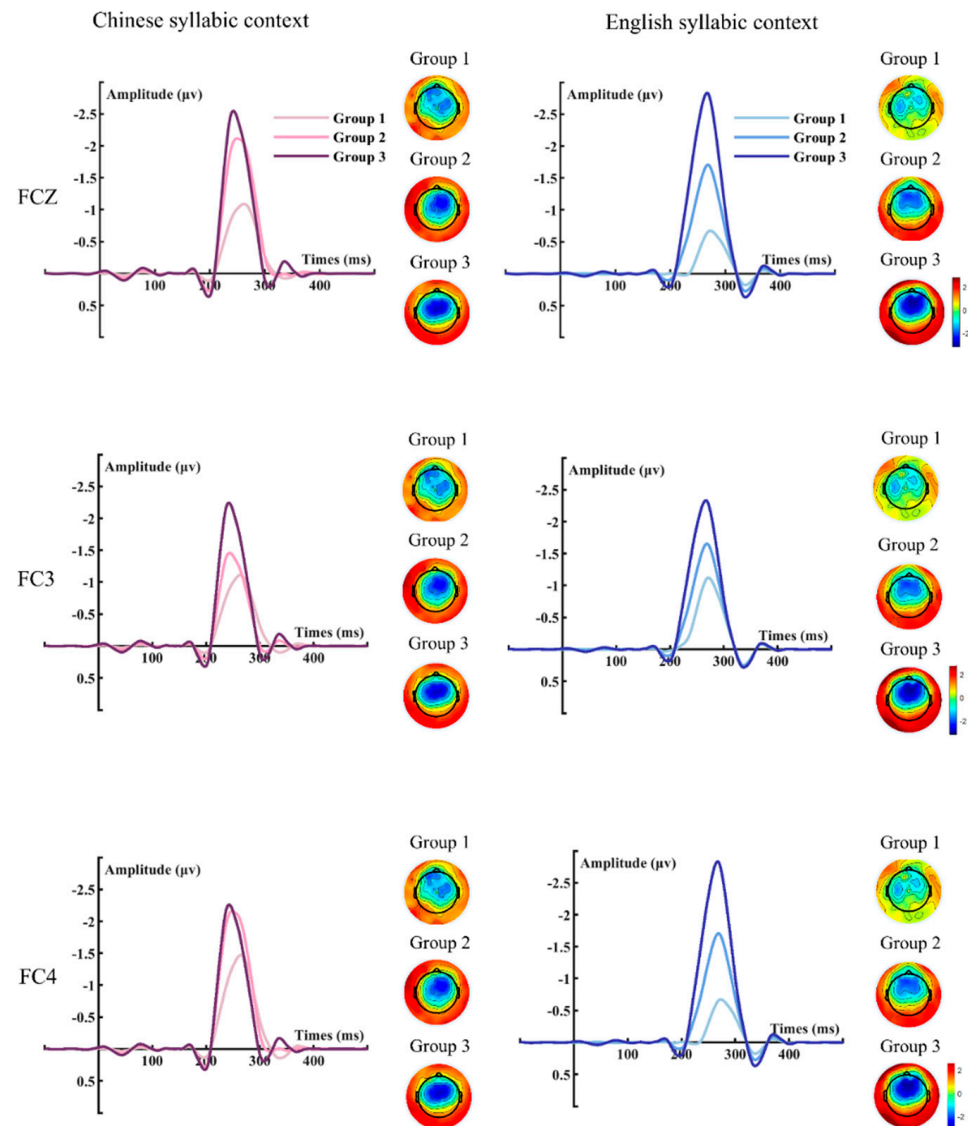


Figure 3. The average amplitudes and topographies of the three groups at Fcz, Fc3, and Fc4 electrodes for the Chinese syllabic and English syllabic contexts. The time window for the topography of Chinese syllabic context was set at 230–260 ms and the time windows for the topography of English syllabic context was set at 250–280 ms. Each topography depicts the average scalp distribution of electrical potentials, with color gradients indicating neural activity magnitude and polarity. The map is characterized by a circular outline marking the reference plane, with the positions of the nose and ears providing a reference for the placement of electrodes.

4. Discussion

The present study investigated the development of MMN response to the Chinese lexical tonal contrast from childhood to adolescence and young adulthood and aimed to clarify whether the development pattern was affected by the native (i.e., NT) versus nonnative (i.e., CT) syllable context on which the lexical tones were superimposed. Our

results revealed a significant difference in the MMN amplitude between the child and adolescent groups but a nonsignificant difference between the adolescent and adult groups, indicating that lexical tone perception develops continuously in childhood and does not fully mature until adolescence. Moreover, the same developmental pattern was observed in both NT and CT conditions, indicating that phonological knowledge of lexical tones is easily transferrable and affects the perception of pitch contours in the nonnative and nonsensical syllabic context.

Speech perception continues to develop from birth influenced by both the language environment and biological maturational changes to the auditory and cognitive systems [54]. However, the form and complexity of behavioral tasks make it challenging to track the development of children's speech perception skills. For example, young children are less adept than older children and adults at focusing on long and complex sequences of sounds and often lack the motivation and sustained attention to follow the task instructions. Consequently, it is often unclear in behavioral studies of speech perception whether poorer performance in children is due to inadequate speech perception or the underdevelopment of cognitive systems related to executive processing (e.g., decision making) [55]. Electrophysiological measures of brain responses are valuable for clarifying the source of children's speech perception performance, as they allow the examination at levels preceding behavioral responses. In particular, a great number of studies have adopted the oddball paradigm and ERP components (e.g., MMN and P3a) to examine speech perception by children of varying ages [56–65].

Previous findings have demonstrated that the MMN response is sensitive to a variety of acoustic and phonological contrasts. However, it remains unclear whether the typical early MMN is present in young children and how maturation influences the amplitude and topography of MMN [16,32,34,61,66–69]. In the present study, the typical early MMN response was observed in all the three groups. This result contrasts with previous research indicating that only the adults showed the typical early MMN in response to a Chinese tonal contrast, while two groups of children (preschool- and school-aged) showed negative responses only in the later time window [34]. This discrepancy might be partly due to differences in the tonal contrasts used in the two studies: Tone1 versus Tone4 has a large tonal difference in the present study, and Tone2 versus Tone3 involves a relatively small tonal difference in Liu et al.'s study. Interestingly, a recent study found that in preschool-aged children, the lexical tonal contrast with a large tonal difference (Tone1 versus Tone3) elicited the typical early MMN whereas the tonal pair with a small tonal difference (Tone2 versus Tone3) did not produce significant MMNs in either the early or late time windows [70]. Furthermore, the results of the present study showed that the amplitude of MMN in the adolescent group was significantly larger than in the child group, revealing for the first time that the ability to discriminate Chinese lexical tones (at least those with a large tonal difference) is an ongoing process from childhood to adolescence.

MMN is sensitive to both acoustic and phonological changes with phonological changes eliciting larger MMNs than equivalent acoustic changes [23,36]. In the present study, the detection of change in pitch contours elicited similar MMNs irrespective of whether the pitch contours were superimposed on Chinese syllables to form Chinese lexical tones or on English syllables to form tonal chimeras. This finding indicates that the phonological knowledge of lexical tones is easily transferrable and affects pitch perception. Our results are consistent with previous findings that native Chinese speakers' perception of pitch in nonnative speech and even nonspeech sounds is influenced by their linguistic experience with lexical tones [23,38–40]. For example, Bent et al. found that native language experience affected the identification of nonspeech pitch contours [39]. Specifically, native Chinese speakers were more likely to misidentify flat and falling pitch contours than native

English speakers and these misidentifications were linked to the specific features of Chinese lexical tones. In a study by Xu et al., activation of the left planum temporale, a brain region associated with phonological processing, was observed when native Thai speakers discriminated Thai tones superimposed on a Chinese syllable (tonal chimeras) [38]. The present study further indicates that lexical tone knowledge affects the perception of pitch contours in nonnative speech sounds, even in children below the age of 8 years old. These results underscore the adaptability and transferability of linguistic knowledge. Children who are exposed to and learn tonal language patterns can apply this knowledge to recognize pitch variations in nonnative speech, such as in tonal chimeras. The ability to transfer phonological knowledge to nonnative contexts suggests that early linguistic experiences shape not only the perception of native speech but also how children process and interpret unfamiliar or modified speech signals. This ability to generalize tonal processing is significant as it highlights the far-reaching impact of early language exposure on auditory perception and the development of cognitive processing related to lexical tone discrimination.

Interestingly, the detection of pitch contour changes in natural speech (230–260 ms time window) was faster than that in English syllables (250–280 ms time window). This advantage in detection of pitch contour changes in native language syllables is consistent with the finding of Xi et al. that the MMN response to pitch changes in native syllables peaked earlier than that in nonspeech harmonics [23]. This effect might be attributed to the participants' higher familiarity with Chinese syllables relative to English syllables. Specifically, the participants' phonetic processing of the native syllables and familiarity with the NT stimuli might have led to faster detection of the difference in pitch contours.

While the present study provides insights into the development of the MMN response to Chinese lexical tonal contrasts and the transferability of phonological knowledge to nonnative pitch perception, several limitations should be acknowledged. First, the study's participant groups were limited to specific age ranges (children aged 6–8 years, adolescents aged 12–14 years, and young adults aged 18–24 years). A larger and more diverse sample that includes a broader range of ages and linguistic backgrounds could provide more comprehensive insights into how lexical tone perception develops across different populations. Second, while the use of tonal chimeras allowed for controlled comparisons between native and nonnative syllables, future research could benefit from including nonspeech tonal contrasts or speech stimuli that more closely mimic natural speech variability. This could help determine if the observed results generalize to real-world speech perception scenarios. Third, although the use of t-PCA and component polarity realignment improved the quantification of the MMN response, potential limitations remain in the PCA approach, such as the choice of components and the subjective nature of component selection. Additional verification methods, such as cross-validation techniques, can be implemented to ensure the robustness of the findings. Furthermore, we propose adopting alternative dimensionality reduction and component analysis methods and implementing more automated approaches for component selection to reduce the inherent subjectivity in the process. These approaches might further strengthen the methodological rigor and enhance the reproducibility of our findings.

The observation that children can transfer their phonological knowledge of lexical tones to perceive pitch contours in nonnative syllables underscores the impact of linguistic experience on auditory processing. This suggests that early exposure to tonal languages may enhance general auditory processing abilities that extend beyond native speech. Understanding that phonological knowledge can be transferred to nonnative contexts may support the development of training programs that leverage a child's existing language skills to facilitate learning of additional languages, especially those with different intonation or lexical tone structures. Future studies should explore the long-term impact of early

tonal language exposure on auditory processing and cognitive development. It would also be valuable to examine how the development of MMN responses to tonal contrasts might influence other cognitive functions, such as memory and attention, and to assess whether similar patterns are observed in individuals with varying levels of exposure to tonal languages or various difficulties with linguistic prosody processing [44,71–73]. Additionally, investigating the effects of environmental factors, such as bilingualism or socioeconomic status, could provide deeper insights into the variability of MMN development and tonal perception across different populations.

5. Conclusions

In conclusion, our findings provide evidence that the MMN response to the Chinese lexical tonal contrast continues to develop and mature from childhood to adolescence, suggesting a gradual refinement of lexical tone perception capabilities over time. This neural development pattern indicates that the ability to detect and differentiate pitch changes, which is fundamental for recognizing tonal contrasts, is an ongoing process that evolves as children age and their auditory and cognitive systems mature. The observed differences in MMN amplitude between the child and adolescent groups further emphasize that the ability to process and respond to lexical tone contrasts becomes more robust during the transition from early to late childhood. Moreover, the phonological knowledge of lexical tones extends to the perception of pitch contours in both native and nonnative speech sounds from childhood.

Author Contributions: Conceptualization, H.W. and L.Z.; Methodology, Y.Z. (Yixiao Zhang); Software, Y.Z. (Yixiao Zhang); Validation, H.W., L.Z. and Y.Z. (Yang Zhang); Formal Analysis, H.W.; Investigation, S.Z.; Resources, H.S.; Data Curation, Y.L.; Writing—Original Draft Preparation, H.W.; Writing—Review and Editing, L.Z. and Y.Z. (Yang Zhang); Visualization, S.Z.; Supervision, H.S.; Project Administration, L.Z.; Funding Acquisition, L.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Social Science Fund of China (20BYY092) to L.Z.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the ethics review board at Beijing Language and Culture University (approval code: #2019-11-01, 1 November 2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request due to restrictions (privacy, legal and ethical reasons).

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Kuhl, P.K.; Stevens, E.; Hayashi, A.; Deguchi, T.; Kiritani, S.; Iverson, P. Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Dev. Sci.* **2006**, *9*, F13–F21. [[CrossRef](#)] [[PubMed](#)]
2. Lovcevic, I.; Tsuji, S. The developmental pattern of native and non-native speech perception during the 1st year of life in Japanese infants. *Infant Behav. Dev.* **2024**, *76*, 101977. [[CrossRef](#)] [[PubMed](#)]
3. Chládková, K.; Paillereau, N. The What and When of Universal Perception: A Review of Early Speech Sound Acquisition. *Lang. Learn.* **2020**, *70*, 1136–1182. [[CrossRef](#)]
4. Best, C.C.; McRoberts, G.W. Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Lang. Speech* **2003**, *46*, 183–216. [[CrossRef](#)]
5. Tsao, F.M.; Liu, H.M.; Kuhl, P.K. Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *J. Acoust. Soc. Am.* **2006**, *120*, 2285–2294. [[CrossRef](#)] [[PubMed](#)]

6. Kuhl, P.K.; Conboy, B.T.; Padden, D.; Nelson, T.; Pruitt, J. Early Speech Perception and Later Language Development: Implications for the “Critical Period”. *Lang. Learn. Dev.* **2005**, *1*, 237–264. [[CrossRef](#)]
7. Kuhl, P.; Rivera-Gaxiola, M. Neural substrates of language acquisition. *Annu. Rev. Neurosci.* **2008**, *31*, 511–534. [[CrossRef](#)] [[PubMed](#)]
8. Moore, J.K.; Guan, Y.L. Cytoarchitectural and axonal maturation in human auditory cortex. *J. Assoc. Res. Otolaryngol.* **2001**, *2*, 297–311. [[CrossRef](#)]
9. Moore, J.K.; Linthicum, F.H. The human auditory system: A timeline of development. *Int. J. Audiol.* **2007**, *46*, 460–478. [[CrossRef](#)] [[PubMed](#)]
10. Werker, J.F.; Tees, R.C. Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Dev. Psychobiol.* **2005**, *46*, 233–251. [[CrossRef](#)] [[PubMed](#)]
11. Jusczyk, P.W. Toward a model of the development of speech perception. In *Invariance and Variability in Speech Processes*; Psychology Press: Hove, UK, 2014; pp. 1–35.
12. Hong, T.; Shuai, L.; Frost, S.J.; Landi, N.; Pugh, K.R.; Shu, H. Cortical responses to Chinese phonemes in preschoolers predict their literacy skills at school age. *Dev. Neuropsychol.* **2018**, *43*, 356–369. [[CrossRef](#)] [[PubMed](#)]
13. Vandewalle, E.; Boets, B.; Ghesquiere, P.; Zink, I. Auditory processing and speech perception in children with specific language impairment: Relations with oral language and literacy skills. *Res. Dev. Disabil.* **2012**, *33*, 635–644. [[CrossRef](#)] [[PubMed](#)]
14. Hazan, V.; Barrett, S. The development of phonemic categorization in children aged 6–12. *J. Phon.* **2000**, *28*, 377–396. [[CrossRef](#)]
15. Chen, F.; Peng, G.; Yan, N.; Wang, L. The development of categorical perception of Mandarin tones in four- to seven-year-old children. *J. Child Lang.* **2017**, *44*, 1413–1434. [[CrossRef](#)]
16. Lee, C.Y.; Yen, H.L.; Yeh, P.W.; Lin, W.H.; Cheng, Y.Y.; Tzeng, Y.L.; Wu, H.C. Mismatch responses to lexical tone, initial consonant, and vowel in Mandarin-speaking preschoolers. *Neuropsychologia* **2012**, *50*, 3228–3239. [[CrossRef](#)]
17. Näätänen, R. The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology* **2001**, *38*, 1–21. [[CrossRef](#)]
18. Turgeon, C.; Lazzouni, L.; Lepore, F.; Ellemberg, D. An objective auditory measure to assess speech recognition in adult cochlear implant users. *Clin. Neurophysiol.* **2014**, *125*, 827–835. [[CrossRef](#)] [[PubMed](#)]
19. Cheour, M.; Korpilahti, P.; Martynova, O.; Lang, A.H. Mismatch negativity and late discriminative negativity in investigating speech perception and learning in children and infants. *Audiol. Neuro-Otol.* **2001**, *6*, 2–11. [[CrossRef](#)]
20. Schulte-Körne, G.; Deimel, W.; Bartling, J.; Remschmidt, H. Speech perception deficit in dyslexic adults as measured by mismatch negativity (MMN). *Int. J. Psychophysiol.* **2001**, *40*, 77–87. [[CrossRef](#)] [[PubMed](#)]
21. Werwach, A.; Maennel, C.; Obrig, H.; Friederici, A.D.; Schaadt, G. Longitudinal trajectories of electrophysiological mismatch responses in infant speech discrimination differ across speech features. *Dev. Cogn. Neurosci.* **2022**, *56*, 101127. [[CrossRef](#)] [[PubMed](#)]
22. Chen, A.; Peter, V.; Burnham, D. Development of neural discrimination of pitch across speech and music in the first year of life, a mismatch response study. *Lang. Cogn. Neurosci.* **2022**, *37*, 1153–1168. [[CrossRef](#)]
23. Xi, J.; Zhang, L.; Shu, H.; Zhang, Y.; Li, P. Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience* **2010**, *170*, 223–231. [[CrossRef](#)] [[PubMed](#)]
24. Näätänen, R.; Paavilainen, P.; Rinne, T.; Alho, K. The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clin. Neurophysiol.* **2007**, *118*, 2544–2590. [[CrossRef](#)] [[PubMed](#)]
25. Ceponiene, R.; Lepistö, T.; Alku, P.; Aro, H.; Näätänen, R. Event-related potential indices of auditory vowel processing in 3-year-old children. *Clin. Neurophysiol.* **2003**, *114*, 652–661. [[CrossRef](#)] [[PubMed](#)]
26. Cheng, Y.Y.; Wu, H.C.; Tzeng, Y.L.; Yang, M.T.; Zhao, L.L.; Lee, C.Y. The Development of Mismatch Responses to Mandarin Lexical Tones in Early Infancy. *Dev. Neuropsychol.* **2013**, *38*, 281–300. [[CrossRef](#)]
27. Csepe, V. On the origin and development of the mismatch negativity. *Ear Hear* **1995**, *16*, 91–104. [[CrossRef](#)] [[PubMed](#)]
28. Kristin, L.; Jennifer, M.; Hannah, D.; Ruth, R. Mapping development of the MMN and P3a potentials during adolescence: A longitudinal investigation of healthy individuals and individuals at-risk for psychosis. In *Proceedings of the Conference Abstract: XII International Conference on Cognitive Neuroscience (ICON-XII)*, Brisbane, Australia, 27–31 July 2014; Volume 9. [[CrossRef](#)]
29. Maurer, U.; Bucher, K.; Brem, S.; Brandeis, D. Development of the automatic mismatch response: From frontal positivity in kindergarten children to the mismatch negativity. *Clin. Neurophysiol.* **2003**, *114*, 808–817. [[CrossRef](#)]
30. McGee, T.J.; King, C.; Tremblay, K.; Nicol, T.G.; Cunningham, J.; Kraus, N. Long-term habituation of the speech-elicited mismatch negativity. *Psychophysiology* **2001**, *38*, 653–658. [[CrossRef](#)]
31. Kraus, N.; McGee, T.; Carrell, T.; Sharma, A.; Micco, A.; Nicol, T. Speech-evoked cortical potentials in children. *J. Am. Acad. Audiol.* **1993**, *4*, 238–248. [[PubMed](#)]
32. Shafer, V.L.; Yu, Y.H.; Datta, H. Maturation of Speech Discrimination in 4- to 7-Yr-Old Children as Indexed by Event-Related Potential Mismatch Responses. *Ear Hear* **2010**, *31*, 735–745. [[CrossRef](#)] [[PubMed](#)]
33. Bishop, D.V.; Hardiman, M.J.; Barry, J.G. Is auditory discrimination mature by middle childhood? A study using time-frequency analysis of mismatch responses from 7 years to adulthood. *Dev. Sci.* **2011**, *14*, 402–416. [[CrossRef](#)] [[PubMed](#)]

34. Liu, H.M.; Chen, Y.; Tsao, F.M. Developmental changes in mismatch responses to mandarin consonants and lexical tones from early to middle childhood. *PLoS ONE* **2014**, *9*, e95587. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Joanisse, M.F.; Zevin, J.D.; McCandliss, B.D. Brain mechanisms implicated in the preattentive categorization of speech sounds revealed using fMRI and a short-interval habituation trial paradigm. *Cereb. Cortex* **2007**, *17*, 2084–2093. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Näätänen, R.; Lehtokoski, A.; Lennes, M.; Cheour, M.; Huottilainen, M.; Iivonen, A.; Vainio, M.; Alku, P.; Ilmoniemi, R.J.; Luuk, A.; et al. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* **1997**, *385*, 432–434. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Vandermosten, M.; Boets, B.; Luts, H.; Poelmans, H.; Golestani, N.; Wouters, J.; Ghesquière, P. Adults with dyslexia are impaired in categorizing speech and nonspeech sounds on the basis of temporal cues. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 10389–10394. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Xu, Y.S.; Gandour, J.; Talavage, T.; Wong, D.; Dzemidzic, M.; Tong, Y.X.; Li, X.J.; Lowe, M. Activation of the left planum temporale in pitch processing is shaped by language experience. *Hum. Brain Mapp.* **2006**, *27*, 173–183. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Bent, T.; Bradlow, A.R.; Wright, B.A. The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *J. Exp. Psychol.-Hum. Percept. Perform.* **2006**, *32*, 97–103. [\[CrossRef\]](#)
40. Francis, A.L.; Ciocca, V.; Ng, B.K.C. On the (non)categorical perception of lexical tones. *Percept. Psychophys.* **2003**, *65*, 1029–1044. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Stilp, C.E.; Shorey, A.E.; King, C.J. Nonspeech sounds are not all equally good at being nonspeech. *J. Acoust. Soc. Am.* **2022**, *152*, 1842–1849. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Faul, F.; Erdfelder, E.; Buchner, A.; Lang, A.G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods* **2009**, *41*, 1149–1160. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Lehmann, D.; Skrandies, W. Spatial analysis of evoked potentials in man—A review. *Prog. Neurobiol.* **1984**, *23*, 227–250. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Wang, X.Y.; Wang, S.P.; Fan, Y.B.; Huang, D.; Zhang, Y. Speech-specific categorical perception deficit in autism: An Event-Related Potential study of lexical tone processing in Mandarin-speaking children. *Sci. Rep.* **2017**, *7*, 43254. [\[CrossRef\]](#)
45. Bishop, D.V.M.; Hardiman, M.J.; Barry, J.G. Lower-Frequency Event-Related Desynchronization: A Signature of Late Mismatch Responses to Sounds, Which Is Reduced or Absent in Children with Specific Language Impairment. *J. Neurosci.* **2010**, *30*, 15578–15584. [\[CrossRef\]](#)
46. Rivera-Gaxiola, M.; Silva-Pereyra, J.; Klarman, L.; Garcia-Sierra, A.; Lara-Ayala, L.; Cadena-Salazar, C.; Kuhl, P. Principal Component Analyses and scalp distribution of the auditory P150–250 and N250–550 to speech contrasts in Mexican and American infants. *Dev. Neuropsychol.* **2007**, *31*, 363–378. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Partanen, E.; Pakarinen, S.; Kujala, T.; Huottilainen, M. Infants' brain responses for speech sound changes in fast multifeature MMN paradigm. *Clin. Neurophysiol.* **2013**, *124*, 1578–1585. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Dien, J. Applying Principal Components Analysis to Event-Related Potentials: A Tutorial. *Dev. Neuropsychol.* **2012**, *37*, 497–517. [\[CrossRef\]](#)
49. Zhang, G.H.; Ristaniemi, T.; Cong, F.Y. Objective Extraction of Evoked Event-Related Oscillation from Time-Frequency Representation of Event-Related Potentials. *Neural. Plast.* **2020**, *2020*, 8841354. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Cong, F.; Huang, Y.; Kalyakin, I.; Li, H.; Huttunen-Scott, T.; Lyytinen, H.; Ristaniemi, T. Frequency-response-based wavelet decomposition for extracting children's mismatch negativity elicited by uninterrupted sound. *J. Med. Biol. Eng.* **2012**, *32*, 205–213. [\[CrossRef\]](#)
51. Wang, X.-D.; Liu, A.-P.; Wu, Y.-Y.; Wang, P. Rapid extraction of lexical tone phonology in Chinese characters: A visual mismatch negativity study. *PLoS ONE* **2013**, *8*, e56778. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Liang, X.Y.; Guo, Z.H.; Wang, X.D.; Guo, X.T.; Sun, J.W.; Wang, M.; Li, H.W.; Chen, L. Event-Related Potential Evidence for Involuntary Consciousness During Implicit Memory Retrieval. *Front. Behav. Neurosci.* **2022**, *16*, 902175. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Luck, S.J. *An Introduction to the Event-Related Potential Technique*; MIT Press: Cambridge, MA, USA, 2014.
54. Werner, L.A. Issues in human auditory development. *J. Commun. Disord.* **2007**, *40*, 275–283. [\[CrossRef\]](#)
55. Coady, J.A.; Kluender, K.R.; Evans, J.L. Categorical perception of speech by children with specific language impairments. *J. Speech Lang. Hear. Res.* **2005**, *48*, 944–959. [\[CrossRef\]](#)
56. Kraus, N.; McGee, T.; Sharma, A.; Carrell, T.; Nicol, T. mismatch negativity event-related potential elicited by speech stimuli. *Ear Hear* **1992**, *13*, 158–164. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Korpilahti, P.; Lang, H.A. Auditory ERP components and mismatch negativity in dysphasic children. *Electroencephalogr. Clin. Neurophysiol.* **1994**, *91*, 256–264. [\[CrossRef\]](#)
58. Kurtzberg, D.; Vaughan, H.G.; Kreuzer, J.A.; Flieger, K.Z. Developmental Studies and Clinical Application of Mismatch Negativity: Problems and Prospects. *Ear Hear* **1995**, *16*, 105–117. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Ceponiene, R.; Service, E.; Kurjenluoma, S.; Cheour, M.; Naatanen, R. Children's performance on pseudoword repetition depends on auditory trace quality: Evidence from event-related potentials. *Dev. Psychol.* **1999**, *35*, 709–720. [\[CrossRef\]](#)

60. Gomot, M.; Giard, M.H.; Roux, S.; Barthelemy, C.; Bruneau, N. Maturation of frontal and temporal components of mismatch negativity (MMN) in children. *Neuroreport* **2000**, *11*, 3109–3112. [[CrossRef](#)] [[PubMed](#)]
61. Shafer, V.L.; Morr, M.L.; Kreuzer, J.A.; Kurtzberg, D. Maturation of mismatch negativity in school-age children. *Ear Hear* **2000**, *21*, 242–251. [[CrossRef](#)]
62. Uwer, R.; von Suchodoletz, W. Stability of mismatch negativities in children. *Clin. Neurophysiol.* **2000**, *111*, 45–52. [[CrossRef](#)] [[PubMed](#)]
63. Gumenyuk, V.; Korzyukov, O.; Alho, K.; Winkler, I.; Paavilainen, P.; Naatanen, R. Electric brain responses indicate preattentive processing of abstract acoustic regularities in children. *Neuroreport* **2003**, *14*, 1411–1415. [[CrossRef](#)] [[PubMed](#)]
64. Shestakova, A.; Huottilainen, M.; Ceponiene, R.; Cheour, M. Event-related potentials associated with second language learning in children. *Clin. Neurophysiol.* **2003**, *114*, 1507–1512. [[CrossRef](#)] [[PubMed](#)]
65. Huttunen, T.; Halonen, A.; Kaartinen, J.; Lyytinen, H. Does mismatch negativity show differences in reading-disabled children compared to normal children and children with attention deficit? *Dev. Neuropsychol.* **2007**, *31*, 453–470. [[CrossRef](#)] [[PubMed](#)]
66. Dehaenelambertz, G.; Dehaene, S. Speed and cerebral correlates of syllable discrimination in infants. *Nature* **1994**, *370*, 292–295. [[CrossRef](#)] [[PubMed](#)]
67. Morr, M.L.; Shafer, V.L.; Kreuzer, J.A.; Kurtzberg, D. Maturation of mismatch negativity in typically developing infants and preschool children. *Ear Hear* **2002**, *23*, 118–136. [[CrossRef](#)] [[PubMed](#)]
68. Cheour, M.; Leppanen, P.H.T.; Kraus, N. Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clin. Neurophysiol.* **2000**, *111*, 4–16. [[CrossRef](#)] [[PubMed](#)]
69. Themaz, L.; Lippus, P.; Padrik, M.; Kask, L.; Kreegipuu, K. Maturation of the mismatch response in pre-school children: Systematic literature review and meta-analysis. *Neurosci. Biobehav. Rev.* **2023**, *153*, 105366. [[CrossRef](#)] [[PubMed](#)]
70. Yang, W.Q.; Xiao, R.; Liang, D.D. Lexical tone perception mechanism in 2- to 4-year-old Mandarin-speaking children in the pre-attention stage. *Acta Psychol. Sin.* **2020**, *52*, 730–741. [[CrossRef](#)]
71. Yu, K.K.; Li, L.; Chen, Y.; Zhou, Y.C.; Wang, R.M.; Zhang, Y.; Li, P. Effects of native language experience on Mandarin lexical tone processing in proficient second language learners. *Psychophysiology* **2019**, *56*, e13448. [[CrossRef](#)]
72. Yu, L.D.; Huang, D.; Wang, S.P.; Zhang, Y. Reduced Neural Specialization for Word-level Linguistic Prosody in Children with Autism. *J. Autism Dev. Disord.* **2023**, *53*, 4351–4367. [[CrossRef](#)]
73. Zhang, H.; Ma, W.; Ding, H.W.; Peng, G.; Zhang, Y. Phonological Awareness and Working Memory in Mandarin-Speaking Preschool-Aged Children with Cochlear Implants. *J. Speech Lang. Hear. Res.* **2022**, *65*, 4485–4497. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.