An Electrophysiological Investigation of Linguistic Pitch Processing in Tonal-language-speaking Children with Autism

A Dissertation
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Luodi Yu

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Yang Zhang, PhD

September 2018

Acknowledgements

I would like to thank my advisor Yang Zhang. It has been a great honor to be his student. His dedication and diligence have always been motivational to me through the highs and lows of my PhD. He has educated me how to be a solid researcher in many ways. I am grateful for all his generous contributions of time, ideas, and funding to make my PhD research productive and rewarding.

I want to thank my committee members, Drs. Benjamin Munson, Sheri Stronach, and Jason Wolff, for their guidance, encouragement, and patience. I appreciate the generous offering of their time to support me through the preliminary exams and other milestones, and sharing their knowledge on phonological development (BM), neurodevelopment (JW) and cross-cultural issues in autism (SS) to help me improve my work.

The Zhanglab members have been a source of friendship. I would like to thank Tess Koerner, Chieh Kao for their support and companionship that made my PhD filled with fond memories. Additional thanks to Aparna Rao for her kindness and caring.

Special thanks to Dr. Suiping Wang at South China Normal University, Dan Huang at Guangzhou Rehabilitation & Research Center for Children with ASD, and all study participants. This dissertation project would not have been possible without their enormous support. Thanks also goes to the many students and colleagues in Guangzhou for their assistance.

I would like to offer my gratitude to the faculty and staff in the Department of Speech-Language-Hearing Sciences. They have been extremely helpful and welcoming since my first day in the PhD program as a new international student. I am also grateful for the world-class education and research facilities offered by the University of Minnesota. My academic skills have grown substantially from working with experts in diverse fields in CATSS, CMRR, CNBD, etc.

At last, I would like to thank my family and friends for their unconditional love and support. Particularly, I thank my parents for everything, friends for being on my side unconditionally, and Tim for being the best fiancé and the sweetest human being.

This dissertation project was made possible by funding from the Bryng Bryngelson Research Fund, the College of Liberal Art Brain Imaging Research Project award and the University of Minnesota Grand Challenges Research Grant to Yang Zhang. Research activities of this dissertation were also partially supported by grants from the Natural Science Foundation of China overseas collaboration grant (31728009) to Yang Zhang.

Abstract

Speech perception is a fundamental skill interfacing sound to meaning; however, systematic characterization of autism in relation to this issue is still lacking, presumably due to insufficient consideration of the language-specific nature of speech processing. Although nearly 70% of world languages are tonal, tonal language users have been significantly under-represented in autism research. An overview of the limited literature reveals that there is a trend of distinct patterns across different language users (i.e., tonal language vs. non-tonal language), indicating potentially disrupted neural specialization for linguistic structures in individuals with autism.

This dissertation examined the rapid cortical processing of pitch patterns varying in linguistic status in native Chinese school-age children with autism and age-matched typically developing (TD) peers using electroencephalography (EEG). The auditory stimuli were nonsense speech and nonspeech sounds presented in passive listening conditions. In comparison with the TD group, the autism group displayed neural timing issues at various levels of information processing as indicated by neural response latency. Moreover, the autism group displayed not only hyposensitivity for native vs. nonnative (or prototypical vs. non-prototypical) *difference* in the early information processing stage but also hypersensitivity in the later processing stage accompanied by diffusive scalp distribution with a rightward dominance. The results collectively support the idea of disrupted neural specialization for linguistic structures in autism. The findings underscore the proposition that autism is bound with auditory and phonological atypicalities in addition to the syndromic social and communication deficits, which have important

implications for requiring language-specific considerations in autism research and clinical practice.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ASD: Autism spectrum disorder

AEP: Auditory evoked potential

ABR: Auditory brainstem response

FFR: Frequency following response

EEG: Electroencephalography

ERP: Event-related potential

MEG: Magnetoencephalography

GFP: Global field power

MMN: Mismatch negativity

LNR: Late negative response

TD: Typically developing

DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition

ITPC: Inter-trial phase coherence

NVIQ: Nonverbal IQ

LME: Linear mixed-effects

ANOVA: Analysis of variance

CP: Categorical perception

VOT: Voice onset time

TMS: Transcranial magnetic stimulation

CV: Consonant-vowel

HFA: High-functioning autism

LFA: Low-functioning autism

LH: Left hemisphere

RH: Right hemisphere

NLNC: Native language neural commitment

T2: Tone 2

ISI: Inter-stimulus interval

ICA: Independent component analysis

F0: Fundamental frequency

Chapter 1: Introduction

1.1 Overview

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by persistent social communication deficits and restricted & repetitive behaviors (American Psychiatric Association, 2013). Unless specified, the word "autism" will be used to refer to ASD henceforth. Delayed speech and language development constitute one of the first symptoms of most infants who are later diagnosed with autism (Tager-Flusberg, Paul, & Lord, 2005; Zwaigenbaum et al., 2005). Moreover, language acquisition is one of the strongest predictors of the long-term outcomes of academic and social performance in children with autism (Gillberg, 1991; Gillespie-Lynch et al., 2012; Howlin, 2005). As a fundamental skill interfacing sound to meaning, speech perception has long been a topic of interest in autism research. However, systematic characterization of potential speech perception problems in autism is lacking.

Importantly, aberrant auditory processing has been identified as a key factor underlying the pathological speech perception and language processing in individuals with autism (Alcantara, Weisblatt, Moore, & Bolton, 2004; Čeponienė et al., 2003; Siegal & Blades, 2003). A large body of research in the past two decades was dedicated to clarifying this issue and refining theories of the complex speech and language profile of autism. Some researchers link autism with an inherently altered auditory system, and the manifestation of aberrant auditory processing is unbalanced depending on the acoustic dimension (spectral or temporal) (Alcantara, Cope, Cope, & Weisblatt, 2012; Groen et al., 2009) and stimulus complexity (Mottron, Dawson, Soulieres, Hubert, & Burack, 2006; Samson, Mottron, Jemel, Belin, & Ciocca, 2006). Others propose a domain-

specific (speech vs. nonspeech) viewpoint that speech learners with autism fail to engage or develop specialized networks for vocal processing and linguistic learning (Eyler, Pierce, & Courchesne, 2012; Lindell & Hudry, 2013; Sperdin & Schaer, 2016). Occasionally, some researchers take into account both the stimulus-specific aspect (spectral or temporal) and domain-specific aspect (speech or nonspeech), stating that neural specialization in autism is altered in a way that some attributes of sounds (e.g., pitch, vowel) are selectively over-processed and some (e.g., duration, syllable) are under-processed (Haesen, Boets, & Wagemans, 2011). However, little of the above mentioned work has factored in the language-specificity of speech perception, although the language-general to language-specific transition in early development has been a well-documented phenomenon in numerous studies of speech perception and production in neurotypical populations (Kuhl, 2010; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Munson, Edwards, & Beckman, 2012; Näätäneiv et al., 1997; Werker & Curtin, 2011; Werker & Tees, 1984).

The human auditory system transforms acoustic information into meaningful phonological units for speech understanding and does so in a language-specific manner. Tonal languages such as Mandarin Chinese employ syllabic pitch variation (i.e., lexical tone) to signal meaning differences; quantity languages such as Finnish employ vowel length contrast phonemically to signal semantic distinction. Typically developing (TD) infants show language-specific perceptual patterns in the first year of life. In general, from 6 to 12 months of age, infants' discrimination ability of nonnative phonemic contrasts declines, demonstrating a perceptual narrowing effect (Werker & Tees, 1984). In the meantime, their discrimination ability for native phonemic contrasts enhances

(Kuhl et al., 2006). Nontonal language learners demonstrate perceptual decline for lexical tone contrasts even before 6 months of age (Mattock, Molnar, Polka, & Burnham, 2008; Yeung, Chen, & Werker, 2013). One candidate brain mechanism for such perceptual re-organization is the neural commitment process. That is, neutrally committed architectures are developed to promote detection of auditory patterns in the native language; meanwhile, the specialized neural system may interfere with the processing of nonnative auditory patterns (Kuhl, 2010; Zhang & Wang, 2007). Disruptions in the auditory processes that are constrained by language experience and pathological conditions can have detrimental effects on one's ability to extract relevant information from speech sounds, causing life-long communication difficulties (Kujala, 2007; Ramus, 2014).

Research has provided rich knowledge on the development of lexical tone perception in TD children. In contrast with nontonal language learners who show a perceptual decline for lexical tones at 4-6 months of age, Mandarin- and Cantonese-acquiring infants maintain their discriminability for the same sounds, suggesting a primitive phonological representation of native lexical tone categories (Yeung et al., 2013). Beyond this initial stage of phonetic acquisition, nontonal language learners are expected to disregard syllabic pitch variation and separate it from word meanings, whereas tonal language learners have to incorporate these variations into word learning, aka the lexical integration of lexical tones (Singh & Fu, 2016). There is evidence that this integration takes place between 18~ 24 months of age, during which English-learning infants no longer treat nonnative lexical tone change as lexically relevant, whereas Chinese-learning infants continue to utilize it to guide their word learning (Singh, Hui, Chan, & Golinkoff, 2014). Such a language-specific shift strongly indicates a qualitative

change in the phonological representation of lexical tones in tonal language learners. Up to school age, Chinese-speaking children's phonological skill for lexical tones improves considerably with formal literacy instruction (Shu, Peng, & McBride-Chang, 2008). By the age of 10, Chinese-speaking children show a phonemic representation of lexical tones similar to adults as evidenced by behavioral categorical perception and electrophysiological response (e.g., mismatch negativity indexing auditory discrimination sensitivity; Zhang et al., 2012) .

Tonal languages constitute nearly 70% of languages spoken worldwide (Yip, 2002); however, tonal language users have been disproportionally under-represented in autism research. This disparity is directly relevant to the cross-linguistic/cultural validity of existing models and clinical instruments, due to potentially differential expression of autism symptomologies across languages and cultures (Caron, Schaaf, Benevides, & Gal, 2012; Cuccaro et al., 2007; Freeth, Sheppard, Ramachandran, & Milne, 2013; Koh & Milne, 2012; Stronach & Wetherby, 2017; Wakabayashi, Baron-Cohen, Wheelwright, & Tojo, 2006). For example, if speech perception in individuals with autism is constrained by an inherently altered auditory system with an inclination for over-representing pitch information while under-representing duration information (Haesen et al., 2011), one would predict speech perception of Chinese-speaking children with autism to be advantageous, as Chinese speech understanding relies on syllabic-level pitch contrasts. However, recent studies conducted with Mandarin Chinese-speaking children with autism do not seem to support this hypothesis. In particular, Chinese children with autism displayed reduced discriminatory sensitivity for linguistic tonal categories compared to TD controls as indicated by MMN, suggesting a disadvantage in lexical tone perception

(Wang, Wang, Fan, Huang, & Zhang, 2017; Yu et al., 2015). This result is inconsistent with existing language-general framework derived exclusively from nontonal language users. The cross-linguistic discrepancies suggest that speech perception impairment in autism might be language-specific to some extent and that language-nonspecific frameworks that only emphasize general auditory processing mechanisms may have limited explanatory power cross-linguistically. For tonal-language-learning children with autism, it is necessary to investigate mechanisms of how linguistic pitch information is represented in their brain, which could be different from processing pitch information in nonspeech stimuli. Answers to this question will help pinpoint the potential neurophysiological underpinnings of speech impairment for theory refinement and inform future neurodevelopmental investigation for children with autism.

1.2 Electrophysiological Methods

Electroencephalography (EEG) is a widely used tool in clinical and research settings to measure neural activity in the brain. By placing sensors on the participant's scalp, EEG records summed postsynaptic potentials generated by neuronal populations non-invasively (Luck, 2014). EEG provides excellent temporal resolution up to a sub-millisecond scale, and is therefore capable of capturing the time and frequency varying features of rapid neural processing of auditory stimulation (Bastiaansen, Mazaheri, & Jensen, 2012). The current research employed waveform analysis of event-related potential as well as time-frequency analysis of EEG oscillations averaged from responses to the stimulus trials. This section will introduce briefly these two approaches.

1.2.1 Event-related Potentials

The term event-related potential (ERP) is used to describe acceptable EEG responses relative to a reference event averaged across trials. In many cases, the reference event is the onset of a stimulation, for example, a sound or an image (Luck, 2014). Since the voltage of ERP elicited by a single stimulus or trial is too small to be differentiated from the ongoing background EEG, ERP of many repeating trials must be averaged in order to obtain a visible neural process. The assumption behind this is that the evoked brain activity has a fixed latency across trials such that irrelevant random brain activities can be cancelled out by averaging. The resulting ERP waveform has distinctive deflections with positive and negative polarities which are called ERP components. The naming of an ERP component is conventionally determined by its polarity and peak latency, for example, P1 is a component with positive polarity peaking around 100 ms after stimulus onset. These time-locked components allow us to investigate different stages of information processing affected by stimulus manipulation and pathological conditions such as autism. In speech perception research, ERP has been successfully used to delineate language-independent auditory processes and language-dependent phonological processes (Kaan, Barkley, Bao, & Wayland, 2008; Phillips, 2001; Swaab, Ledoux, Camblin, & Boudewyn, 2012).

The morphology of auditory ERP (or AEP) has a protracted maturational course extending into adolescence (Čeponienė, Rinne, & Näätänen, 2002; C. W. Ponton, Eggermont, Kwong, & Don, 2000; Sharma, Kraus, McGee, & Nicol, 1997). In adults, auditory stimuli elicit the obligatory P1(or P50 in some literature)-N1-P2 complex, whereas in children, P1-N2 (or N250 in some literature) are the most prominent and reliable AEP components (Čeponienė, Cheour, & Näätänen, 1998; C. W. Ponton et al.,

2000; Sharma et al., 1997). Adult-like N1 in children is found to be present only with slow stimulus presentation (Čeponienė et al., 1998; Čeponienė, Rinne, et al., 2002).

The obligatory AEP components are characterized by their polarity and topographical distributions across the scalp and can be quantified by peak amplitude and latency. AEP latency can index the efficiency of neural transmission in the processing of auditory stimuli (Näätänen & Picton, 1987). In particular, P1 latency is considered to mark the time delay of thalamocortical neural transmission (Eggermont, Ponton, Don, Waring, & Kwong, 1997). TD children's P1 amplitude increases with greater acoustic complexity (Čeponienė et al., 2001), and is affected by arousal and sound onset features (Hillel Pratt, 2012). Following P1 is a large negative deflection N2 peaking around 250 ms (or N250) (Čeponienė, Rinne, et al., 2002). With faster stimulus presentation (i.e., ISI < 1 s), the fronto-centrally distributed child N2 partially overlaps with developmentally emerging N1 (Čeponienė et al., 1998; Čeponienė, Rinne, et al., 2002). Studies with typical adults have demonstrated that N1-P2 components can reflect neural coding of characteristic acoustic features of various consonant and vowel categories (Digester, Wohlberedt, & Hoppe, 2009; Sharma & Dorman, 1999; Tavabi, Obleser, Dobel, & Pantev, 2007; Zaehle, Jancke, & Meyer, 2007; Zhang, Kuhl, Imada, Kotani, & Tohkura, 2005). Unlike P1 which reflects primarily low-level sensory detection, the auditory N2 is considered to reflect higher-level sensory integration and phonetic perception (Almeqbel & McMahon, 2015; Čeponienė, Cummings, Wulfeck, Ballantyne, & Townsend, 2009; Čeponienė et al., 2001; Čeponienė, Torki, Alku, Koyama, & Townsend, 2008; Karhu et al., 1997). N2 latency in children has been shown to indicate the neural coding of voice-

onset-time (VOT) that defines stop consonant categories (Almeqbel & McMahon, 2015; King et al., 2008).

As opposed to the obligatory AEP components, late ERP responses that typically occur after 200 ms are thought to indicate endogenous neural activities related to higher-order cognitive processes. For example, N400 has been shown to differ between semantically nonsense words as opposed to familiar words in adults (Mills, Coffey-Corina, & Neville, 1993). Late negative responses have been found to index processing of word meaning and higher-order phonological structures in children and adults (Friedrich & Friederici, 2005; Mills, CoffeyCorina, & Neville, 1997; Mills, Plunkett, Prat, & Schafer, 2005). Moreover, hemispheric asymmetry in the ERP topographic distribution has also been shown to reflect neural processing of phonotactic organization and linguistic prosody by children and adults (Friedrich & Friederici, 2005; Mills et al., 2004).

1.2.2 Oscillation Activities

One important aspect of the EEG signal is the rhythmic activity or oscillation at various frequencies. Oscillation is a product of synchronized local field potentials generated across neural populations. ERP waveform alone that is averaged across stimulus trials provides little information about the oscillation dynamics of ongoing EEG because it only captures the time-locked and phase-locked activity in the EEG whereas the non-phase-locked oscillation frequency and phase information are disregarded from the ERP averaging process. In other words, if the oscillation phase is highly jittered across trials, the non-phase locked activities may cancel each other out, even though these activities might be of relevance to information processing.

It has been found that phase alignment of EEG oscillations in response to a stimulus, especially in theta (4~7 Hz) and alpha (8~12 Hz) bands, drives the generation of human evoked potentials (Edwards et al., 2009; Klimesch, Sauseng, Hanslmayr, Gruber, & Freunberger, 2007). Time-frequency analysis offers a way to characterize such neural synchrony or variability using the inter-trial phase coherence (ITPC) in frequencies of interest. Smaller ITPC values represent poorer consistency in the phase alignment of oscillations or larger amount of neural "jittering" across trials. Trial-by-trial phase alignment of ongoing EEG oscillations measured by ITPC can provide in-depth information about the neural coding of speech sounds and linguistic processing in TD individuals (Edwards et al., 2009; Klimesch et al., 2007; Koerner & Zhang, 2015; Koerner, Zhang, Nelson, Wang, & Zou, 2016) and in individuals with autism (Edgar, Fisk Iv, et al., 2015; Edgar, Khan, et al., 2015; Jochaut et al., 2015; Simon & Wallace, 2016). In particular, infants' theta activity has been shown to be modulated by linguistic experience (Radicevic, Vujovic, Jelicic, & Sovilj, 2008) and phonetic salience (Zhang et al., 2011). In adults, theta ITPC significantly predicts auditory N1-P2 amplitude to CV (consonant-vowel) syllables (Koerner & Zhang, 2015). Furthermore, theta activities are found responsible for syllabic-level speech encoding and believed to be crucial for speech parsing and understanding (Doelling, Arnal, Ghitza, & Poeppel, 2014; Giraud & Poeppel, 2012; Morillon et al., 2010; Peelle, Gross, & Davis, 2013).

1.3 Auditory Processing and Speech Perception in Autism Spectrum Disorder

In the literature, the majority of studies on this topic were conducted with nontonal language speakers (i.e., English, Finnish, French). Some papers did not provide information about the participants' language background. Efforts have also been made in

the investigation of auditory processing and speech perception with phonologically different language cohorts. This section will review findings from the nontonal language studies and recent tonal language studies.

1.3.1 Evidence from Nontonal Language Users

Auditory sensitivity in individuals with autism has been shown as highly unbalanced, mainly in the processing of spectral (tonal) vs. temporal (durational), and simple vs. complex stimuli. The estimated prevalence of absolute pitch ability is much higher among musically naïve people with autism than in musically naïve TD people (DePape, Hall, Tillmann, & Trainor, 2012; Heaton, 2003; Heaton, Hermelin, & Pring, 1998). A series of behavioral studies with high-functioning (HF) individuals with autism have demonstrated enhanced discrimination of simple pure tones (Bonnel et al., 2010; Bonnel et al., 2003), musical pitch intervals (Heaton, 2005), and pitch contours embedded in spoken sentences (Jarvinen-Pasley & Heaton, 2007; Jarvinen-Pasley, Pasley, & Heaton, 2008; Jarvinen-Pasley, Wallace, Ramus, Happe, & Heaton, 2008). Enhanced pitch perception is further verified using electrophysiological methods measuring mismatch negativity (MMN). MMN is obtained via oddball paradigm by subtracting the ERP to frequently occurring standard stimuli from the ERP to infrequently occurring deviant stimuli (i.e., oddballs) in a sound train. MMN is typically present around 200 ms in the difference ERP waveform. It reflects pre-attentive auditory change detection and is used as an indicator of auditory discrimination ability (Näätänen, Kujala, & Winkler, 2011). Mirroring the behavioral findings, children and adolescents with autism display increased MMN amplitude (Ferri et al., 2003) and shortened MMN latency (Gomot, Giard, Adrien, Barthelemy, & Bruneau, 2002) to frequency change of

pure tones. Additionally, Finnish-speaking children with autism show enhanced MMN to pitch changes of both syllables and complex nonspeech sounds (Lepisto et al., 2008; Lepistö et al., 2005; Lepistö et al., 2006).

Meanwhile, deficits in the discrimination of auditory temporal features have been commonly observed. Behaviorally, studies have shown impaired sound duration judgment (Brodeur, Gordon Green, Flores, & Burack, 2014), and extended temporal binding window in people with autism (Foss-Feig et al., 2010; Kwakye, Foss-Feig, Cascio, Stone, & Wallace, 2011). These findings imply that reduced auditory temporal resolution may be associated with the disorder. Electrophysiological studies have reported diminished MMN to both nonspeech and vowel duration changes in Finnish-speaking children (Lepisto et al., 2008; Lepistö et al., 2005; Lepistö et al., 2006). Moreover, delayed MMN latency to vowel duration changes was also observed in Japanese-speaking adults with autism (Kasai et al., 2005). It is necessary to mention that both Finnish and Japanese are quantity languages that utilize vowel duration change as a phonemic cue. As a result, users of these languages rely on the categorization of vowel duration contrasts in their phonological system for successful speech perception. Therefore, deficits in the neural sensitivity to auditory temporal details may hinder the acquisition of relevant phonological structures in young children. Indeed, it has been found that toddlers with autism show reduced MMN response for syllable discrimination marked by formant transition difference, a subtle but critical temporal acoustic cue for consonant categories (Kuhl, Coffey-Corina, Padden, & Dawson, 2005).

Acoustic complexity is also thought to play a role in the expression of auditory atypicality in autism. The neural complexity hypothesis ascribed atypical auditory

processing in autism to disrupted neural hierarchy. That is, compared to TD controls, individuals with autism tend to perform better with spectro-temporally simple sounds but more poorly with complex sounds or tasks (Samson, Mottron, Jemel, Belin, & Ciocca, 2006). The presence of temporal details in sounds can further complicate the auditory perception in autism. Behavioral studies have shown that speech-in-noise performance of HF individuals with autism drops significantly whenever temporal dips were present in the background noise, suggesting impaired "dip listening" ability (Alcantara et al., 2004). This finding is supported by another speech-in-noise study where HF children with autism showed less improvement from temporal dips in the background noise in comparison with the TD controls (Groen et al., 2009). Functional imaging data showed that the primary auditory cortex of adolescents and adults with autism was involved in the processing of temporally complex sounds, whereas nonprimary auditory cortex was recruited for the same task in the TD controls; however, this group difference was not present for spectrally complex sounds (Samson et al., 2011). These studies suggest difficulties in coping with acoustic complexity and spectro-temporal integration in individuals with autism.

Another important aspect of speech acquisition and language development is hemispheric lateralization for the processing of speech sounds and linguistic content. Although at first the right hemisphere is preferred for processing sub-lexical phonetic information before any related linguistic systemization is acquired (Mills et al., 1997; Seery, Vogel-Farley, Tager-Flusberg, & Nelson, 2013), leftward migration of cortical activity in 2nd year of life is thought to mark a qualitative change in word experience and linguistic sophistication (Mills, Conboy, & Paton, 2005). Nonetheless, it is important *not*

to consider this as the endpoint of the cortical organization for speech. There has been strong evidence of bilateral activation for phonetic processing without apparent asymmetry in the adult brain, while the higher-order linguistic process is left lateralized (Hickok & Poeppel, 2007). Moreover, the two hemispheres differ in computational properties such that asymmetry may occur depending on temporal and spectral properties of the auditory stimuli. That is, fast-changing input and temporal information (e.g., duration) drive left hemispheric activity, and slow-changing input and spectral information (e.g., pitch) are preferably handled by the right hemisphere (Boemio, Fromm, Braun, & Poeppel, 2005; Schonwiesner, Rubsamen, & von Cramon, 2005; Zaehle, Wustenberg, Meyer, & Jancke, 2004).

Brain lateralization to speech sounds can be observed in TD toddlers at 12-36 months that natural speech predominantly activated left hemisphere, whereas toddlers with autism did not show such lateralization (Eyler et al., 2012). Others reported rightward asymmetry in ERP responses to words in young children with autism (Coffey-Corina, Padden, & Kuhl, 2008; Kuhl et al., 2013). Taken together, lack of hemispheric lateralization or even rightward asymmetry during development might be underlying a broad range of speech and language impairments in individuals with autism (Haesen et al., 2011; Lindell & Hudry, 2013)

1.3.2 Evidence from Tonal Language Users

As mentioned earlier, understanding auditory processing and speech perception in autism would benefit from investigations on language-specific phonological features. Evidence from nontonal language speaking individuals with autism suggests a domain-general pattern whereby atypical temporal/spectral processing seems to correspond with

atypicality in the perception of tonal and durational speech stimuli. However, as each language employs a unique set of phonological features, speech perception deficits in autism may take different forms depending on the sound category and language. Speech acquisition in TD listeners involves an implicit neural commitment process to map out the phonological system of their native language by perceptually 'tuning out' irrelevant acoustic information, a process termed "native language neural commitment" or NLNC (Kuhl, 2004). If something went wrong in the early development of such neural specialization for native linguistic structures, individuals with autism who have a different perceptual weighting system of spectral and temporal cues would have deficits in representing the phonological categories in their native language.

Motivated by the phonological roles of pitch and duration in different languages, recent studies with Chinese children with autism have provided new information to the auditory and speech processing in autism. In a recent study, Chinese-speaking children with autism showed increased MMN amplitude to pitch changes in nonspeech stimuli but diminished responses to lexical tone contrast of speech sounds (Yu et al., 2015). One implication of this result is that hypersensitivity to spectral details may hinder the proper acquisition of tonal categories in tonal language speaking children with autism. To verify this hypothesis, a follow-up study specifically tested categorical perception (CP) of the lexical tones in Chinese-speaking children with autism. Greater mismatch response to between-category lexical tone differences than that to within-category changes would indicate CP. The results showed reduced CP for lexical tone in the speech condition in the autism group but typical-like CP in the nonspeech harmonic tone condition. Moreover, the lack of CP for speech in the autism group was driven by heightened

sensitivity to within-category pairs (Wang et al., 2017). The co-existence of 'pitch superiority' for nonspeech stimuli and language-specific pitch processing deficit in Chinese-speaking children with autism is in support of a NLNC abnormality for the higher-order phonological representation of lexical tones (Wang et al., 2017; You, Serniclaes, Rider, & Chabane, 2017; Yu et al., 2015).

These findings suggest that the bottom-up or cue-specific hypersensitive processing style in autism might be counter-effective or even detrimental in a linguistic environment. The supporting evidence also comes from an intonation study. The study showed that Chinese-speaking HF individuals with autism did poorer at speech intonation identification compared with the TD controls, despite having enhanced pitch perception of melodic contours (Jiang et al., 2015).

Vowel duration change can indicate word meaning change in Finnish, but duration does not play a phonemic role in the Chinese language. Using similar MMN method, a recent study has found that Chinese children with autism displayed duration discrimination deficit only for pure tones but not for vowels (Huang et al., 2018). This finding is in contrast with the domain-general duration deficit observed in the Finnish and English studies. Together, it appears that impaired discrimination exists in detecting duration differences involving a phonemic contrast (Finnish vowels, English formant transition) but not nonphonemic duration contrast (Chinese vowels), which again suggest an altered perceptual weighting system.

Based on the above evidence, speech perception impairment in autism can be not only feature-specific and domain-specific but also language-specific. Furthermore, individuals with autism tend to display heightened perceptual skills associated with right

hemisphere function and impaired perceptual skills associated with left hemisphere function, which is reflected in their enhanced auditory spectral processing (RH) but impaired temporal processing (LH) (Groen et al., 2009; Haesen et al., 2011; Huang et al., 2018) as well as enhanced non-linguistic pitch processing (RH) but diminished linguistic pitch processing (LH) (Jiang, Liu, Wan, & Jiang, 2015; Wang et al., 2017; Yu et al., 2015).

1.4 Research Questions

The language-specific speech perception atypicality and altered hemispheric asymmetry in autism point to a potentially disrupted neural specialization for native language. One way to characterize neural specialization for linguistic structures is to compare neural responses to native vs. nonnative speech, which allows the examination of bottom-up acoustic and top-down linguistic processes at the same time. So far, few studies have examined speech perception from a direct cross-linguistic (native vs. nonnative) approach, with the exception of a recent EEG study looking at the perceptual narrowing effect in phonetic learning with at-risk infants for autism (Seery et al., 2013). But this study provided little evidence to the initial delay of phonetic learning in the high-risk group. However, the results did not negate the possibility of disrupted neural specialization later in life, given the fact that mechanisms of joint attention as a prerequisite for phonetic learning is not substantially affected until first birthday (E. J. Jones, Gliga, Bedford, Charman, & Johnson, 2014; Seery et al., 2013). Given that the phonological space of native language undergoes continuous refinement through school-age (Munson et al., 2011; Shu et al., 2008; Singh & Fu, 2016), data from older children

with more stabilized phonological representation are needed to pinpoint where the issues might be.

Although we have found some neurophysiological evidence of impaired discrimination of tonal categories in Chinese learners with autism, detailed information is still lacking as to how language-specific pitch patterns are represented in the brain affected by autism. The derived MMN measure provides information about the perceptual discrimination and auditory sensory memory (Näätänen et al., 2012); however, it does not capture the early sensory encoding such as encoding of the transient pitch onset characterizing Chinese lexical tones, or higher-order phonological integration and abstraction at later processing stages. Going beyond MMN is necessary to determine important questions of the relationship between audition and language in autism. That is, whether early cortical responses to lexical tones show language-experience-dependent attunement in children with autism, as reflected in the obligatory auditory responses in adult native listeners (Krishnan, Gandour, & Suresh, 2014). Moreover, very little is known on the neural representation of prosodic phonology conveyed by syllable-to-syllable pitch variations at the word level in individuals with autism who speak a tonal language, despite having intonation impairment (Jiang et al., 2015). The present research aimed at examining neural specialization for the linguistically relevant pitch at both syllabic and supra-syllabic level in school-age Chinese-speaking children with autism. Electrophysiological measures of EEG oscillation synchrony in addition to ERP measures were employed to provide further information about atypical auditory response and its time course. The reported work in this dissertation represented four original journal manuscript submissions in verbatim, including one that has been published.

Study 1 (Chapter 2, reprinted from Yu et al, 2018, Clin Neurophysiol)

Research question:

- This study aimed to provide evidence for the underlying role of neural phase locking in atypical auditory ERPs in children with autism, and to provide a feasibility test of time-frequency analysis for this dissertation research. The research question was: can trial-by-trial neural variability in the phase alignment of cortical oscillations measured by inter-trial phase coherence (ITPC) account for atypical patterns in the basic auditory ERP responses in children with autism compared with TD peers?

Expected outcome:

- As reduced synchrony of neuronal activities in autism may lead to overall reduced evoked power (Simon & Wallace, 2016), we expected that reduced or increased theta ITPC within corresponding windows of ERP components (Edwards et al., 2009; Klimesch et al., 2007), would be predictive of ERP amplitude reduction or increment in children with autism.

Study 2 (Chapter 3)

Research question:

- How does linguistically relevant pitch trajectory affect ERP and ITPC in response to syllabic-level pitch pattern in Chinese-speaking children with autism, compared with TD peers?

Expected outcome:

- The effect of pitch trajectory (curvilinear vs. linear T2) was expected to be greater in the TD group compared to that in the autism group, indicating less language-specific attunement for linguistic pitch pattern in the children with autism.

Study 3 (Chapter 4)

Research question:

- This study examined normal adults to test the feasibility of using the native and nonnative supra-syllabic stimuli, as such cross-linguistic sounds had not been used previously in EEG studies. The research question was: can ERP and/or ITPC measures index neural processing of prosodic phonology conveyed by supra-syllabic pitch variations in Chinese-speaking TD adults?

Expected outcome:

- Late ERP responses but not early obligatory sensory ERPs, and the associated theta ITPC would index distinction of prosodic phonology (speech) but not prosodic acoustics (nonspeech hum), indicating higher-order abstraction and integration of linguistic structures beyond the language-nonspecific auditory analysis of complex sound.

Study 4 (Chapter 5)

Research question:

- How is prosodic phonology conveyed by supra-syllabic pitch variation represented in the brain of Chinese children with autism compared to the TD peers?

Expected outcome:

- The TD children may or may not show adult-like responses for prosodic phonology distinction. In comparison, the autism group was expected to show bilateral or rightward dominance in the distinction of native vs. nonnative prosodic phonology, indicating impaired or delayed cortical specialization for linguistic structures (Kuhl et al., 2013). Given the increased cue-specific auditory sensitivity for spectral

information associated with right hemisphere function in autism, the autism group may also display response differentiation for the prosodic acoustics (nonspeech hum stimuli) with pitch variations.

Chapter 2: Examining Auditory Evoked Potentials (AEPs) and Inter-trial Phase Coherence of Cortical Oscillations in Children with Autism

Section 1-5 are reprinted from:

Yu, L., Wang, S., Huang, D., Wu, X., & Zhang, Y. (2018). Role of inter-trial phase coherence in atypical auditory evoked potentials to speech and nonspeech stimuli in children with autism. *Clin Neurophysiol, 129*(7), 1374-1382.

2.1 Introduction

Sensory processing abnormality constitutes a core feature of autism spectrum disorders (American Psychiatric Association, 2013). Atypical auditory evoked potentials (AEPs) in autism have been reported with mixed findings. School-age children with autism often display prolonged neural response latency around 100 ms (N1, or M100 in magnetoencephalography) to tonal stimuli compared with typically developing (TD) children (Bruneau, Bonnet-Brilhault, Gomot, Adrien, & Barthelemy, 2003; Gandal et al., 2010; Oram Cardy, Flagg, Roberts, & Roberts, 2008; Roberts et al., 2010). However, the opposite pattern with the autism group showing earlier N1 response has also been found (Ferri et al., 2003). Attenuated P1 amplitude to pure tones (Donkers et al., 2015; Orekhova et al., 2008), speech and complex tonal stimuli (Lepistö et al., 2005), as well as attenuated N1c to clicks (Orekhova et al., 2009) have been discovered in children with autism. Moreover, some researchers observed attenuated N2 amplitude in children with autism in response to vowels (Whitehouse & Bishop, 2008) and complex tones (Lepistö et al., 2005), whereas others did not (Čeponienė et al., 2003; Gomot et al., 2011).

The mixed AEP findings suggest the variable nature of sensory perception in individuals with autism, which may partly result from methodological differences across studies. First, the sample age range varied from study to study with the widest being 11 years. As the AEP morphology in terms of amplitude and latency of the peak components

show age-dependent maturational changes, it is not surprising to see large variations in the results across individuals and studies. Second, differences in stimulus presentation rate across studies could contribute to the variations in the reported AEP results, as it is known to affect the presence and morphology of P1-N1 (or N2) in children (Čeponienė et al., 1998; Čeponienė, Rinne, et al., 2002; Cunningham, Nicol, Zecker, & Kraus, 2000). Third, atypical sensory perception in autism may show different hyper- or hypo- sensitive patterns depending on the category of auditory stimuli (i.e., whether the stimulus is speech or nonspeech) (Huang et al., 2018; Wang et al., 2017; Yu et al., 2017), and it is still unclear how these atypical sensory patterns as assessed by the AEP responses may change over time in child development. For instance, when tested with speech sounds, children with autism tend to display a categorical perception deficit with poorer discrimination of phonemic contrasts and enhanced sensitivity for within-category differences or allophonic variations (Wang et al., 2017; You et al., 2017). When tested with nonspeech stimuli, children with autism often exhibit enhanced discrimination of certain acoustic dimensions such as pitch (Lepistö et al., 2005; Wang et al., 2017; Yu et al., 2015) but reduced sensitivity for other dimensions such as sound duration (Huang et al., 2018; Lepistö et al., 2005). These findings point to the necessity of exercising caution when interpreting abnormal AEPs in autism, as they may reflect different neural processes of acoustic, phonetic, and phonological analysis. Fourth, abnormal AEPs in autism are often intertwined with stimulus complexity. It is suggested that the spectro-temporal complexity in verbal stimuli may compromise the neural dynamics in individuals with autism (Samson et al., 2006). Moreover, the inclusion of ASD subgroups of autistic disorder, Asperger's Syndrome and individuals with potentially only

subclinical symptomology may further introduce heterogeneity to the sample (Smith, Reichow, & Volkmar, 2015). The differences in subject characteristics, stimulus property, and experimental task add complexity to the synergistic interpretation of the existing AEP findings across the autism studies.

Among the frequently studied AEP components, P1 and N2 are the most prominent and reliable responses in childhood through adolescence (Čeponienė et al., 1998; C. W. Ponton et al., 2000; Sharma et al., 1997). In TD children, the auditory P1 is a stimulus-driven response peaking at around 100 ms post-stimulus. P1 latency is thought to mark the time delay of thalamocortical neural transmission in the auditory pathway (Eggermont et al., 1997). TD children's P1 amplitude increases with greater acoustic complexity (Čeponienė et al., 2001), and is affected by arousal and sound onset features (Hillel Pratt, 2012). Following the P1 is a large negative deflection N2 peaking around 250 ms (Čeponienė, Rinne, et al., 2002). With faster stimulus presentation (i.e., inter-stimulus interval < 1 s), the fronto-centrally distributed N2 in children partially overlaps with the developmentally emerging adult-like N1 (Čeponienė et al., 1998; Čeponienė, Rinne, et al., 2002). Unlike the P1 which reflects primarily low-level sensory detection, the auditory N2 is considered reflecting higher-level sensory integration and phonetic perception (Čeponienė et al., 2009; Čeponienė et al., 2001; Čeponienė et al., 2008; Karhu et al., 1997). It has been shown that N2 amplitude increases with repetition of identical sound, which might be instrumental in forming auditory-learning induced neural representations of sounds (Fujioka, Ross, Kakigi, Pantev, & Trainor, 2006; Karhu et al., 1997).

Although the AEP results can provide insights on the neural basis of variable sensory perception in autism in terms of amplitude and latency measures, the quantification of AEP components focuses solely on the time-domain information averaged across trials while discarding the trial-by-trial variations in the time-frequency domain. It has been found that phase alignment of EEG oscillations in response to a stimulus, especially in theta (4~7 Hz) and alpha (8~12 Hz) band, drives the generation of human evoked potentials (Edwards et al., 2009; Klimesch et al., 2007). Such evoked neural synchrony or phase-locking can be computed as inter-trial phase coherence (ITPC) in frequencies of interest. Smaller ITPC values represent poorer consistency in the phase alignment of oscillations or larger amount of neural "jittering" across trials.

One prevailing theory on the neurophysiology of autism is the increased variability in cortical responses to complex sensory input (Haigh, 2018), which can be tested through the inter-trial oscillatory synchrony. Indeed, recent data have demonstrated reduced task-related ITPC across multiple frequency bands in children (Edgar, Fisk Iv, et al., 2015; Edgar, Khan, et al., 2015; Gandal et al., 2010) and adults (Buard, Rogers, Hepburn, Kronberg, & Rojas, 2013; Jochaut et al., 2015; Milne, 2011; Sun et al., 2012; van Noordt et al., 2017) with autism. Critically, such jittered oscillation rhythm may result in an overall reduction in evoked response power in the EEG signal (Simon & Wallace, 2016). In this regard, one viable question to ask is whether attenuated AEP amplitude in autism obtained from averaging over EEG trials can be explained by increased inter-trial variability in the phase alignment of oscillations for both speech and nonspeech sounds. Answer to this question can provide insightful information regarding the neural basis of altered auditory processing.

The current study aimed to characterize atypicality in AEPs to speech and nonspeech stimuli in children with autism in the light of oscillatory variability. We focused on neural oscillations in the theta band, because theta synchrony not only is an important generator of AEP but also holds additional significance for speech and linguistic processing. In particular, infants' theta activity has been shown to be modulated by linguistic experience (Radicevic et al., 2008) and phonetic salience (Zhang et al., 2011). In adults, theta ITPC significantly predicts auditory N1-P2 amplitude to CV (consonant-vowel) syllables (Koerner & Zhang, 2015). Furthermore, theta activities are believed to be responsible for syllable-level speech encoding, which are crucial for successful speech comprehension (Doelling et al., 2014; Giraud & Poeppel, 2012; Morillon et al., 2010; Peelle et al., 2013). In the autism literature, however, the neurobehavioral consequences of theta rhythm in the processing of speech and nonspeech stimuli have rarely been tested. We were able to find only one study, in which adults and adolescents with autism showed weak synergistic theta activity for sentence-level speech processing in the auditory cortex (Jochaut et al., 2015). The relationships between trial-to-trial variations in theta activities and the AEP components for speech and nonspeech sounds remain unknown.

For the current investigation, we aimed to have a better control of some of the confounding factors to understand the AEP characteristics in children with autism with two primary objectives. First, we examined cortical processing of simple tones and complex speech sounds in school-age (8 to 13 years old) children with autism and age-matched TD children. Second, we were interested in exploring the relationships between the trial-to-trial variability/consistency of theta phase-locking measured by ITPC and the

P1-N2 amplitude within each group. We targeted school-age children in order to control potential age-related confounds due to maturational AEP changes. A short passive listening procedure with a fast stimulus presentation was used to elicit auditory P1-N2 components for both speech and nonspeech stimuli in these children. Amplitude/latency of P1 and N2 and the associated ITPCs were analyzed as a function of subject group and stimulus type.

We aimed to test two main hypotheses. First, if children with autism exhibit atypical sound onset feature detection and subsequent higher-level sound processing, we may observe atypical AEP amplitude and theta ITPC in the autism group at the processing stages of both P1 and N2 components compared with the TDs. Moreover, as deficiency in perceptual learning of auditory patterns in autism is considered to be speech-specific (Kujala, Lepistö, & Näätänen, 2013), we expect to find the group difference to be more pronounced in the word condition than in the pure tone condition. Second, as reduced synchrony of neuronal oscillations in autism may lead to overall reduced evoked power (Simon & Wallace, 2016), we hypothesized that reduced theta ITPC might be significantly correlated with reduced AEP amplitude in the group of autism.

2.2 Methods

2.2.1 Participants

Fifteen children with autism (14 boys and 1 girl, age mean = 9.6 years, *SD* = 1.7 years, range 8~12.4 years) were recruited from a local school for children with autism following recruitment and screening protocols established in our previous autism studies (Huang et al., 2018; Wang et al., 2017; Yu et al., 2015). The numbers of children with

autism included in data analysis were 15 for the pure tone condition and 14 for the word condition as one boy was not able to complete EEG recording for the word condition. The diagnoses were established by pediatricians according to the DSM-4 criteria for Autistic Disorder (American Psychiatric Association, 1994). As the Autism Diagnostic Observation Schedule (ADOS) (Lord, Rutter, DiLavore, & Risi, 2001) has not been officially validated and adopted in mainland China, we confirmed the diagnoses using the Chinese version of the Gilliam Autism Rating Scale–Second Edition (GARS-2) (Gilliam, 2006). Sixteen age-matched TD controls (13 boys and 3 girls, age mean = 9.8 years, *SD* = 1.4 years, range 7.8 – 12.9 years) were recruited from a local elementary school. One boy and one girl were excluded from analysis for the word condition due to noisy EEG signal, resulting in 16 children for the pure tone condition and 14 for the word condition in the TD group.

All participants were native speakers of Mandarin Chinese. They were screened for hearing loss using pure tone audiometry and met the criteria for normal hearing. All children in the autism group were verbal with limited language ability. Specifically, eight out of 15 had documented delayed onset of speech measured by their use of two-word utterances (this information was unavailable for 3 children); the average verbal IQ in the autism group was 64 (*N* = 13, information unavailable for 2 children, *SD* = 19, range 45 – 96) measured by the Wechsler Intelligence Scale for Children (WISC-IV) (Wechsler, 2003). All participants were native Mandarin speakers. None of the children had a known or diagnosed genetic, mental, or additional neurological condition, and were unmedicated at the time of the study. Nonverbal IQ (NVIQ) of each child was measured using the Raven's Standard Progressive Matrices Test (Raven & Court, 1998). The autism group

scored lower (mean = 88, *SD* = 19) compared to the TD group (mean = 106, *SD* = 15) ($t(29) = 3.33$, $p < .01$). The lower nonverbal scores in the autism group were expected and consistent with the reported IQ profiles in the literature (Dawson, Soulieres, Gernsbacher, & Mottron, 2007). Informed consent was obtained from each child's parent following a protocol approved by the local institutional review board.

2.2.2 Stimuli and Procedure

The pure tone was a 216 Hz sinusoidal wave. The word /bai2/ ("white" in Chinese) was uttered by a female talker, recorded using Neundo 4 (Steinberg Media Technologies, Germany). The sounds were edited to have a duration of 350 ms including 10 ms rise and fall time. The sound intensity was normalized to be equal for the two stimuli. The sound editing was completed using Praat (Boersma & Weenink, 2014) and Goldwave (http://www.goldwave.com). The two stimulus conditions were presented in two separate blocks with 500 ms inter-stimulus interval (ISI). Each block contained 500 stimuli. The stimuli were presented via AKG K518 headphones at approximately 70 dB SPL. The participants were seated and asked to watch a muted self-chosen cartoon and to ignore the auditory stimuli during the experiment.

2.2.3 EEG Recording and Data Analysis

Continuous EEG data were recorded with a 32-channel BrainAmps DC amplifier system (Brain Products, Germany). The sampling rate was 500 Hz. The left mastoid and AFz were used as the reference and ground, respectively. Ocular activities were monitored with electrodes placed below the right eye and the outer corner of the left eye. Electrode impedance was kept below 10 kΩ. Data analysis was performed using EEGLAB (Delorme & Makeig, 2004). In offline analysis, the continuous EEG data were

first visually inspected before preprocessing to identify overly noisy or flat segments and remove them from analysis. The data were high-pass filtered with 0.5 Hz cut-off. Ocular and muscle artifacts were removed using independent component analysis (ICA) for AEP waveform analysis. Epochs with 800 ms length were extracted including a 200 ms pre-stimulus baseline. Trials with instantaneous values exceeding ±100 μV were rejected.

Waveform analysis of the P1 and N2 components were performed with data band-pass filtered at 0.5-30 Hz. Based on the grand mean waveforms, P1 and N2 peaks were searched within post-stimulus windows of 70-150 and 200-350 ms, respectively. Mean amplitudes of P1 and N2 were computed for a 20 ms window around the peak of each participant. The Fz electrode was used for statistical analysis based on the topographical distribution.

Trial-by-trial time-frequency analysis was carried out in EEGLAB (Delorme et al. , 2004). Inter-trial phase coherence (ITPC) in the theta band (4~7 Hz) was computed using the "newtimef" function: $ITPC(f,t) = \frac{1}{n}\sum_{k=1}^{n}\frac{F_k(f,t)}{|F_k(f,t)|}$. In this function, $F_k(f,t)$ is the spectral estimate of trial k at frequency *f* and time *t* obtained using short-time Fourier transformation (STFT), and || represents the complex norm of trial k. The modified STFT (with Hanning tapers) in EEGLAB uses overlapping sliding windows that are adaptive to the target frequency bins (i.e., the time window decreases linearly as frequency increases), which is recommended to overcome limitations of conventional fixed window in estimating low frequency activities. The frequency range analyzed was 0.5 ~ 50 Hz. Zero-padding was applied to windows without sufficient number of sample points with a padratio of 16 with a frequency spacing of 0.5 Hz. ITPC value of a given frequency at a given time point can range from 0 to 1. Larger ITPC values indicate higher phase

consistency across trials, and smaller values indicate lower consistency or larger neural "jittering". For the calculation of theta ITPC, the ITPC data were first averaged across the frequencies within the theta range for further processing. Then the maximum theta ITPC values within the designated time windows of pre-stimulus baseline (-200~0), P1 (70~140) and N2 (150-250 ms) were identified for each participant for statistical analysis. Spectral power in the theta band was also computed for both the pre-stimulus baseline and the response portion of the epochs using the *spectopo* function in EEGLAB based on Welch's power spectral density estimate (oversampling ×8). Similar time-frequency analysis procedures were used in published studies (Koerner & Zhang, 2015; Koerner et al., 2016). The number of trials for analysis in the autism group were 332 (range 165-466) for the pure tone condition and 300 (172- 383) for the word condition. The numbers in the TD group were 309 (207- 396) for the pure tone condition and 334 (234- 403) for the word condition. These numbers did not differ between groups (pure tone, $t(29) = 1.01$, $p = .321$; word, $t(26) = -1.48$, $p = .15$).

To examine the main effects and interaction of group (autism vs. TD) and stimulus condition (pure tone vs. word), linear mixed effect (LME) regression was performed for each outcome measure, namely, amplitude and latency of the P1 and N2 components, theta ITPC within the baseline and corresponding windows of P1-N2, spectral theta power in the baseline and response. The mixed-effects model approach is considered superior to traditional methods statistical analysis such as repeated measures analysis of variance and Pearson's correlation (Gueorguieva & Krystal, 2004a; Koerner & Zhang, 2017). In each LME model, NVIQ, group, stimulus condition, and group*stimulus condition interaction were included as fixed effects, among which NVIQ

was regarded as a controlled covariate; participant was included as a random effect. Additionally, to examine the relationships between P1-N2 amplitude and ITPC within the corresponding windows in each group, LME model with theta ITPC as a predictor variable was fit for P1 and N2 amplitude for each group separately. In each LME model, theta ITPC was entered as a fixed effect, and subject as a random effect; NVIQ was included as a covariate and stimulus condition as a blocking variable. Two-tailed significance level was used for all the statistical analyses.

2.3 Results

2.3.1 AEP Amplitude and Latency

The LME regression model showed significant effects of group ($F(1,54) = 5.19$, $p < .05$) and stimulus condition ($F(1,54) = 16.13$, $p < .01$) on the P1 amplitude while controlling for NVIQ ($F(1,54) = 0.12$, $p = .731$), indicating larger P1 response in the autism group compared with the TD group and increased P1 response to word stimuli compared to pure tones in both groups (Fig. 1 & Tables 1 and 4). No significant group*condition interaction was found for the P1 amplitude ($F(1,54) = 0.46$, $p = .500$). For the P1 latency, there was a significant group effect ($F(1,54) = 6.93$, $p < .05$) while controlling for NVIQ ($F(1,54) = 0.38$, $p = .538$), indicating prolonged P1 latency in the autism group compared with the TD group. No condition effect ($F(1,54) = 0.24$, $p = .624$) or group*condition interaction ($F(1,54) = 0.69$, $p = .410$) on P1 latency was observed.

There was a nonsignificant trend of group effect on the N2 amplitude ($F(1,54) = 2.88$, $p = .096$) while controlling for the effect of NVIQ ($F(1,54) = 3.75$, $p = .058$) (Table 4). When the NVIQ was excluded from the regression, the group effect became significant ($F(1,55) = 6.04$, $p < .05$), suggesting a diminished N2 response in the autism

group compared with the TD group. No condition effect ($F(1,54) = 0.52$, $p = .474$) or group*condition interaction ($F(1,54) = 0.12$, $p = .915$) was found for the N2 amplitude. For the N2 latency, no significant effect was found (group, $F(1,54) = 1.17$, $p = .284$; condition, $F(1,54) = 0.00$, $p = .999$; group*condition interaction, $F(1,54) = 1.43$, $p = .237$; NVIQ ($F(1,54) = 0.01$, $p = .913$).

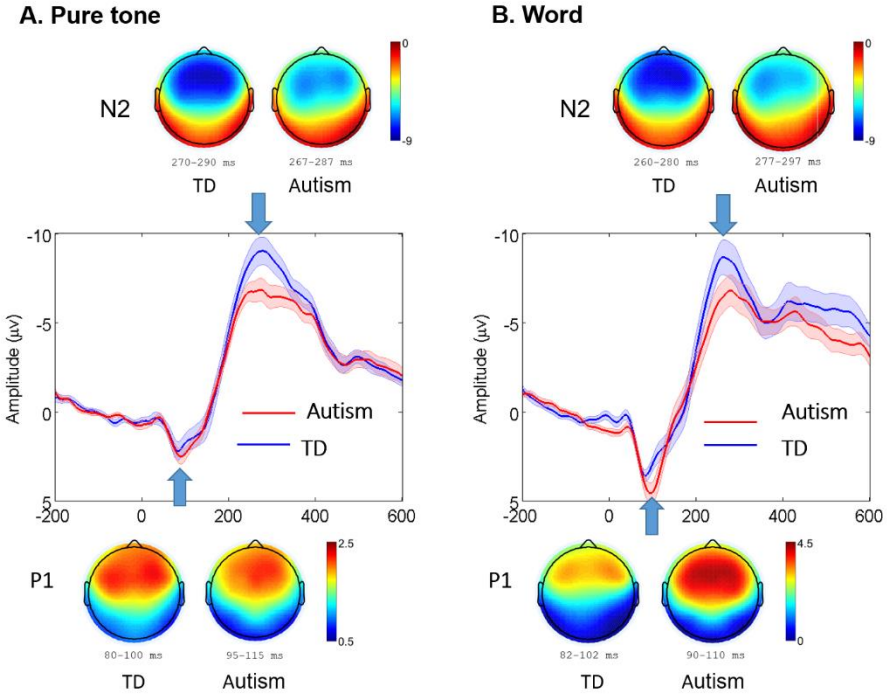

Figure 1. Grand averaged waveforms at Fz and topographical maps of mean amplitude in the P1 and N2 windows in the pure tone condition (A) and word condition (B). Shaded areas along the waveforms represent standard error at all sample points.

Table 1.

Mean and standard deviation (SD) of P1 and N2 amplitude (μV) and latency (ms) in the autism group and the TD group.

	Group	P1		N2	
		Amplitude	Latency	Amplitude	Latency
Pure tone	Autism	2.82 (1.06)	104 (19)	-7.54 (1.99)	277 (43)
	TD control	2.39 (2.00)	91 (14)	-9.50 (2.83)	279 (31)
Word	Autism	4.74 (1.83)	99 (12)	-7.09(3.30)	287 (24)
	TD control	3.75 (1.11)	92 (15)	-8.86 (3.45)	269 (16)

2.3.2 Inter-trial Phase Coherence and Spectral Power

Group and stimulus condition effects on theta ITPC were examined separately for the P1 and N2 windows (Fig. 2 & Tables 2 and 4). In the P1 window, the LME regression model revealed significant effects of group ($F(1,54) = 11.00$, $p < .01$) and condition ($F(1,54) = 5.69$, $p < .05$) on ITPC, controlling for NVIQ ($F(1,54) = 0.30$, $p = .587$). In particular, the autism group displayed increased ITPC in the theta band than the TD group for this early time window, and the word stimuli elicited greater theta ITPC than the pure tones for both groups. No interaction between group and condition ($F(1,54) = 0.43$, $p = .518$) was found for this window.

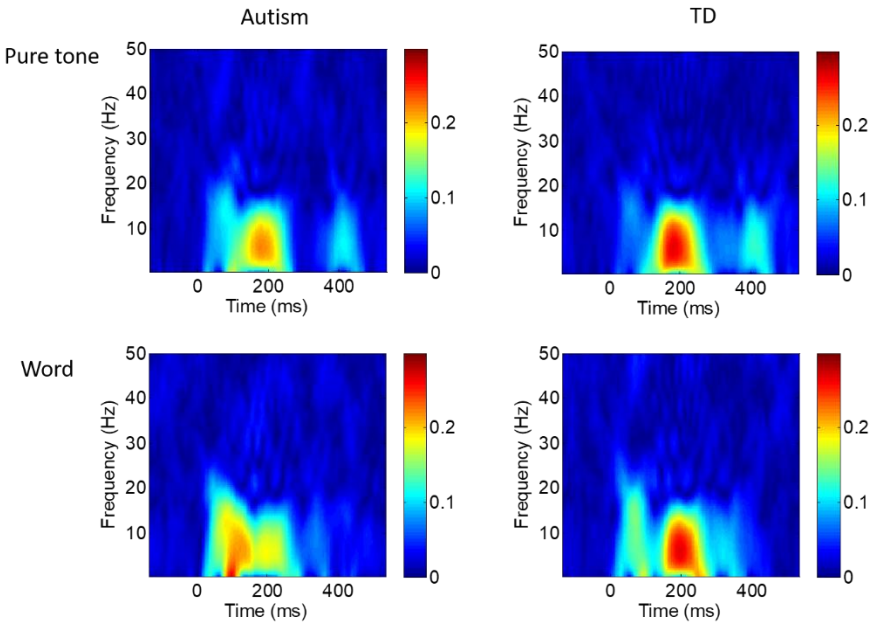

Figure 2. Time-frequency representations showing trial-to-trial phase-locking measured by ITPC in the pure tone condition and word condition.

Table 2.

Mean and standard deviation (SD) of theta ITPC in the pre-stimulus baseline and post-stimulus windows.

	Pure tone			Word		
	Baseline	P1	N2	Baseline	P1	N2
Autism	0.09(0.03)	0.19(0.08)	0.25(0.09)	0.08(0.03)	0.24(0.07)	0.22(0.09)
TD	0.09(0.02)	0.16(0.08)	0.29(0.10)	0.08(0.03)	0.19(0.05)	0.30(0.09)

Analysis for the N2 response revealed a nonsignificant trend of group effect ($F(1,54) = 3.23$, $p = .078$) on the ITPC while controlling for significant NVIQ ($F(1,54) = 4.38$, $p < .05$, $\beta = -0.03$) (Table 4). Similar to the amplitude result in this window, the group effect became significant ($F(1,55) = 6.88$, $p < .05$) after excluding NVIQ from the regression model, indicating reduced N2-associated theta ITPC in the autism group

compared to the TDs. No effect of condition ($F(1,54) = 0.21$, $p = .647$) or group*condition interaction ($F(1,54) = 0.52$, $p = .472$) was observed for this measure. The baseline ITPC before the stimulus onset did not differ between conditions ($F(1,54) = 0.39$, $p = .537$) or groups ($F(1,54) = 0.33$, $p = .568$), nor show any group*condition interaction ($F(1,54) = 0.23$, $p = .637$).

Theta power for the pre-stimulus baseline and the response was examined using similar LME regression (Tables 3 and 4). The results did not show any significant condition effect (baseline, $F(1,54) = 0.01$, $p = .933$; response, $F(1,54) = 0.21$, $p = .648$), group effect (baseline, $F(1,54) = 0.72$, $p = .399$; response, $F(1,54) = 0.01$, $p = .924$), or group*condition interaction (baseline, $F(1,54) = 0.05$, $p = 826$; response, $F(1,54) = 0.10$, $p = .759$).

Table 3.

Mean and standard deviation (SD) of theta spectral power (dB) of the pre-stimulus baseline and the response portion of the epochs on a trial-by-trial basis.

	Pure tone		Word	
	Baseline	Response	Baseline	Response
Autism	23.08(1.23)	11.08(1.18)	23.15(1.12)	11.29(1.01)
TD	23.30(0.81)	11.33(0.77)	23.26(0.64)	11.37(0.98)

Table 4.

Summary of F-statistics of main effects and interaction on the AEP, theta ITPC and power measures, controlling for NVIQ.

	Amp		Lat		ITPC			Spectral power	
	P1	N2	P1	N2	Baseline	P1	N2	Baseline	Response
Group	5.19*	2.88†	6.93*	1.17	0.33	11.00**	3.23†	0.72	0.02
Condition	16.13**	0.52	0.24	0.00	0.39	5.69*	0.21	0.02	0.21
Group* condition	0.46	0.12	0.69	1.43	0.23	0.43	0.52	0.05	0.10

** $p < .01$; * $p < .05$; † $p < .1$.

2.3.3 Relationships between AEP and ITPC

The LME regression revealed that theta ITPC in the autism group was a significant predictor of P1 amplitude ($F(1,25) = 4.39$, $p < .05$) and N2 amplitude ($F(1,25) = 6.28$, $p < .05$) across stimulus conditions regardless of NVIQ (Table 5). The regression coefficients (β) indicated positive relationships between these two types of measures, indicating stronger theta phase-locking within corresponding windows predicted larger P1-N2 response amplitude. The results were similar in the TD group that greater theta ITPC predicted larger amplitude of both P1 ($F(1,26) = 5.04$, $p < .05$) and N2 ($F(1,26) = 8.43$, $p < .01$) across stimulus conditions.

Table 5.

F-statistics and regression coefficients (β) indicating the relationships between AEP amplitude and theta ITPC in the autism group and the TD group.

Group	P1 amp		N2 amp	
	F	*β*	F	*β*
Autism	4.39*	7.78	6.28*	-13.81
TD	5.04*	10.47	8.43**	-16.71

** p < .05; ** p < .01*

2.4 Discussion

2.4.1 AEP findings

The waveform analysis of AEPs confirmed our first hypothesis regarding abnormality in the P1-N2 components. Our first observation is the prolonged P1 latency in the autism group compared with the TD group. The amount of delay we found here is consistent with reports by other researchers examining auditory M50 and M100 to pure tones, which was approximately 10% when compared with TD children (Edgar, Khan, et al., 2015; Gage, Siegel, Callen, & Roberts, 2003; Gandal et al., 2010; Roberts et al., 2014; Roberts et al., 2010). The delayed P1 response likely indicates lower neural transmission speed along the ascending auditory pathway (Eggermont et al., 1997). Previous work has demonstrated that TD children's early auditory response (M100) shortens with increasing age, and this trend corresponds with the developmental changes in thalamocortical white matter integrity (Roberts et al., 2009). However, age-related latency shortening in individuals with autism appears to be uncoupled with thalamocortical white matter properties (Roberts et al., 2013). This uncoupling might indicate other mechanisms such as synaptic transmission and maturation as contributing

factors to the latency lag in children with autism (Eggermont et al., 1997; Roberts et al., 2013). Interestingly, the P1 latency delay was observed across domains but the effect for words was not as prominent as in the pure tone condition (Table 1), which seemingly contradicts the fact that autism is often characterized with pronounced language impairment. We speculate that this result might be attributable to the robustness of stimulus onset. Unlike the pure tone, which rises sharply from silence to peak amplitude with a fixed slope at the same sound frequency, the onset of the CV syllable is not as uniformly defined as the consonant portion contains multiple spectral components and a much longer interval for the nonlinear consonant-to-vowel transition. As P1 marks the neural transmission time for sound onset detection rather than registration of fine-grain content aspect of sounds (Čeponienė, Alku, Westerfield, Torki, & Townsend, 2005; Eggermont et al., 1997), stimuli with more clearly defined onset acoustic features such as pure tones are expected to be more sensitive to show the between-group P1 latency differences than complex sounds that do not have the same robust onset.

The fact that words elicited larger P1 than pure tones in both groups of children indicates stronger neural activity evoked by the acoustically rich broadband signal. Interestingly, there was also a group effect that the autism group had overall increased P1 amplitude than the TD group. This finding appears to be inconsistent with previous work showing attenuated P1 to speech in school-age children with autism (Čeponienė et al., 2003; Lepistö et al., 2005; Whitehouse & Bishop, 2008). The inconsistency might arise from coding onset differences in the physical properties of speech stimuli. Our study used a naturally recorded CV syllable starting with a less well-defined acoustic transient, whereas the previous studies used computer-synthesized steady-state vowels. P1

amplitude in adults has been found to be modulated by duration of consonant-vowel transition (CVT) (Čeponienė et al., 2008). The larger P1 to words in autism in our study might suggest a hypersensitive reaction to the acoustic transient of sound onset compared to the age-matched TD children. As syllable-evoked P1 amplitude typically decreases from childhood to adolescence (Cunningham et al., 2000; Sharma et al., 1997), enlarged P1 might also reflect delayed maturation of the auditory system in processing acoustically complex sounds. Additionally, the overall larger P1 amplitude may suggest deficits in auditory habituation to the large amount of sound repetition, as P1 amplitude reduction to repetitive stimulus can index sensory gating and habituation (Grunwald et al., 2003), which has been found impaired in adults (J.S. Buchwald et al., 1992) and children with autism (Orekhova et al., 2008).

The N2 component in the autism group was attenuated across speech and nonspeech stimulus conditions. The lack of any speech-specific difference for N2 measures seems to suggest rather domain-general atypicality in this processing stage. According to the theory by Karhu et al. (1997), children's N2 might reflect auditory "sensitization" in building up neural representations of sound features. Thus, the attenuated N2 in the autism group may reflect some basic abnormalities in auditory learning of repeating sound pattern. However, in the context of speech sounds, N2 reflects complex processes involving not only acoustic processes but also phonetic analysis of fine-grain content aspect of sounds (Čeponienė et al., 2009; Čeponienė et al., 2008; Fujioka et al., 2006). In this perspective, the autism group' reduced N2 amplitude to words might indicate some phonological deficits beyond just auditory learning.

2.4.2 Trial-to-trial synchrony of theta oscillations

One primary finding of the current study is the increased ITPC in autism within the P1 window but reduced ITPC in the N2 window, compared with the TD children. The N2-associated ITPC clearly indicates increased neural "jittering" or variable neural response in the children with autism, whereas the increased P1-associated ITPC is counterintuitive and inconsistent with the idea that overall reduced neural synchrony may underlie the sensory and cognitive abilities in individuals with autism. For instance, individuals with autism were found to display reduced alpha ITPC during visual processing of Gabor patches (Milne, 2011), and reduced beta ITPC during picture naming (Buard et al., 2013), as well as reduced theta ITPC in feedback processing of rewards and errors (van Noordt et al., 2017). In auditory processing, reduced gamma ITPC to tones has been frequently reported in children with autism (Edgar, Fisk Iv, et al., 2015; Edgar, Khan, et al., 2015). However, when taking into account processing stage, the literature suggests a trend that lower frequency bands below gamma generally show reduced synchrony beyond 200 ms post-stimulus but not in earlier windows (~100 ms) (Buard et al., 2013; Edgar, Fisk Iv, et al., 2015; Milne, 2011; van Noordt et al., 2017), whereas gamma bands often show reduced synchrony from early windows (Edgar, Khan, et al., 2015; Gandal et al., 2010).

Our results are the first to demonstrate increased theta ITPC during early stage of auditory processing in children with autism. More work is needed to clarify the role of neural synchrony in the different frequency bands during various tasks and stages of information processing in autism. There are some notable methodological differences across studies. While previous auditory ITPC studies on autism presented stimuli with jittered ISI, the current study used the fixed ISI protocol as the majority of child AEP

studies did. As timing regularity of auditory input influences the onset-driven phase alignment of ongoing EEG (Barry, 2003), it is possible that choice of ISI settings (i.e., jittered vs. fixed) can critically affect the inter-trial phase-locking. Unpublished EEG data in our lab using randomized ISIs in the range of 800-1200 ms did show reduced low-frequency ITPC within the P1 window in school-age children with autism, suggesting that neural phase-locking across trials in autism is subject to the modulation of stimulus timing predictability. Nevertheless, the current results of increased ITPC within the early P1 window is compatible with the sensory hyperactivity account (Mottron et al., 2013). It is proposed that synchronous cortical responses rely on thalamic regulation and thalamocortical connectivity (Malekmohammadi, Elias, & Pouratian, 2015; Winer, Miller, Lee, & Schreiner, 2005). The current ITPC result for the P1 response might reflect enhanced feedforward thalamocortical connectivity ascending to the primary auditory cortex. In the domain of somatosensory processing, recent work has provided indirect evidence of enhanced feedforward thalamocortical connectivity with primary sensory cortex (S1) in individuals with autism, which is linked with their enhanced response and phase-locking to the onset of somatosensory input (Khan et al., 2016).

In our study, the two subject groups differed in theta ITPC and showed opposite patterns depending on the time course of the AEP component (i.e., enhanced P1 vs. attenuated N2 in autism group). Similar results have been observed in somatosensory processing (Khan et al., 2016). In that study, individuals with autism displayed increased transient phase-locking in S1 to the onset of tactile stimulation, but decreased activities during the later steady-state portion of the response. These findings underscore the possibility that sensory hypersensitivity might negatively impact subsequent

communications between sensory and higher-order brain regions (Isler, Martien, Grieve, Stark, & Herbert, 2010).

2.4.3 Relationships between ITPC and AEPs

Consistent with our second hypothesis, theta ITPC value was a significant predictor of AEP amplitude for both P1 and N2 in both subject groups. These patterns are consistent with findings from normal adults (Koerner & Zhang, 2015). One additional piece of information is that the P1 hyper-sensitivity in the children with autism was associated with their heightened theta synchrony while their hypo-sensitivity in the N2 was associated with the reduced theta synchrony. The latter is in line with the notion that reduced inter-response synchrony in EEG oscillation may lead to overall reduced power in autism (Simon & Wallace, 2016).

An implicated question is why there is hyper-sensitive (relative to TDs) theta oscillations for onset detection but hypo-sensitive for subsequent sound content analysis in the children with autism. In normal adults, stronger theta phase-locking tends to attenuate neural excitability during speech onset (segment-level) encoding and increase excitability during the encoding of steady-state portion of the signal, an operation fundamental for syllable-level speech parsing and encoding (Giraud & Poeppel, 2012). In the current data, it appears that the neural synchrony across trials is not universally decreased in children with autism relative to TD, but rather weighted depending on the time course of the AEP components associated with low-level vs. higher-level and segmental vs. syllabic processing. This conjecture is supported by the ITPC patterns for words in the two subject groups. Judging from the data in Table 2, word-evoked theta synchrony increased from the P1 window to the N2 window in the TD group but not in

the autism group. The dynamic phase control is instrumental for efficient signal processing by the brain (Schroeder & Lakatos, 2009). Broadly speaking, individuals with autism may not have the highly organized neural rhythm to entrain to specific sensory input, which also means some sensory details or ambient signal might be unselectively amplified by the neural system.

2.4.4 Implications and Limitations

The coexistence of hyper- and hypo-sensitive responses depending on the stage of information processing highlights the complexity of abnormal sensory perception in autism. Moreover, the fact that these hyper- and hypo-sensitivities in AEPs are associated with neural phase synchrony enhancement or attenuation in the trial-by-trial theta oscillatory activities points to a more nuanced view of disrupted underlying neural mechanisms for atypical auditory information processing in autism. The exact implication of this disruption remains to be further explored in future work. It is possible that when the timing of stimulus presentation is regular and highly predictable, children with autism are less susceptible to neural adaptation to the detection of sound onset as reflected in the P1, but their later N2 responses to processing higher-level content aspect of the auditory stimuli were not as robust as the age-matched controls. In particular, individuals with autism are subject to abnormal thalamocortical function in regulating sensory filtering (Hegarty et al., 2018). This may explain some of the sensory adaptation problems with individuals with autism. In this regard, neuromodulatory (e.g., transcranial magnetic stimulation) and neurofeedback treatment targeting oscillatory dynamics such as improving the hierarchical organization of oscillation activities might be a promising

avenue for improving sensory-related behavioral issues in children with autism (Simon & Wallace, 2016).

Several limitations of our work need to be acknowledged. First, the relatively small sample size and the inclusion of mostly male participants limit the generalizability of findings. Replications are needed with larger groups of children with more balanced sex/gender ratio and children with different language backgrounds. Second, results from the addition of a complex nonspeech stimulus condition could help tease apart the influences of stimulus complexity, social relevance, and semantic significance on cortical oscillations. Third, the current study does not provide longitudinal data to verify the maturational delay interpretation. The developmental changes in children with autism may not follow the same steps as in typical development (Mottron, 2017). Moreover, further studies are needed to examine the potential impact of these sensory atypicalities on speech and language development in children with autism.

2.5 Conclusions

The current study provides confirmatory AEP and novel ITPC evidence of sensory abnormalities in school-age children with autism using simple pure tones and complex speech stimuli. School-age children with autism displayed increased theta ITPC for sound onset detection in the P1 window and attenuated theta ITPC in the subsequent N2 window. The theta ITPCs in the corresponding windows were significant predictors of P1 and N2 amplitudes, indicating disrupted oscillation synchrony as a neural generator of abnormal AEPs in children with autism. These data indicate that neural synchrony in individuals with autism might be sub-optimally organized depending on the stage of information processing that support low-level sensory vs. higher-level perceptual coding.

Chapter 3: Neural coding of linguistically relevant syllabic-level pitch pattern in children with autism

3.1 Introduction

Linguistic relevance shapes the neural coding of pitch patterns in speech and nonspeech sounds. For example, native Chinese listeners displayed greater MMN amplitude to perceptually challenging lexical tone contrast (T2/T3) compared with native English listeners, whereas such language experience effect was absent for perceptually less challenging contrast (T1/T3) (Chandrasekaran, Krishnan, & Gandour, 2007b). Furthermore, when the deviant tone became less naturalistic, that is, a linearly rising T2 instead of a prototypical curvilinear T2, only native Chinese listeners but not native English listeners displayed reduced MMN amplitude, suggesting heightened neural sensitivity to the subtle acoustic difference of pitch trajectory (Chandrasekaran, Krishnan, & Gandour, 2007a). Thai is another typical tonal language. It was found that native Thai speakers showed greater MMN amplitude to a lexical tone presented in Thai syllable (native) than the similar tone presented in Chinese syllable (nonnative) (Sittiprapaporn, Chindaduangratn, Tervaniemi, & Khotchabhakdi, 2003). These cross-linguistic findings provide strong support for the language-specific neural attunement for linguistic pitch features in terms of pre-attentive auditory discrimination.

To characterize fine-scaled sensory encoding of critical pitch features by listeners with long-term tonal language experience, Krishnan et al. (2014) compared early obligatory auditory ERPs in native Chinese listeners and English listeners in response to nonspeech noise stimuli. The stimuli included a prototypical Chinese T2 with curvilinear pitch trajectory and a linearly accelerated T2 that does not exist in any language. Instead of using speech sounds, the tones were carried by iterated ripple noise (IRN). The

nonspeech IRN carrier can eliminate influence from speech context and allows us to examine more precisely the acoustically driven cortical encoding of pitch features. The results showed that both groups displayed stronger ERP response to the prototypical T2 relative to the linear T2, and the effect was stronger in the Chinese group at the right temporal site. Assuming that right hemisphere hosts the auditory region with higher spectral resolution (Hyde, Peretz, & Zatorre, 2008), such effects of linguistic status imposed by pitch trajectory might reflect a linguistically shaped, specialized neural network in the right hemisphere for pitch processing.

To our knowledge, such language-dependent auditory attunement has not been examined in autism. The current study employed the similar nonspeech stimuli used in Krishnan et al. (2014) varying in pitch trajectory to examine more closely the cortical encoding of linguistically relevant pitch information in tonal-language-learning in children with autism. It was also aimed to provide additional information for the neural correlates of syllabic-level linguistic pitch processing in TD children, as such information on children has been sparse with the exception of a few MMN studies (Meng et al., 2005; Wang et al., 2017; Zhang et al., 2012). Following the developmental trend of auditory ERPs described in Chapter 1 Section 2.1, we expected to observe P1 and N2 (or N250) components to be elicited in these school-age children. According to Krishnan et al. (2014), prototypical T2 would produce greater ERP magnitude relative to the unnaturalistic linear T2 in native Chinese users. However, we hypothesized such pitch trajectory effect to be smaller in the autism group, as supporting evidence of reduced neural specialization for linguistic pitch pattern. Additionally, given the neural timing delay identified in many studies, the autism group may show prolonged ERP latency

relative to the TD controls, indicating slower neural processing speed. Following Study 1 (Chapter 2) finding that information processing stage played a role in AEP amplitude increment or reduction in autism, we hypothesized a similar trend of attenuated later ERP but similar or enhanced early ERP in the autism group in comparison with the TD group.

3.2 Methods

3.2.1 Participants

The participants in the autism and TD groups were recruited from the same sites as the Study 1, following the same screening protocol. Twenty-two children with autism (20 boys and 2 girls, age M =12.1 years, SD = 1.4, range = 9.5~ 14.5) and 27 age-matched TD controls (21 boys and 6 girls, age M = 11.7 years, SD = 0.8, range = 10~ 13.5) completed the study. Nonverbal IQ (NVIQ) of each child was measured using the Raven's Standard Progressive Matrices Test (Raven & Court, 1998). The autism group scored lower (M = 84, SD = 12) compared to the TD group (M = 103, SD = 11) ($t(47) = 6.99$, $p < .001$). All the children with autism were verbal with language impairment, and reportedly relied on speech and vocalizations as primary ways of communication. Fourteen out of these 22 children only spoke in single words, phrases, or simple sentences. Others in the group can speak some complex sentences but limited in range. The children in the TD group were screened to rule out being in the autism spectrum using the Social Communication Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003) . Two out of the 27 children scored above the cutoff score of 15 points for possible ASD (20, 16). Follow-up evaluation of these two children was conducted by pediatricians and determined that they were typically developing. Informed consent was obtained from each child's parent following a protocol approved by the local institutional review board.

3.2.2 Stimuli and Procedure

Two nonspeech sounds carrying two types of Tone 2 (T2) pitch contours were constructed. The naturalistic T2 was extracted from a synthesized Chinese syllable /da2/ with a male voice. As shown in Figure 3, the T2 trajectory is curvilinear and dynamically changing. Its frequency falls slightly from the beginning and accelerates at 0.33 Hz/s throughout the rest of the sound. The frequency turning point occurs around 30% through the sound duration. The frequency ranges from lowest at 91 Hz to highest at 140 Hz. These acoustic parameters are consistent with and well within the typical variations in Mandarin Chinese (Keating & Kuo, 2012; Krishnan, Gandour, Smalt, & Bidelman, 2010; Shen, Lin, & Yan, 1993; Xu, 1997), thus can represent a prototypical exemplar of T2. The unnatualistic T2 was a linearly rising frequency contour. The linear T2 shares the onset and offset frequency of the prototypical T2, thus coarsely approximated the frequency trajectory of T2. To construct the nonspeech carrier, first, an iterated rippled noise (IRN) with 100 Hz frequency was created using 32 iteration steps (Swaminathan, Krishnan, Gandour, & Xu, 2008), then the pitch tier was replaced with the above mentioned fundamental frequency (F0) extracted from T2 and the linear frequency ramp. The same stimulus generation method was used in Yu and Zhang (2018). Each sound was 250 ms in duration, including a 5 ms fade-in/out.

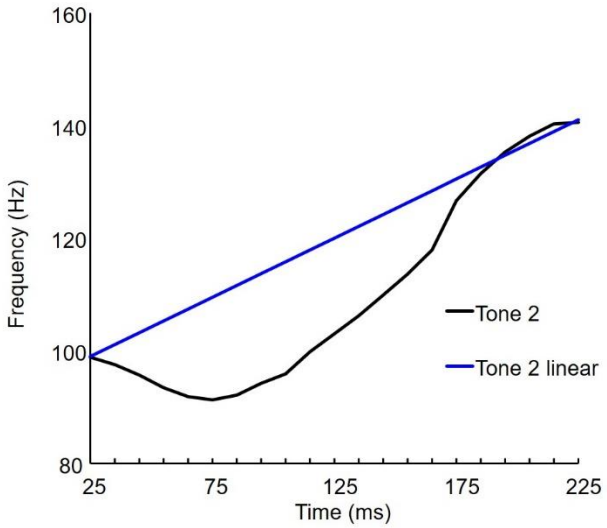

Figure 3. Pitch contours of the stimuli. The black curve represents the frequency of the prototypical Tone 2; the blue line represents the linear Tone 2.

Each sound was presented 120 times in 3 short blocks, resulting in 6 blocks in total. The presentation order of the blocks was randomized for each participant. The ISI was jittered between 800 ms and 1200 ms. The stimuli were delivered through ER-1 ear inserts at 70 dB SPL. The participant was instructed to sit still and watch a silent cartoon of his/her choice while ignoring the sounds. The total presentation time was about 7 min.

3.2.3 EEG Recording and Data Analysis

Continuous EEG was recorded with a 32-channel BrainAmps DC amplifier system at a 1000 Hz sampling rate (Brain Products, Germany). The left mastoid and AFz were used as the reference and ground, respectively. Electrode impedance was kept below 10 kΩ. Data analysis was performed using EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014). The data were high-pass filtered with 0.5 Hz cut-off. Epochs with 800 ms length (200 ms baseline) were extracted. Trials with instantaneous values exceeding ±100 μV were rejected.

Grand mean ERP waveforms and global field power (GFP) were obtained for each condition. GFP is computed as the standard deviation of amplitude across all electrodes and can serve as an objective measure of neural activities independent of scalp location (Lehmann & Skrandies, 1984). GFP is useful in determining the presence of ERP components without bias from electrode selection (Sussman, Steinschneider, Gumenyuk, Grushko, & Lawson, 2008).

ERP peak detection windows were determined based on the grand mean ERP and GFP waveforms (Figure 4&5). P50, P1, and N250 peaks were searched within post-stimulus windows of 50- 90 ms, 100-190 ms, and 190- 380 ms, respectively. In addition to the peak measure of these identifiable components, mean amplitude was computed for the time window of 180- 230 ms to better characterize the morphological difference between conditions and groups. Nine electrode sites were used for statistical analysis: Frontal region F3/Fz/F4, central region C3/Cz/C4, and parietal region P3/Pz/P4. Trial-by-trial time-frequency analysis of inter-trial phase coherence (ITPC) was carried out following the methods described in Study 1 (Chapter 2). The maximum theta ITPC values within the designated time windows of the ERP components were identified for each participant for statistical analysis.

LME regression was performed for each outcome measure. For data with repeated measures, LME is considered advantageous over repeated measures ANOVA, as it accounts for the within-subject correlations across multiple repeated measures by including subject factor as a random effect (Gueorguieva & Krystal, 2004b; Koerner & Zhang, 2017). In each regression, group (autism vs. TD), stimulus condition (prototypical T2 vs. linear T2), region (frontal, central, parietal), and interactions among variables were

entered as fixed effects; subject was entered as a random effect. Post-hoc analysis was conducted whenever necessary using the least square mean method (Lenth, 2016). Standard error bars on the bar plot were corrected for within-subject design (Cousineau, 2005).

To examine the relationships between ERP amplitude and ITPC within the corresponding windows in each group, LME model with theta ITPC as predictor variable was fit for each amplitude measure. In the LME model, stimulus condition and region were first entered as blocking variables, then theta ITPC was entered as a fixed effect and subject as a random effect. Two-tailed significance level was used for all statistical analyses.

3.3 Results

3.3.1 ERP Measures

Several ERP components were identified in the global field power (GFP) waveform (Fig. 4). The P50 response peaked around 75 ms with relatively low amplitude was followed by P1 response peaked around 130 ms. These two early components had differential spatial distribution on the scalp. P50 was distributed posteriorly and P1 was anterior. The later responses displayed some visible differences between conditions in both groups. That is, immediately following the P1 component, the ERP seemed to differ between conditions in the TD group but not so much in the autism group; both groups had a negative response with latency around 250- 300 ms, namely, the N250, with larger amplitude for the prototypical T2 compared to the linear T2.

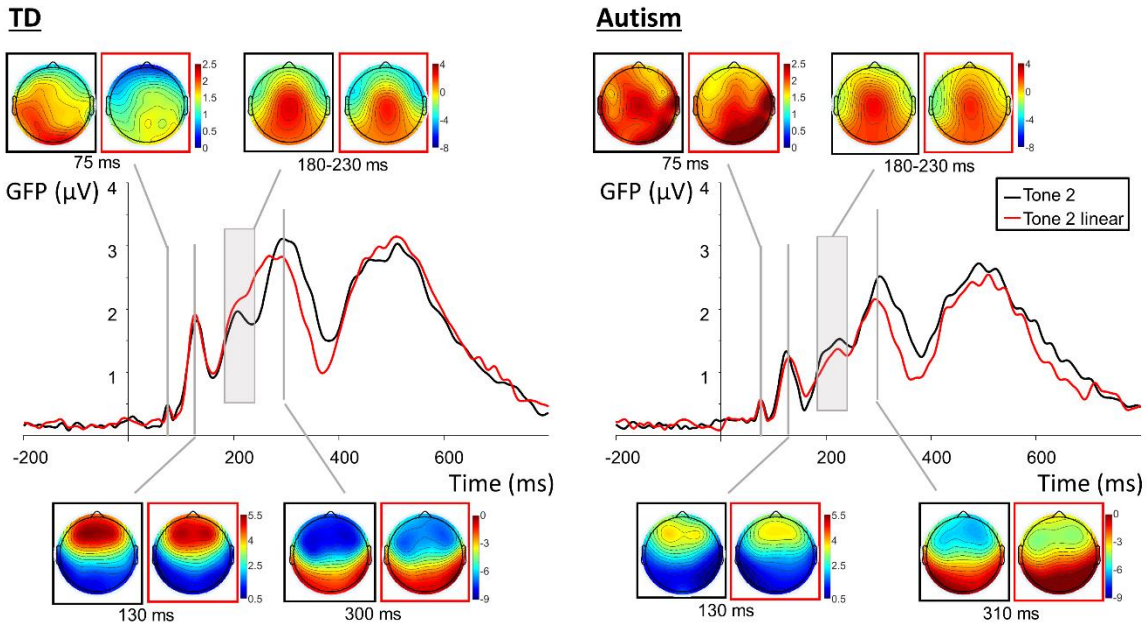

Figure 4. Global field power (GFP) waveforms of the prototypical T2 (black) and linear T2 (red) plotted against each other. The scalp maps show the topographical distribution of the ERP responses of interest (color bar: amplitude in μV). Each pair of maps represent the two stimulus conditions.

3.3.1.1 P50 Response

The LME regression for P50 latency showed significant effect of region ($F(2,823) = 9.58$, $p < .001$). Post-hoc analyses indicated shorter latency at frontal electrode sites (frontal< central = parietal) (Table 6, Fig. 5). No effect of group ($F(1,47) = 0.10$, $p = .778$), condition ($F(1,823) = 0.22$, $p = .638$), or interaction among variables was found (group*condition, $F(1,823) = 0.01$, $p = .927$; group*region, $F(2,823) = 0.16$, $p = .856$; condition*region, $F(2,823) = 0.3$, $p = .741$; 3-way, $F(2,823) = 1.11$, $p = .895$).

LME for the P50 amplitude showed a significant condition effect ($F(1,823) = 20.54$, $p < .001$), that the prototypical T2 elicited greater P50 amplitude than the linear T2 across groups (Table 6, Fig. 5). No other effect was observed (group, $F(1,47) = 0.5$, $p = .483$), region, $F(2,823) = 0.62$, $p = .537$, group*condition, $F(1,823) = 0.56$, $p = .456$;

group*region, $F(2,823) = 0.21$, $p = .809$; condition*region, $F(2,823) = 0.10$, $p = .907$; 3-way, $F(2,823) = 0.11$, $p = .896$).

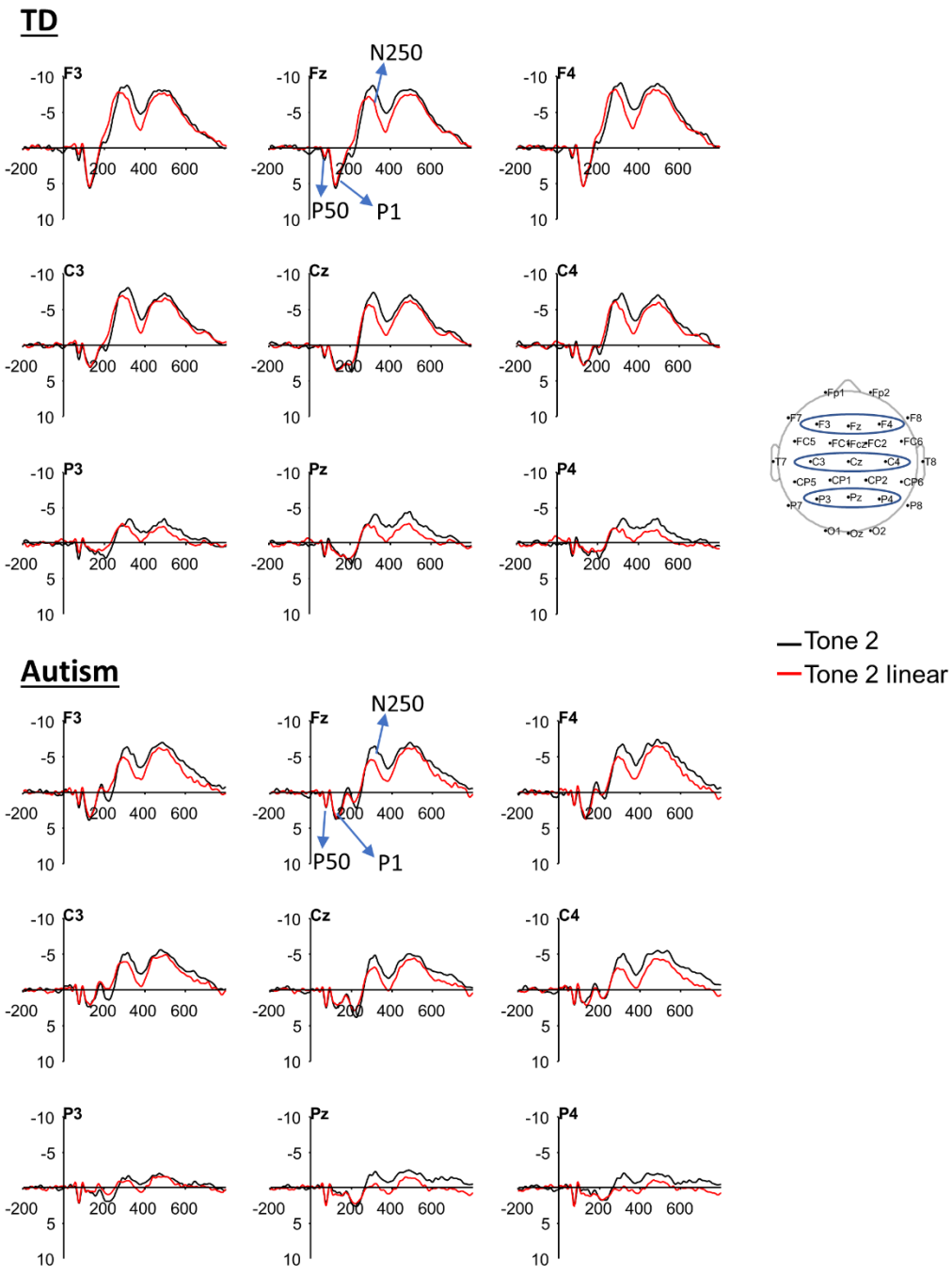

Figure 5. ERP waveforms at the electrodes of interest in the autism group and the TD group.

3.3.1.2 P1 Response

The LME regression for P1 latency revealed a significant group*region interaction ($F(2,823) = 3.69$, $p < .05$). Post-hoc analyses demonstrated that latency in the autism group was significantly shorter than the TD group at the central region ($F(1,47) = 5.39$, $p < .05$) but not at frontal ($F(1,47) = 0.49$, $p = .487$) or parietal region ($F(1,47) = 1.12$, $p = .296$) (Table 6, Fig. 5). No other effect was found for latency (group, $F(1,47) = 2.58$, $p = .115$; condition, $F(1,823) = 2.21$, $p = .138$; group*condition, $F(1,823) = 2.82$, $p = .094$; condition*region, $F(2,823) = 2.74$, $p = .065$; 3-way, $F(2,823) = 0.49$, $p = .611$).

For P1 amplitude, there was a significant group*region interaction ($F(2,823) = 6.17$, $p < .01$), associated with reduced P1 amplitude in autism than in the TD controls at the frontal sites (frontal, $F(1,47) = 5.08$, $p < .05$; central ($F(1,47) = 2.75$, $p = .104$; parietal, $F(1,47) = 1.51$, $p = .225$) (Table 6, Fig. 5). No other effect was found for P1 amplitude (condition, $F(1,823) = 2.00$, $p = .157$; group*condition, $F(1,823) = 2.56$, $p = .110$; condition*region, $F(2,823) = 0.40$, $p = .685$; 3-way, $F(2,823) = 1.05$, $p = .352$).

Table 6.

ERP latency and amplitude measures.

Group	Tone	P50		P1		180- 230 ms	N250	
		Amplitude	Latency	Amplitude	Latency	Mean amp.	Amplitude	Latency
Autism	T2	2.86(3.64)	74(10)	3.44(2.91)	134(21)	1.04 (4.16)	-5.47 (3.93)	320(23)
	T2 li	2.51(3.19)	74(10)	3.48(2.71)	135(22)	0.71 (3.86)	-4.25 (3.73)	303(31)
TD	T2	2.35(2.72)	73(11)	4.76(3.14)	142(22)	0.87 (5.18)	-7.83 (5.30)	310(25)
	T2 li	1.86(2.94)	74(11)	4.41(3.13)	139(21)	-0.06 (4.69)	-6.92 (5.06)	292(33)

SD in parentheses

3.3.1.3 Response of 180- 230 ms

The LME model revealed significant interactions of group*condition ($F(1,823) = 3.86$, $p < .05$) and group*region ($F(2,823) = 7.57$, $p < .001$). Post-hoc of these interactions demonstrated that the effects were mainly from the TD group. That is, the amplitude difference between conditions was significant in the TD group ($F(1,458) = 15.12$, $p < .001$) but not in the autism group ($F(1,373) = 2.53$, $p = .113$) (Table 6, Fig. 5&Fig. 6); the topographical distribution in the TD group was more focalized ($F(1,457) = 47.86$, $p < .001$) relative to that in the autism group ($F(1,372) = 11.71$, $p < .001$) (Fig. 4). No other main effect or interaction was found (group, $F(1,47) = 0.17$, $p = .684$; condition*region, $F(2,823) = 0.24$, $p = .784$; 3-way, $F(2,823) = 0.83$, $p = .436$).

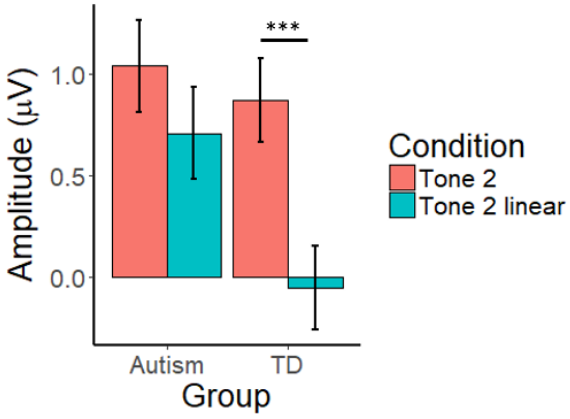

Figure 6. Bar graph of ERP mean amplitude of the 180-230 ms interval.

3.3.1.4 N250 Response

The LME regression for N250 latency revealed significant effects of group ($F(1,47) = 5.41$, $p < .05$), condition ($F(1,823) = 111.00$, $p < .001$), and region ($F(2,823) = 4.65$, $p < .01$). Further examination indicated slower N250 response in the autism group; The linear T2 elicited overall earlier response than the prototypical T2; the frontal latency was significantly shorter than that of the parietal latency (Table 6, Fig. 5). No significant interaction was found (group*condition, $F(1,823) = 0.10$, $p = .749$; group*region, $F(2,823) = 1.27$, $p = .282$; condition*region, $F(2,823) = 0.65$, $p = .520$; 3-way, $F(2,823) = 0.76$, $p = .467$).

There were significant effects of group ($F(1,47) = 5.95$, $p < .05$), condition ($F(1,823) = 50.68$, $p < .001$), and region ($F(2,823) = 407.94$, $p < .001$) on the N250 amplitude. The autism group had reduced response amplitude compared with the TD group; across groups, the prototypical T2 produced greater response than the linear T2; the N250 reached its maximum over frontal region. Additionally, there was a group*region interaction ($F(2,823) = 5.50$, $p < .01$) which revealed increasingly more significant group difference going from anterior to posterior (frontal, $F(1,47) = 3.98$, p

=.052; central, $F(1,47) = 6.73$, $p < .05$; parietal, $F(1,47) = 7.26$, $p < .01$) (Table 6, Fig. 5). No other interaction was found (group*condition, $F(1,823) = 1.00$, $p = .318$; condition*region, $F(2,823) = 0.06$, $p = .943$; 3-way, $F(2,823) = 0.24$, $p = .785$).

3.3.2 ITPC and Its Relationship with ERP Amplitude

LME regression for theta ITPC in the P50 window showed significant effect of region ($F(2,235) = 19.84$, $p < .001$, frontal > central > parietal). There was a trend of increased theta ITPC in the autism group in comparison with the TD group ($F(1,47) = 3.92$, $p = .054$) (Table 7, Fig. 7). No other effect was found (condition, $F(1,235) = 0.08$, $p = .777$; group*condition, $F(1,235) = 0.17$, $p = .683$; group*region, $F(2,235) = 1.70$, $p = .184$; condition*region, $F(2,235) = 1.01$, $p = .367$; 3-way, $F(2,235) = 0.21$, $p = .809$). However, the LME model with theta ITPC as a predictor variable for amplitude did not show any significant relationship between the two in either group (TD, $F(1,131) = 0.57$, $p = .453$; autism, $F(1,106) = 1.18$, $p = .280$).

Theta ITPC in the P1 window again showed significant region effect ($F(1,235) = 58.94$, $p < .001$, frontal > central > parietal). No other effect was found for this measure (group, $F(1,47) = 2.53$, $p = .119$; condition, $F(1,235) = 0.99$, $p = .321$; group*condition, $F(1,235) = 0.03$, $p = .874$; group*region, $F(2,235) = 1.68$, $p = .189$; condition*region, $F(2,235) = 0.67$, $p = .510$; 3-way, $F(2,235) = 0.18$, $p = .836$) (Table 7, Fig. 7). In both groups, theta ITPC was a significant predictor of P1 amplitude across conditions and electrode sites (TD, $F(1,131) = 6.89$, $p < .01$, $\beta = 6.53$; autism, $F(1,106) = 10.96$, $p < .01$, $\beta = 8.55$). The regression coefficient β indicated that greater theta synchrony predicted greater P1 amplitude in these children.

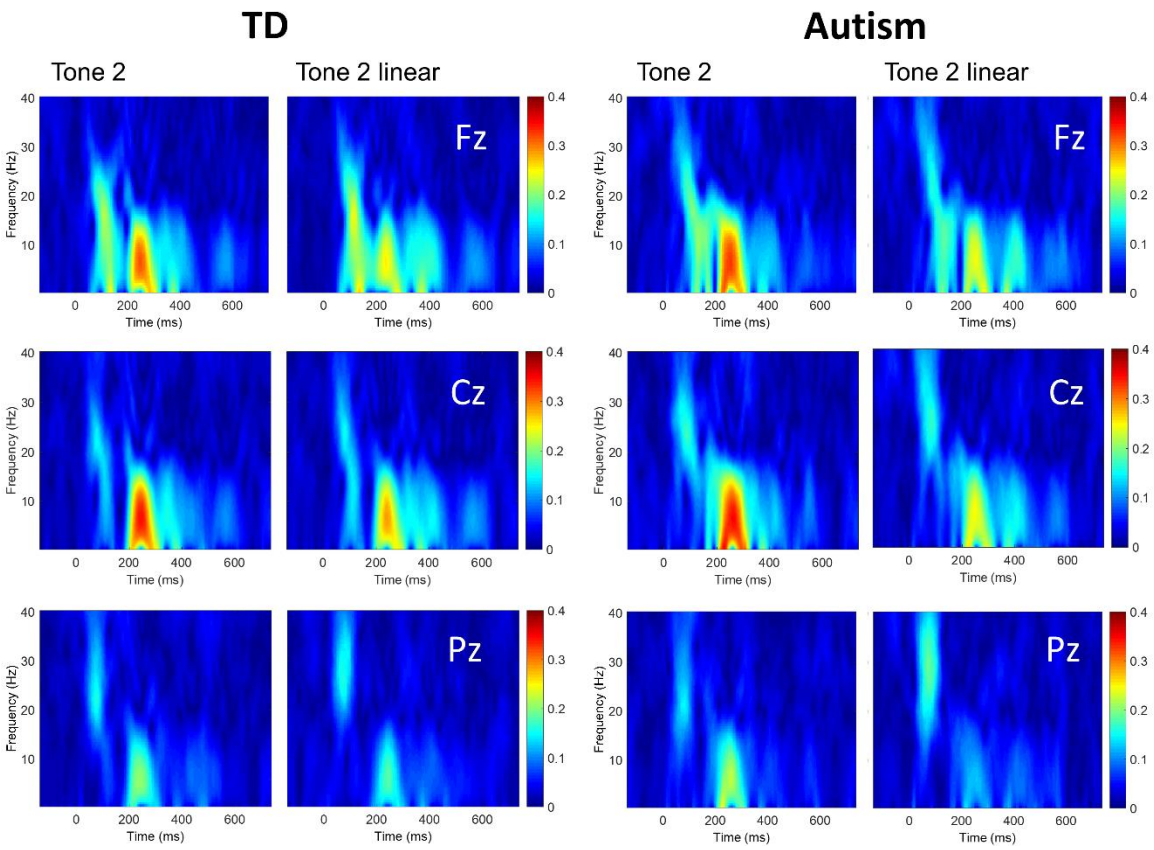

Figure 7. Time-frequency representations showing trial-by-trial phase-locking measured by ITPC at midline channels Fz, Cz, and Pz.

In the 180- 230 ms window, there were significant effects of condition with greater theta synchrony associated with the prototypical T2 ($F(1,235) = 11.91$, $p < .001$), accompanied by significant region effect ($F(2,235) = 64.78$, $p < .001$) (Table 7, Fig. 7). No other effect was found (group, $F(1,47) = 0.25$, $p = .622$; group*condition, $F(1,235) = 0.07$, $p = .790$; group*region, $F(2,235) = 0.15$, $p = .865$; condition*region, $F(2,235) = 0.13$, $p = .875$; 3-way, $F(2,235) = 0.28$, $p = .755$). Further LME modeling revealed that theta ITPC was predictive of ERP amplitude in this time window only in the autism group ($F(1,106) = 8.88$, $p < .01$, $\beta = 8.20$) but not in the TD group ($F(1,131) = 0.18$, $p = .672$, $\beta = -1.15$).

Table 7.

ITPC values in the corresponding windows of ERP.

Group	Tone	P50	P1	180- 230 ms	N250
Autism	T2	0.17(0.07)	0.28(0.08)	0.29(0.10)	0.33(0.10)
	T2 li	0.17(0.07)	0.27(0.09)	0.26(0.10)	0.29(0.10)
TD	T2	0.14(0.06)	0.24(0.11)	0.30(0.12)	0.33(0.11)
	T2 li	0.14(0.07)	0.24(0.11)	0.27(0.12)	0.30(0.11)

SD in parentheses

For the N250 window, there was a significant condition effect on theta ITPC ($F(1,235) = 23.83$, $p < .001$) such that the prototypical T2 produced greater theta synchrony than the linear T2, along with a significant region effect ($F(1,235) = 85.09$, $p < .001$, frontal = central > parietal) (Table 7, Fig. 7). No other effect was observed for the N250-associated ITPC (group, $F(1,47) = 0.06$, $p = .806$; group*condition, $F(1,235) = 0.03$, $p = .860$; group*region, $F(1,235) = 0.31$, $p = .737$; condition*region, $F(1,235) = 0.26$, $p = .772$; 3-way, $F(1,235) = 0.83$, $p = .436$). Further LME models showed that in both groups, greater theta ITPC was correlated with greater N250 amplitude (TD, $F(1,131) = 52.29$, $p < .001$, $\beta = -21.53$; autism, $F(1,106) = 19.15$, $p < .001$, $\beta = -11.79$).

3.4 Discussion

3.4.1 ERP Evoked by Pitch-carrying Complex Noise

The IRN stimuli elicited several auditory ERP components in both groups of school-age children. In addition to the P1 component, P50 was also present in these children. Although close in latency, these two early components had differential topographical distributions, indicating different neural processes for stimulus onset detection. In the ERP literature, P50 and P1 are often treated as the same such as in the P1(P50)-N1-P2 complex of adults and older children under specific listening conditions

(Čeponienė, Rinne, et al., 2002; Gilley, Sharma, Dorman, & Martin, 2005; Korzyukov et al., 2007; Sharma et al., 1997). The separated P50 and P1 here are worth noting, as it supports the idea of multiple sources of P50 or P1. The putative generator of P50 is the primary auditory cortex within the Heschl's gyrus (Korzyukov et al., 2007; Liegeois-Chauvel, Musolino, Badier, Marquis, & Chauvel, 1994; C. Ponton, Eggermont, Khosla, Kwong, & Don, 2002). However, studies have found rather complex sources, including reticular formation system (J. S. Buchwald, Rubinstein, Schwafel, & Strandburg, 1991) auditory cortex (Howard & Volkov, 2000; Tavabi et al., 2007), hippocampus and prefrontal cortex (Grunwald et al., 2003). The current observation of earlier and occipitally distributed P50 might reflect arousal reaction, and the frontally distributed P1 could be of greater auditory function.

Immediately following the P1, both groups displayed a deflection somewhere around 180- 230 ms. For the TD group, it appears only the prototypical T2 elicited the deflection but not the linear T2 (Fig. 4&5). This deflection cannot be classified as child N2 because of its morphology and scalp distribution. Coincidentally, in an ERP study with 10-year-old Chinese-speaking TD children (Zhang et al., 2012), a similar deflection was present in response to a prototypical speech T2. Critically, in that study, the deflection was absent for T1(flat tone) or T4 (falling tone). One possibility is that the ERP deflection is related to neural coding of the defining pitch feature of T2. Unlike linear T2, the frequency slope of prototypical T2 is time-varying with a prominent turning point, which requires the tracking of continuous change in the brain. That said, the response might also reflect a secondary onset response to the acoustic transient or auditory change complex (the C-process) (S. J. Jones & Perez, 2001).

We saw in the later processing stage, a large response N250 with the time range and scalp pattern of a typical child N2. Regardless of pitch conditions, the autism group had prolonged latency and reduced amplitude in this window compared to the TD group, which is consistent with our hypothesis concerning neural timing and the impact of processing stage. As mentioned in the Introduction and Study 1 (Chapter 1&2), unlike P1 or P50 which is primarily a low-level sensory component, the auditory N2 reflects more of higher-order sensory integration and fine-grain analysis of sound content (Almeqbel & McMahon, 2015; Čeponienė et al., 2009; Čeponienė et al., 2001; Čeponienė et al., 2008; Karhu et al., 1997). N2 latency shortening in children is linked with training-induced improvement of auditory skills (Cunningham et al., 2000; Fujioka et al., 2006). The fact that the curvilinear T2 produced longer N250 latency and much larger amplitude than the linear T2 might be reflective of its greater acoustic complexity with ever-changing dynamic pitch trajectory that the brain needs more time to process. The N250 enlargement is also consistent with Krishnan et al. (2014) adult finding which was interpreted as language-dependent effect transferred from the speech experience of native Chinese listeners.

3.4.2 ERP Distinction of Linguistically Relevant Pitch Patterns in TD and Autism

We observe some shared patterns of ERP distinction of the tones. Significantly, both groups demonstrated enhanced N250 response with prolonged latency in the processing of prototypical curvilinear T2 relative to linear T2. It is not surprising to see the enhancement, likely due to the recruitment of additional neuronal populations for the perceptual integration of dynamic curvilinear pitch pattern. It is also in agreement with increased activation in the human auditory cortex for sound input with spectro-temporal

modulation (Brechmann, Baumgart, & Scheich, 2002; Hall et al., 2002; Langers, Backes, & Dijk, 2003). Similarly, the reduced N250 latency for linear tone could be due to the faster spectro-temporal integration of relatively steady-state pitch trajectory as opposed to a time-varying one. At this point, it seems that the cortical function for encoding dynamically complex pitch trajectory is intact in the children with autism.

However, we observed a group*condition interaction in the 180-230 ms window. Although this measure varied considerably across individuals, the group-level distinction of pitch trajectories was only significant in the TD controls but not in the children with autism. In fact, the GFP and ERP waveforms indicated that the autism group displayed visible deflection for both tones, whereas the TD group did so for the prototypical T2 only. The interpretation of acoustic transit or turning point can no longer apply to the autism group because there was no such thing in the linear tone. One possibility is that the additional deflection might reflect enhanced general contour coding, which fits well with previous studies suggesting increased salience of pitch contour for children with autism relative to TD controls (Jarvinen-Pasley, Wallace, et al., 2008) and heightened neural sensitivity for within-category pitch contours (Wang et al., 2017; Yu et al., 2015).

3.4.3 Role of Theta ITPC in Syllabic-level Pitch Encoding

Unlike Study 1 (Chapter 2) with enhanced theta ITPC in autism within the P1 window, the groups did not differ on this measure for P50 or P1. For Study 1, we suspected ISI (i.e., jittered vs. fixed) might influence the presentation of theta synchrony in autism for auditory onset response that temporal regularity may produce heightened onset synchrony in children with autism (Yu, Wang, Huang, Wu, & Zhang, 2018). The current result produced by the jittered ISI offers some support for this interpretation.

In both groups, the prototypical T2 produced greater theta synchrony than the linear T2 in 180-230 ms and N250 windows but not in the earlier windows. This finding suggests that theta oscillation is modulated by the linguistically relevant aspect of pitch in the later processing stage. It might also be related to the steeper frequency slope later in the curvilinear T2 which was acoustically more naturalistic for native Chinese listeners compared to the linear T2, thus producing greater neural synchronization across trials. Additionally, synchronized theta oscillation contributed to greater P1 and N250 amplitude in both groups, a relationship established in Study 1 (Chapter 2). However, in the 180-230 ms interval, only the autism group displayed this relationship, suggesting that theta oscillation played a facilitative role for the potentially enhanced pitch relevant response in autism.

3.5 Conclusions

ERP and trial-by-trial theta oscillation synchrony in the Chinese school-age children with autism and TD controls are sensitive to the linguistic pitch patterns at the syllabic level represented by the prototypical and linear Tone 2 trajectories. However, unlike the TD group, the autism group failed to show early ERP distinction of the tones, suggesting a lack of auditory attunement for the linguistically relevant pitch features.

Chapter 4: Neural Coding of Word-level Prosodic Phonology in Chinese-speaking Adults

4.1 Introduction

The studies of pitch processing, among other acoustic properties, in autism have mostly focused on an individual syllable or a singular sound. However, speech perception and comprehension is not about individual sounds but relies on the whole structure of the sound organization. Prosodic cues are primarily constrained by pitch, intensity, and duration of supra-segmental elements. Numerous studies have suggested that knowledge of native prosody is acquired even prenatally by speech learners (DeCasper & Fifer, 1980; DeCasper & Spence, 1986; Mampe, Friederici, Christophe, & Wermke, 2009; Mehler et al., 1988), and has been found to serve major "bootstrapping" functions in language acquisition (Gervain & Mehler, 2010; Morgan & Demuth, 1996). For example, prosodic cues signaling phrasal boundaries are utilized by infants during the first year of life to facilitate word learning (Shukla, White, & Aslin, 2011).

Only a few studies have investigated supra-syllabic-level pitch information processing in autism. Using behavioral methods, some studies reported enhanced pitch perception of sentence-level stimuli when the pitch was low-level acoustic dimension independent from linguistic content (Jarvinen-Pasley & Heaton, 2007; Jarvinen-Pasley, Pasley, et al., 2008; Jarvinen-Pasley, Wallace, et al., 2008). Others reported impairment when the pitch dimension was manipulated to clearly convey emotional or linguistic meaning, i.e., a prosodic cue (Jiang et al., 2015; McCann, Peppe, Gibbon, O'Hare, & Rutherford, 2007). Taken together, these findings suggest that the perceptual deficit for prosody in individuals with autism is language-specific rather than a merely cue-specific

auditory phenomenon constrained by acoustic sensitivity, which is also consistent with previous electrophysiological studies with Chinese-speaking children with autism demonstrating speech-specific deficits of lexical tone perception (Wang et al., 2017; Yu et al., 2015). However, knowledge about the processing of linguistic pitch in supra-syllabic units in autism is lacking, and little is known about the neural basis of language-specific processing of larger speech units such as prosodic phonology in autism. Studies 3 (adult) and 4 (child) were motivated to address these gaps using disyllabic-level native vs. nonnative stimuli.

In TD individuals, late ERP responses that typically occur after 200 ms are considered to indicate endogenous neural activities related to higher-order cognitive-driven processes. Late negative responses have been found to index the processing of word meaning and higher-order phonological structures in children and adults (Friedrich & Friederici, 2005, 2015; Mills et al., 1993; Mills et al., 1997; Mills, Plunkett, et al., 2005). Moreover, hemispheric difference in the ERP topographic distribution has also been shown to reflect the neural processing of phonotactic organization in children and adults (Friedrich & Friederici, 2005; Mills et al., 2004; Wagner, Shafer, Martin, & Steinschneider, 2012). With regard to neural correlates for native prosody, an imaging study showed that cortical activation for sentence-level linguistic prosody (i.e., question or statement) in the native language is specialized to the left hemisphere (Tong et al., 2005). However, it is unclear whether prosodic phonology alone without any semantic activation would show similar leftward dominance.

As we have little information on whether and how the native and nonnative prosodic phonology stimuli can produce neural response differences, Study 3 was

conducted to search for the neural signature(s) of supra-syllabic level native vs. nonnative prosody in the *adult* brain. The adult "norm" data will lay the foundation for the examination of developmental changes from childhood to adulthood as well as for the interpretation of differences between TD children and children with autism.

Chinese (native) and English (nonnative) nonsense two-syllable utterances were used to examine the neural responses to prosodic phonology without the influence of semantic activation. Tonal languages such as Mandarin Chinese have significantly different prosodic characteristics from nontonal languages such as English. English is a stress language in which syllable-by-syllable stress variations are implemented to convey linguistic messages, whereas Chinese is a stress-flexible language in which syllable-by-syllable pitch variations are employed (Chrabaszcz, Winn, Lin, & Idsardi, 2014). In addition to synthesized speech, we included a nonspeech control condition with the hummed version of the speech stimuli. The hum stimuli preserved the critical acoustic features that define native prosody but without the fine phonetic details and segmental features, which allowed us to investigate acoustic vs. phonological analysis in the cortical response. Previous cross-language studies have typically found language-specific response to native phoneme in later stage of neural processing (>200 ms) (Dehaene-Lambertz, Dupoux, & Gout, 2000; Näätäneiv et al., 1997; Sharma & Dorman, 2000; Wagner et al., 2012), whereas early obligatory auditory components (N1-P2) reflect primarily acoustic feature analysis but not phonological processing of the cross-linguistic contrasts (Elangovan & Stuart, 2011; Sharma & Dorman, 2000; Wagner, Shafer, Martin, & Steinschneider, 2013). Therefore, we expected to observe native vs. nonnative

prosodic phonology (speech) distinction in the later ERP, and prosodic acoustics (hum) distinction only in the N1-P2 but not later ERP.

4.2 Methods

4.2.1 Participants

The participants were 18 healthy native Chinese-speaking adults (8 males and 10 females) with normal hearing. All participants were recruited via study flyers posted on college campus. The mean age of the participants was 21.6 years (*SD* = 1.7, range 19~ 24 years). All participants had normal hearing. All received at least six years of English-as-a-second-language education at school.

4.2.2 Stimuli and Procedure

The experimental followed a 2 (speech vs. hum) *2 (native vs. nonnative) design. The stimuli for the speech condition were ten Chinese (native) nonsense disyllabic utterances and ten English (nonnative) ones. In each pair of the Chinese and English utterances, the syllabic structure of the English token matched that of the Chinese one. For example, for a Chinese stimulus /mi3ling2/ in Pinyin, the paired English stimulus was “meeling”. The nonspeech condition was composed of the hummed version of the speech stimuli, which preserved the acoustic features of prosody but without the phonetic context that defines speech. To confirm that hum stimuli can no longer be perceived as speech, we had 5 healthy adults naïve to the stimuli rate the tokens using a 7-point Likert scale (1- definitely not speech, 7- definitely speech). The result showed that speech tokens on average received a score of 6.6 and hum tokens received 2.1. The Chinese and English words were first produced by a synthesized male voice using a commercial text-to-speech program, then digitally edited to match overall intensity and duration. All

tokens had a duration of 550 ms. The quality of the speech tokens was also checked by native Chinese speakers and native English speakers. Figure 8 displays the pitch contours and intensity contours of all speech tokens.

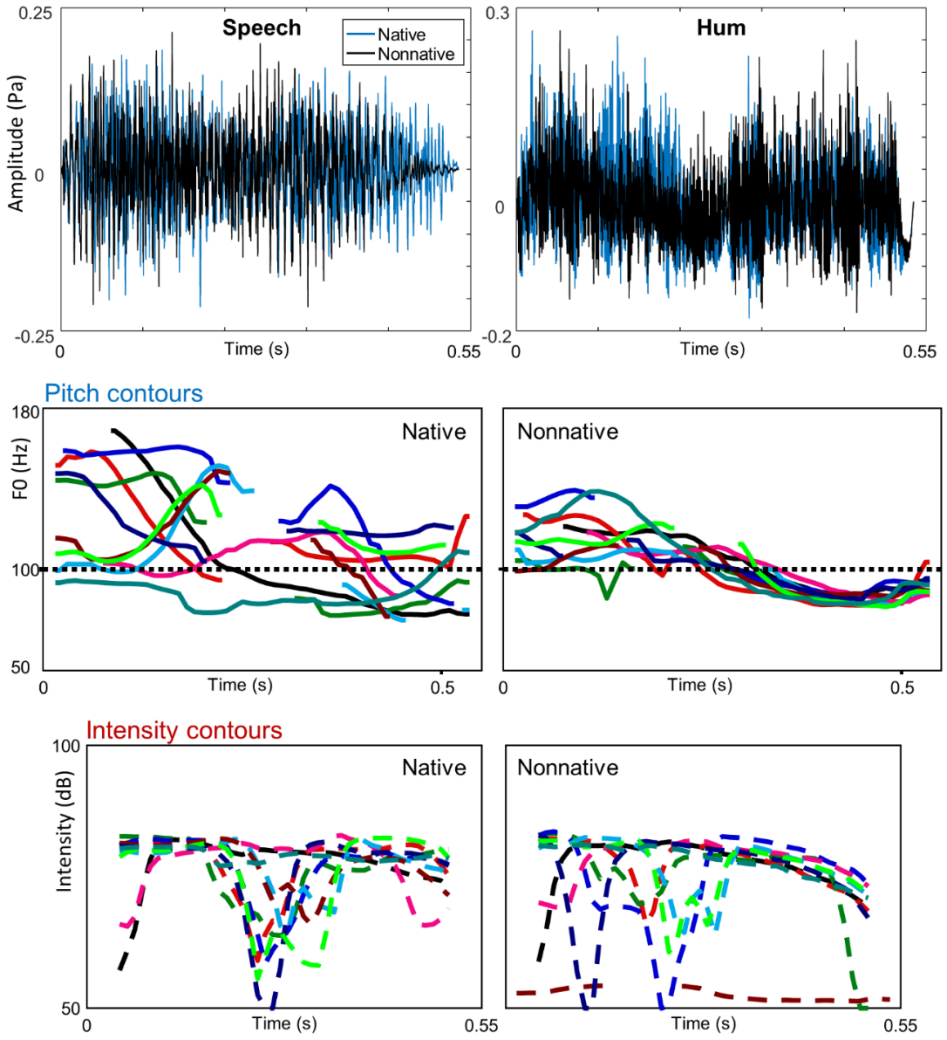

Figure 8. **Top**. Sound waveform averaged across tokens. **Middle.** Pitch contours of all native and nonnative speech tokens. **Bottom.** Intensity contours of all native and nonnative speech tokens.

Each token was presented 10 times, providing 100 trials for each stimulus condition. each condition was presented with 10 short blocks, resulting in 20 blocks in

the speech condition and 20 in the hum condition. The speech and hum stimuli were presented separately with two long blocks. The ISI was randomized between 1000 and 1200 ms. Other presentation settings were kept the same as in Study 2. The length of the presentation was about 15 min for each participant.

4.2.3 EEG Recording and Data Analysis

The EEG recording was the same as Study 2 (Chapter 3). The data were high-pass filtered with 0.1 Hz cut-off. Epochs were extracted with 1000 ms window including 200 ms baseline. Trials with instantaneous values exceeding ±80 μV were rejected. ERP peak detection windows were determined based on the grand mean ERP and GFP waveforms (Fig. 9 & 10). N1 and P2 peaks were searched within the post-stimulus windows of 100-190 ms, and 190- 380 ms, respectively. Mean amplitude of late negative response (LNR) was computed for every 50 ms interval from 300- 800 ms. Fcz electrode where the maximum ERP located on the scalp was chosen for statistical analysis. Trial-by-trial time-frequency analysis of inter-trial phase coherence (ITPC) was carried out following the analysis procedures described in Study 1 (Chapter 2). The maximum theta ITPC values within the designated time windows of the N1 -P2 and mean ITPC in each LNR interval were obtained for each participant in each condition for statistical analysis.

LME regression was performed for each outcome measure. Language (native vs. nonnative), type (speech vs. hum), and language * type interaction were entered as fixed effects; subject was entered as a random effect. Additionally, to examine the relationships between ERP amplitude and ITPC within the corresponding windows in each group, LME model with theta ITPC as predictor variable was fit for each ERP amplitude measure. In each LME model, language and type variables were first included as

blocking variables; then theta ITPC was entered as a fixed effect, and subject as a random effect. Two-tailed significance level was used for all statistical analyses. Standard error bars on all bar plots were corrected for within-subject design (Cousineau, 2005).

4.3 Results

4.3.1 ERP Measures

4.3.1.1 N1 and P2 Responses

For the latency measure, LME regression revealed significant main effects of type that the hum stimuli elicited earlier N1 and P2 responses than the speech sounds (N1, $F(1,51) = 4.45$, $p < .05$; P2, $F(1,51) = 60.12$, $p < .001$) (Table 8). There was no language main effect (N1, $F(1,51) = 0.35$, $p = .558$; P2, $F(1,51) = 0.89$, $p = .350$), or language*stimulus type interaction on latency (N1, $F(1,51) = 0.17$, $p = .734$; P2, $F(1,51) = 0.70$, $p = .408$).

The speech sounds elicited greater response amplitude of both N1 ($F(1,51) = 9.74$, $p < .01$) and P2 ($F(1,51) = 96.60$, $p < .001$). There was a language*type interaction on N1 amplitude ($F(1,51) = 13.24$, $p < .001$). Further testing indicated that nonnative speech elicited greater N1 ($F(1,17) = 9.44$, $p < .01$) whereas the effect was the opposite for the nonspeech hum sounds ($F(1,17) = 7.20$, $p < .05$) (Fig. 10). No significant interaction effect was found for P2 amplitude ($F(1,51) = 3.64$, $p = .062$).

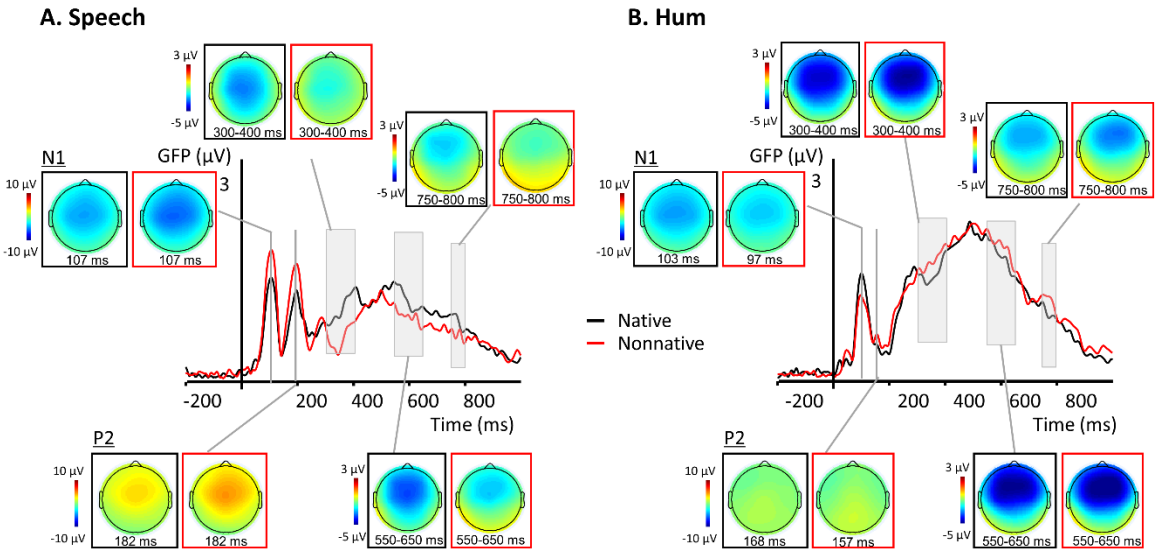

Figure 9. GFP waveforms for the native and nonnative sounds plotted against each other in the (A) speech condition and (B) hum condition. The vertical lines mark the N1-P2, and the shaded areas represent the LNR intervals that differed between native vs. nonnative speech but not native vs. nonnative hum. The topographical distributions are shown with scalp maps of the corresponding time points. Black color represents the native conditions and red represents the nonnative conditions.

Table 8.

Latency, amplitude and theta ITPC measures of N1 and P2 responses

Stimulus type	Language	N1			P2		
		Amplitude	Latency	Theta ITPC	Amplitude	Latency	Theta ITPC
Speech	Native	-5.03(2.58)	109(12)	0.36(0.07)	3.87(3.15)	196(11)	0.32(0.07)
	Nonnative	-6.13(2.73)	108(7)	0.32(0.07)	5.31(3.59)	196(11)	0.38(0.06)
Hum	Native	-5.18(2.98)	105(10)	0.26(0.10)	1.10(3.10)	172(18)	0.27(0.08)
	Nonnative	-4.21(2.53)	104(11)	0.23(0.10)	1.21(2.86)	166(19)	0.25(0.08)

SD in parentheses

4.3.1.2 Late Negative Response (LNR)

Unlike the N1 and P2, the LNR was significantly larger in the hum condition than in the speech sounds in all intervals from 300 to 800 ms (300- 350 ms, $F(1,51) = 102.52$, $p < .001$; 350- 400 ms, $F(1,51) = 47.53$, $p < .001$; 400- 450 ms, $F(1,51) = 69.73$, $p < .001$;

450- 500 ms, $F(1,51) = 50.90$, $p < .001$; 500- 550 ms, $F(1,51) = 47.77$, $p < .001$; 550- 600 ms, $F(1,51) = 47.77$, $p < .001$; 600- 650 ms, $F(1,51) = 49.03$, $p < .001$; 650- 700 ms, $F(1,51) = 18.77$, $p < .001$; 700-750 ms, $F(1,51) = 7.70$, $p < .01$; 750- 800 ms, $F(1,51) = 12.81$, $p < .001$). Language*type interaction was found in five 50-ms-intervals, 300- 350 ms ($F(1,51) = 6.29$, $p < .05$), 350- 400 ms ($F(1,51) = 10.71$, $p < .01$), 550- 600 ms ($F(1,51) = 5.34$, $p < .05$), 600-650 ms ($F(1,51) = 3.92$, $p = .053$), and 750- 800 ms ($F(1,51) = 5.31$, $p < .05$) (Table 9, Fig. 10). Further examination confirmed that in these intervals, native sounds elicited greater LNR than nonnative sounds only in the speech condition but not in the hum condition. In fact, LNR did not differ between the native and nonnative hums.

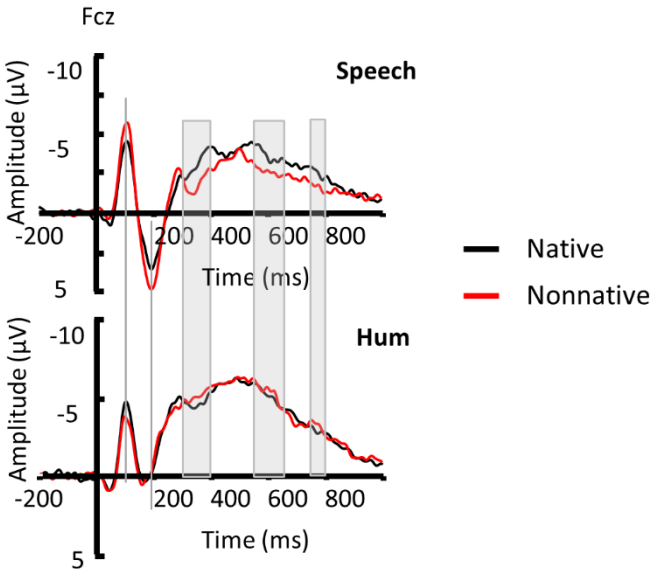

Figure 10. ERP waveforms at Fcz for the native and nonnative sounds plotted against each other. The vertical lines mark the N1-P2, and the shaded areas represent the LNR intervals that differed between native vs. nonnative speech but not native vs. nonnative hum.

Table 9.

Mean amplitude and theta ITPC measures of the LNR intervals affected by language only in the speech condition but not in the hum condition.

Type	Lan.	300- 350 ms		350- 400 ms		550- 600 ms		600- 650 ms		750- 800 ms	
		Amp	ITPC	Amp	ITPC	Amp	ITPC	Amp	ITPC	Amp	ITPC
S	Na	-2.43	0.11	-3.70	0.10	-4.07	0.07	-3.30	0.06	-2.50	0.06
		(2.15)	(0.06)	(2.40)	(0.05)	(2.13)	(0.04)	(2.12)	(0.04)	(1.68)	(0.04)
	Non	-1.46	0.18	-2.26	0.11	-2.78	0.10	-2.46	0.09	-1.62	0.07
		(2.03)	(0.06)	(2.25)	(0.06)	(2.13)	(0.05)	(1.86)	(0.04)	(1.77)	(0.02)
H	Na	-4.48	0.14	-4.80	0.12	-5.51	0.10	-4.71	0.11	-2.83	0.10
		(2.48)	(0.06)	(1.85)	(0.05)	(2.39)	(0.04)	(2.06)	(0.04)	(1.82)	(0.04)
	Non	-4.85	0.12	-5.33	0.11	-5.66	0.10	-4.99	0.13	-3.15	0.09
		(2.88)	(0.05)	(2.53)	(0.05)	(1.96)	(0.05)	(2.22)	(0.06)	(1.84)	(0.04)

SD in parentheses; S, speech; H, hum; Lan, language; Na, native; Non, nonnative.

4.3.2 ITPC and Its Relationship with ERP Amplitude

In the N1 and P2 windows, speech sounds produced greater theta ITPC than the nonspeech hum stimuli (N1, $F(1,51) = 69.88$, $p < .001$; P2, $F(1,51) = 75.65$, $p < .001$). No significant language effect was observed for either window (N1, $F(1,51) = 0.19$, $p = .669$; P2, $F(1,51) = 3.17$, $p = .081$). There was significant language*type interaction on theta ITPC for both components (N1, $F(1,51) = 11.90$, $p < .01$; P2, $F(1,51) = 20.88$, $p < .001$): Consistent with the ERP peak result, nonnative sound evoked greater theta synchrony than the nonnative sound did in the speech condition (N1, $F(1,17) = 14.41$, $p < .01$; P2, $F(1,17) = 37.93$, $p < .001$) but the opposite in the nonspeech condition (N1, $F(1,51) = 9.65$, $p < .01$; P2, $F(1,17) = 7.62$, $p < .05$) (Table 8; Fig. 11). The LME regression model also demonstrated the predictive role of theta ITPC in N1($F(1,51) = 71.72$, $p < .001$, $\beta = -19.40$) and P2 ($F(1,51) = 8.64$, $p < .01$, $\beta = 10.64$) across language and stimulus type conditions. The β value indicated that greater theta synchrony was associated with greater N1 and P2 amplitude.

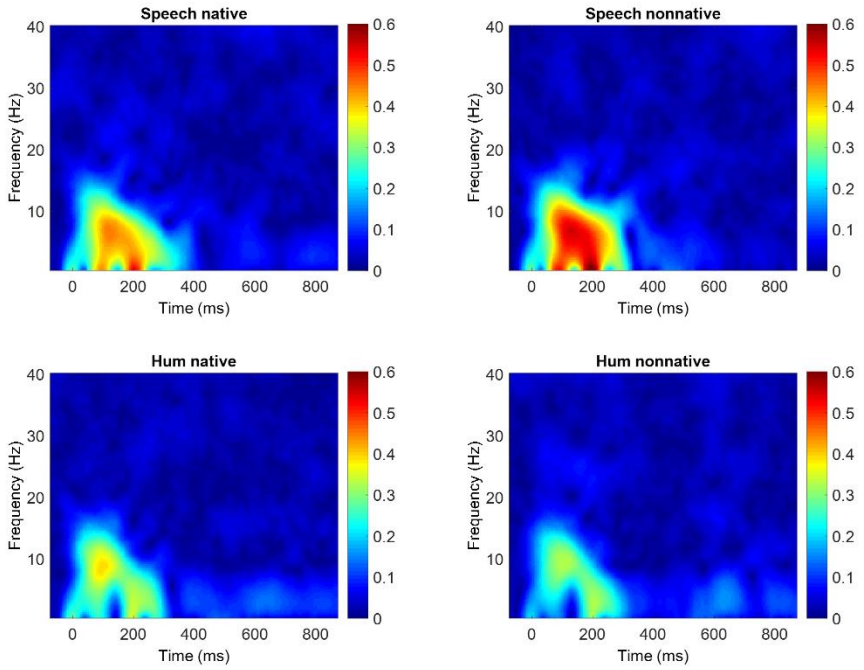

Figure 11. ITPC at 0.5- 40 Hz as a function of time.

The analysis of LNR-associated ITPC was focused on the time intervals where LNR showed language*type interaction (shaded intervals in Fig. 9 & 10). Similar to the LNR findings, hum produced greater theta ITPC than the speech sounds in 600- 650 ms ($F(1,51) = 22.33$, $p < .001$) and 750- 800 ms ($F(1,51) = 15.54$, $p < .001$). There was also a main effect of language that theta ITPC for native sounds was smaller compared to that for nonnative sounds (300- 350 ms, $F(1,51) = 5.03$, $p < .05$; 600- 650 ms, $F(1,51) = 9.38$, $p < .01$). Significant language*type interaction was found only in 300- 350 ms ($F(1,51) = 13.46$, $p < .001$). Post-hoc test indicated greater ITPC for the nonnative speech than for native speech ($F(1,17) = 12.22$, $p < .01$) and no difference in the hum condition ($F(1,17) = 1.70$, $p = .209$). The LME models did not show any predictive relationship between theta ITPC and LNR amplitude (Fig. 12).

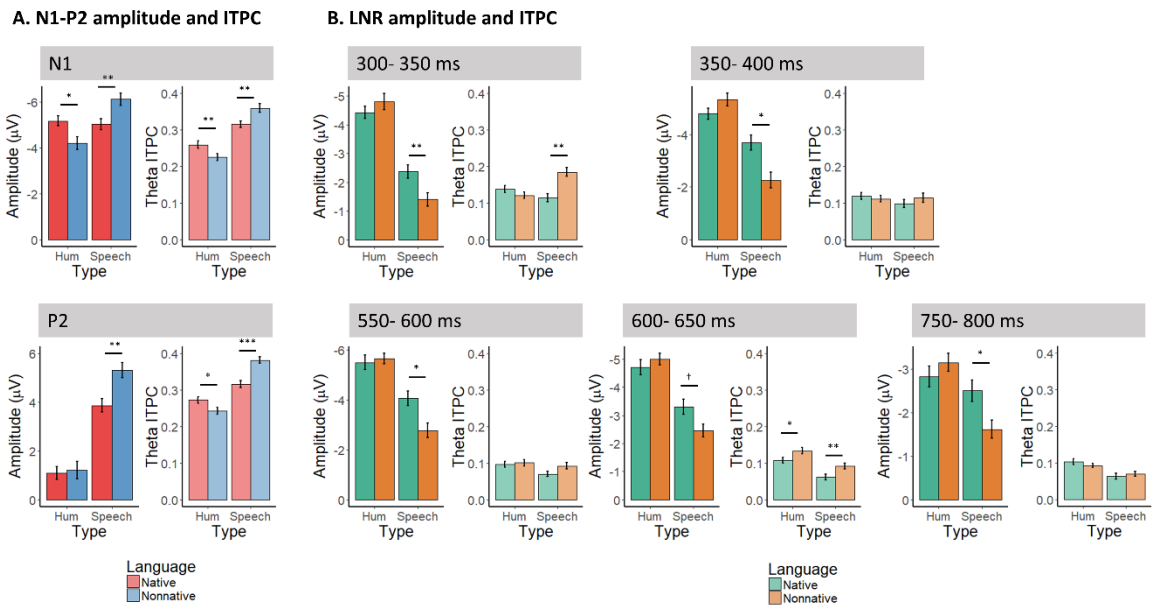

Figure 12. Bar graphs of ERP amplitude plotted in darker colors and the corresponding ITPC plotted in lighter colors. **(A)** N1 and P2; **(B)** LNR intervals showing type*language interaction.

4.4 Discussion

The aim of this study was to examine the neural responses to disyllabic prosodic phonology in healthy Chinese-speaking adults. Three major findings are to be discussed. First, "speechness" (i.e., whether the sound is speech or not) modulated both the earlier N1-P2 and later LNR, but in the opposite directions. Second, as hypothesized, LNR differentiated between the native and nonnative prosodic phonology but not prosodic acoustics. Third, trial-by-trial oscillation synchrony contributed to the earlier N1-P2 amplitude but not later LNR amplitude.

4.4.1 Speech vs. Hum

Speech elicited greater N1 amplitude than the hum stimuli. This finding is not surprising, as N1 is known to be sensitive to sound salience (Čeponienė, Yaguchi, et al., 2002; Näätänen, 1990; Näätänen & Picton, 1987; Pereira et al., 2014) and audibility

(Martin, Kurtzberg, & Stapells, 1999). Despite inevitable differences in the onset acoustics from hum, speech is by nature more salient than nonspeech to the human auditory system. Moreover, P2 was diminished in the hum condition. This result might be due to the less acoustic richness of hum compared to speech, provided that P2 is modulated by acoustic complexity (Shahin, Roberts, Miller, McDonald, & Alain, 2007). However, the fact that P2 was nearly "shut down" by the hum stimuli let us lean towards the sound classification account of P2, that P2 is involved in the auditory system's placement of sound categories (Garcia-Larrea, Lukaszewicz, & Mauguiére, 1992). Similarly, P2 amplitude has been shown to index training-induced improvement in auditory discrimination (Tong, Melara, & Rao, 2009; Tremblay & Kraus, 2002; Tremblay, Kraus, Mcgee, Ponton, & Otis, 2001), prone to conscious and attentional modulation (Crowley & Colrain, 2004). Unlike the speech sounds, hum sounds were probably unclassifiable by the adult auditory system which resulted in the weakened P2 component. This hypothesis can be tested by training listeners with hum sounds assigned with meaning and see if P2 emerges post-training.

Interestingly, these adults displayed increased LNR amplitude to the hum stimuli despite having reduced transient N1-P2. As mentioned earlier, later ERP components reflect fine-grain sound feature analysis beyond the initial sound detection. It is also linked with endogenous cognitive-driven processes as oppose to exogeneous stimulus-driven processes, more likely indexing higher-order sensory integration and abstraction (Čeponienė et al., 2008). From a neural efficiency viewpoint, larger LNR might reflect the struggling neural circuitry in the process of integrating the content of unfamiliar auditory input.

4.4.2 Native vs. Nonnative Prosody

In both the speech and hum contexts, there were significant differences in N1 between native and nonnative sounds, but in the opposite directions. Although we matched the overall intensity, the onset intensity was impossible to control due to the inherently different stress patterns of Chinese and English language. As a matter of fact, as illustrated by the average waveforms of native and nonnative tokens, the onset intensity and slope seem to be higher in the English speech tokens than those in Chinese, and the opposite seems to be true for the hum tokens. N1 amplitude is thought to reflect the amount of acoustic energy transmitted by the auditory system for onset detection (Čeponienė et al., 2005). Thus, we argue that differences in the obligatory N1-P2 between native and nonnative sounds might reflect the brain's automatic registration of acoustic differences before linguistically relevant higher-order processes. A corroborative evidence is from a transcranial magnetic stimulation (TMS) study in which neuromodulatory effects on emotional prosody recognition only occurred at later time windows but not earlier windows, suggesting that speech prosody is beyond the basic acoustic analysis in the earlier information processing stage (Hoekert, Bais, Kahn, & Aleman, 2008).

A robust interaction between stimulus type and language was found in the LNR. Consistent with our hypothesis, LNR only differed between the native and nonnative conditions in the context of speech. In fact, the only native vs. nonnative distinction we observed for hum was N1 amplitude resulted from initial acoustic feature coding. The fact that the LNR effect was speech-specific supports the idea of LNR as a correlate of language-specific processing of prosodic phonology but not prosodic acoustics.

Unlike previous work showing left hemisphere dominance for linguistic prosody processing when contrasted with non-linguistic tasks (Arciuli & Slowiaczek, 2007; Gandour et al., 2004; Kreitewolf, Friederici, & von Kriegstein, 2014; Tong et al., 2005), we did not observe any hemispheric asymmetry in the ERP. But our result is consistent with other studies using passive listening paradigm which does not require active attention to speech content (Zhang et al., 2005) and with non-linguistic prosodic stimuli (Diamond & Zhang, 2016), as well as meaningless word stimuli (Friedrich & Friederici, 2005). In our case, the speech stimuli were not linguistic (e.g., statement vs. question) or emotional, but an abstract representation of the prosodic structure. Therefore, bilateral involvement was reasonable given that phonetic processing is thought to be bilateral in adults (Binder et al., 2000; Hickok & Poeppel, 2007). Altogether, the LNR might be indexing the abstraction of supra-syllabic prosodic phonology or syllable-to-syllable processing scheme in native Chinese users.

4.4.3 Role of Theta ITPC in ERP Amplitude

Theta ITPC significantly contributed to the N1 and P2 amplitude across stimulus types and languages, providing supporting evidence for the oscillation phase concentration as a neural generator of evoked potentials (Fuentemilla, Marco-Pallares, & Grau, 2006; Klimesch et al., 2007). The spectrograms and bar plots indicate that trial-by-trial oscillation synchrony became much weaker onwards from 300 ms. In other words, as the neural process moved up to higher levels in the auditory hierarchy, it becomes less phase-locked across trials. This observation is coherent with the ERP waveform findings. Unlike the obligatory N1-P2 components, LNR had much wider peaks, indicating more sustained and variable neural activation time course. Moreover, the tokens all had

different acoustic and phonological features, which means perceptual integration reflected in the LNR varies in time across stimuli. Regardless of language, hum sounds produced greater theta synchrony than the speech sounds. This might be associated with the fact that segmental differences in the disyllabic speech stimuli were absent in the hum stimuli, making the hum stimuli acoustically less variable thereby resulting in greater trial consistency. However, there was no predictive relationship or shared pattern of language effect between LNR amplitude and associated theta ITPC. Taken together, we concluded that in our case, trial-by-trial oscillation synchrony has tenuous explanatory power for the endogenous phonological processes reflected in the LNR.

4.5 Conclusions

This study investigated Chinese adults' ERP and oscillatory synchrony in response to native vs nonnative prosodic features in the context of speech and nonspeech. The results provided evidence for the late negative response (LNR) but not the associated trial-by-trial oscillation synchrony as a neural signature of language-specific processing of prosodic phonology.

Chapter 5: Examining the Neural Coding of Word-level Prosodic Phonology in Chinese-speaking Children with Autism

5.1 Introduction

Compared with other aspects of auditory and speech perception, the processing of prosodic phonology in autism has been significantly understudied. This is an oversight given the extensive auditory atypicality in autism and the critical role of prosody in speech acquisition and language development. Despite the oversight, efforts have been made to examine the functional use of prosody of individuals with autism, as part of the general language and social cognition function.

Individuals with autism are often reported having unusual speech intonation. From the listeners' perspective, their prosody tends to be monotonal or "sing-songy" (McCann & Peppe, 2003). From acoustic point of view, people with autism have larger global pitch range compared to those without (Diehl, Watson, Bennetto, McDonough, & Gunlogson, 2009; Hubbard, Faso, Assmann, & Sasson, 2017; Nadig & Shaw, 2012), but lack the complex and flexible use of pitch variation (Bonneh, Levanon, Dean-Pardo, Lossos, & Adini, 2011; Depape, Chen, Hall, & Trainor, 2012; Green & Tobin, 2009). Children with autism were shown having difficulty imitating speech prosody (Filipe, Frota, Castro, & Vicente, 2014; Peppe, McCann, Gibbon, O'Hare, & Rutherford, 2007). Increased overall pitch range was also reported with a sample of native Cantonese-speaking children with autism (Chan & To, 2016), indicating cross-linguistic atypicality in expressive prosody.

Behavioral studies have found that school-age children with autism had difficulty identifying question sentences conveyed by prosody and emotional prosody

understanding (Jarvinen-Pasley, Peppe, King-Smith, & Heaton, 2008; Peppe et al., 2007). Moreover, children with autism were more likely to hear the same prosody as different in those two studies. However, such response bias was not supported by electrophysiological data. Using an oddball paradigm, several studies observed diminished MMN to emotional prosody change in Finnish one-word utterances, suggesting reduced neural sensitivity to prosodic features at the pre-attentive level (Korpilahti et al., 2006; Kujala, Lepisto, Nieminen-von Wendt, Näätänen, & Näätänen, 2005; Lindstrom, Lepisto-Paisley, Vanhala, Alen, & Kujala, 2016). This inconsistency suggests that behavioral measures cannot be mapped onto neurophysiological measures at pre-attentive level, instead, some higher-level mechanisms must be involved in the observed behavioral differences. Indeed, Eigsti, Schuh, Mencl, Schultz, and Paul (2012) found that adolescents with autism activated more diffusive cortical regions during emotional and linguistic prosody recognition to achieve the same level of performance as the TD controls. This result was interpreted as greater dependence on cognitive control and active "mentalizing". Similarly, Gebauer, Skewes, Horlyck, and Vuust (2014) tested adults with autism and observed a tendency of increased activation in the right caudate, suggesting greater attentional demand of emotional prosody task for these individuals. Overall, people with autism tend to perform poorly when it comes to the linguistic and social meaning of prosody, but as age increases the performance tends to improve, perhaps associated with acquired top-down executive modulation.

The reviewed work using meaningful speech material provide little information about the auditory and phonological basis of prosody perception in autism because the lower-order phonetic information was confounded with the higher-order linguistic

meaning of the stimuli. Indeed, disrupted neural processing of word meaning has been observed in very young children with autism. Coffey-Corina et al. (2008) tested 2-year old toddlers with autism with known words and unknown words and found that the TD and HFA groups showed known vs. unknown distinction in ERP confined to the left electrodes, whereas the LFA group' distinction was bilaterally distributed with a rightward dominance. This finding suggests disrupted cortical organization for language that may underlay core symptoms. Furthermore, Kuhl et al. (2013) found a predictive relationship between the early ERP measure for words and later language outcome in toddlers with autism. Assuming leftward re-organization of cortical activity for word processing marks the fundamental change of word experience specific to meaning (Mills, Conboy, et al., 2005), one could argue that prosody impairment is mostly related to impairment in linguistic functioning. However, given the auditory atypicality for syllabic pitch and durational information in autism (Chapter 1 Introduction and Chapter 3 Study 2), it is possible that neural coding of supra-syllabic prosodic acoustics is fundamentally altered, such that information conveyed by these acoustic features is harder to comprehend for children with autism.

Based on the neural signatures identified in Study 3 (Chapter 4), the current study was conducted with school-age Chinese children with autism, following the same protocol, to investigate their auditory cortical processing of disyllabic prosodic phonology and acoustics. If increased auditory sensitivity for pitch impedes the formation of language-specific representation of tonal features, we would expect to see reduced neural specialization for phonological structures at the supra-syllabic level as well. The TD children may show similar LNR pattern as in the adult data (Chapter 4) for prosodic

phonology distinction. That is, LNR was expected to be greater for native speech compared to nonnative speech. In comparison, children with autism would show more diffused native vs. nonnative distinction over the scalp as an indicator of reduced cortical specialization for linguistic structure (Kuhl et al., 2013). Given the increased cue-specific auditory sensitivity for spectral information associated with the right hemisphere function in autism, the neural specialization problem might arise from an acoustic mode of processing. If this is the case, they may additionally display LNR differentiation for the native vs. nonnative prosodic acoustics carried by the nonspeech hum stimuli.

5.2 Methods

5.2.1 Participants

The participant recruitment procedure was the same as Study 2 (Chapter 3). Twenty-one children with autism (19 boys and 2 girls, age M =12.1 years, SD = 1.4, range = 9.5~ 14.5) and twenty-five age-matched TD controls (20 boys and 5 girls, age M = 11.7 years, SD = 0.8, range = 10~ 13.5) completed the study. Nonverbal IQ (NVIQ) of each child was measured using the Raven's Standard Progressive Matrices Test (Raven & Court, 1998). The autism group scored lower (M = 84, SD = 12) compared to the TD group (M = 103, SD = 11) ($t(44) = 5.66$, $p < .001$). All the children with autism were verbal with language impairment, and reportedly relied on speech and vocalizations as primary ways of communication. Fourteen out of these 21 children only spoke in single words, phrases, or simple sentences. Others in the group can speak some complex sentences but limited in range. The children in the TD group were screened to rule out being on the autism spectrum using the Social Communication Questionnaire or SCQ

(Rutter et al., 2003). Two children in the TD group scored above the cutoff score for possible ASD and were determined as TD by pediatricians through follow-up evaluation.

5.2.2 Stimuli and Procedure

The stimuli and testing procedures followed those of Study 3 (Chapter 4). Cartoon videos were available for the children to choose from and watch during the cap fitting and stimulus presentation.

5.2.3 EEG Recording and Analysis

The EEG recording procedures were the same as in Study 3. Independent component analysis (ICA) of continuous EEG was performed to remove ocular and muscle artifacts. For other preprocessing details, see Chapter 3 (Study 2). ERP peak detection windows were determined based on the grand mean ERP and GFP waveforms (Fig. 13&14). P1 and N1 peaks were searched within post-stimulus windows of 50- 90 and 100-150 ms, respectively; mean LNR amplitude was computed for every 200 ms interval from 200- 800 ms. Four electrode sites were selected for statistical analysis based on the topographical distributions, including F3/C3 on the left side of scalp and F4/C4 on the right. Trial-by-trial time-frequency analysis of inter-trial phase coherence (ITPC) for P1 and N1 was performed following the methods described in Study 1 (Chapter 2).

LME regression was performed for each outcome measure for the speech and hum conditions separately. In each LME regression, group (autism vs. TD), language (native vs. nonnative), hemisphere (left vs. right), and interactions among variables were entered as fixed effects; subject was entered as a random effect. The relationships between ERP amplitude and ITPC within the corresponding windows were also modeled using LME regression for each group. Language and hemisphere were first entered in the

model as blocking variables; then theta ITPC was entered as a fixed effect and subject as a random effect. Two-tailed significance level was used for all statistical analyses. Standard error bars on all bar plots were corrected for within-subject design (Cousineau, 2005).

5.3 Results

5.3.1 ERP Measures

5.3.1.1 P1 and N1 Responses

In the speech condition (Table 10, Fig. 14), the LMR regression revealed a nonsignificant trend of group*language*hemisphere interaction on P1 latency ($F(2,316) = 2.94$, $p = .054$). Post-hoc analyses indicated right-only latency prolongation for the native sounds in the TD group ($F(1,74) = 5.32$, $p < .05$). In the hum condition (Table 11, Fig. 15), the 3-way interaction was significant ($F(2,316) = 3.83$, $p < .05$), with left-only latency shortening for the native sounds in the TD group ($F(1,74) = 6.18$, $p < .05$). Three-way interaction on P1 peak amplitude was observed for speech but not hum ($F(2,316) = 3.45$, $p < .05$). Post-hoc tests showed that this effect was driven by right-only P1 enhancement for the native speech in the TD group ($F(1,74) = 6.61$, $p < .05$).

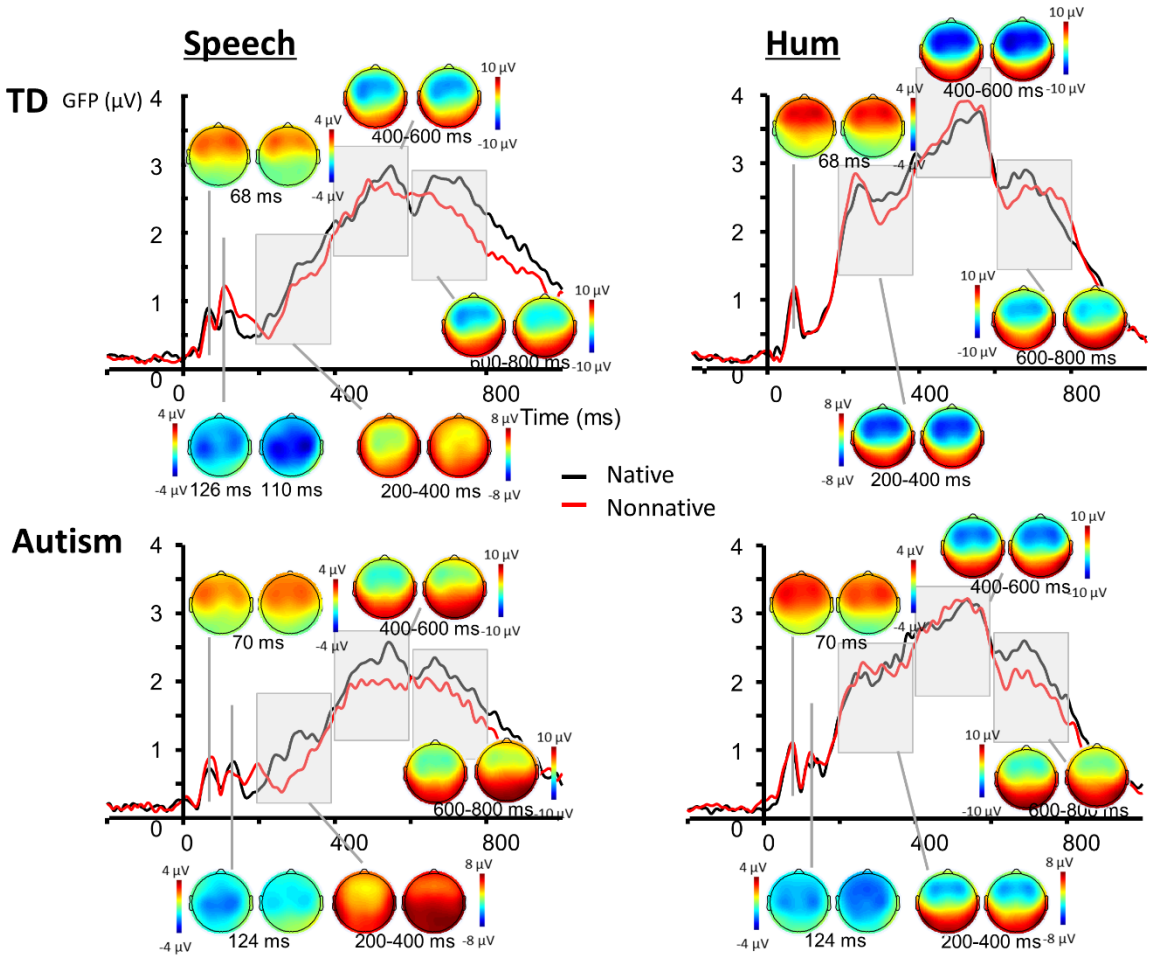

Figure 13. GFP waveforms for the native and nonnative sounds plotted against each other in the **(A)** speech condition and **(B)** hum condition. The vertical lines mark the P1 and N1 components, and the shaded areas represent the LNR windows of interest from 200-800 ms. For each pair of scalp maps, the one on the **left** represents the native condition and one on the right is the nonnative condition.

LME regression of N1 latency to speech showed significant main effects of group ($F(1,44) = 4.58$, $p < .05$) and language ($F(1,316) = 9.85$, $p < .01$) (Table 10, Fig. 14). Specifically, compared to the TD group, the autism group had delayed N1; across groups, nonnative speech produced earlier N1 latency than the native speech. No effect in the hum condition was observed for N1 latency. For the N1 amplitude, LME regression

revealed a significant 3-way interaction of group*language*hemisphere in the speech condition ($F(2,316) = 3.24$, $p < .05$). Post-hoc analyses indicated left-only N1 enhancement for the nonnative speech in the TD group ($F(1,74) = 4.72$, $p < .05$). In the hum condition (Table 11, Fig. 15), the only effect found for N1 amplitude was hemisphere that N1 was greater on the right scalp than on the left ($F(1,316) = 4.19$, $p < .05$).

Table 10.

*P1 and N1 amplitude (μV) and latency (ms), and theta ITPC in the corresponding windows in the **speech** condition.*

Group	Hemi.	Language	P1			N1		
			Amp	Lat	ITPC	Amp	Lat	ITPC
TD	Left	Native	2.63(2.24)	70(10)	0.16(0.06)	-3.36(4.09)	121(17)	0.19(0.07)
		Nonnative	2.45(2.25)	68(10)	0.16(0.07)	-4.14(3.39)	120(17)	0.20(0.07)
	Right	Native	2.88(2.17)	69(12)	0.15(0.06)	-3.68(2.71)	122(16)	0.18(0.07)
		Nonnative	2.09(2.34)	65(10)	0.16(0.06)	-4.37(3.36)	119(16)	0.20(0.07)
Autism	Left	Native	2.81(1.74)	70(11)	0.16(0.07)	-2.40(3.56)	131(13)	0.17(0.07)
		Nonnative	3.07(2.18)	73(12)	0.14(0.08)	-2.24(3.48)	124(15)	0.17(0.08)
	Right	Native	2.47(1.77)	72(11)	0.14(0.07)	-2.61(3.74)	130(12)	0.16(0.07)
		Nonnative	2.79(1.97)	70(11)	0.14(0.07)	-2.41(3.15)	119(16)	0.17(0.08)

SD in parentheses

Table 11.

*P1 and N1 amplitude (μV) and latency (ms), and theta ITPC in the corresponding windows in the **hum** condition.*

Group	Hemi.	Language	P1			N1		
			Amp	Lat	ITPC	Amp	Lat	ITPC
TD	Left	Native	3.33(2.40)	67(11)	0.16(0.07)	-1.97(3.93)	118(20)	0.20(0.07)
		Nonnative	3.30(2.27)	71(11)	0.16(0.07)	-2.04(3.65)	121(18)	0.20(0.08)
	Right	Native	3.33(2.30)	69(10)	0.16(0.07)	-2.24(3.60)	121(19)	0.18(0.07)
		Nonnative	3.15(2.42)	72(9)	0.15(0.07)	-2.31(2.78)	121(17)	0.18(0.07)
Autism	Left	Native	3.43(2.93)	71(8)	0.17(0.09)	-2.84(3.16)	125(14)	0.20(0.10)
		Nonnative	3.00(1.52)	69(10)	0.17(0.08)	-3.05(3.02)	125(14)	0.19(0.09)
	Right	Native	3.31(3.09)	70(9)	0.17(0.08)	-3.21(2.73)	124(16)	0.20(0.09)
		Nonnative	3.04(2.15)	70(10)	0.17(0.08)	-3.54(3.78)	122(16)	0.20(0.09)

SD in parentheses

5.3.1.2 LNR

In the speech condition (Table 12, Fig. 14), LME regression revealed several main effects and interactions in the LNR windows of interest. In the 200- 400 ms windows, a group*language interaction was found ($F(1,316) = 4.00$, $p < .05$). Post-hoc analyses indicated native vs. nonnative LNR difference in both groups but greater effect in the autism group (TD, $F(1,172) = 10.37$, $p < .01$; Autism, $F(1,166) = 21.79$, $p < .001$). In the 400- 600 ms window, there was a significant 3-way interaction ($F(2,316) = 3.99$, $p < .05$), driven by a right-only native vs. nonnative difference in the autism group ($F(1,62) = 10.76$, $p < .01$). In the last 600- 800 ms window, there were main effects of language ($F(1,316) = 30.13$, $p < .001$) and hemisphere ($F(1,316) = 7.81$, $p < .01$). Specifically, native speech elicited larger LNR than nonnative speech, and LNR was greater on the left than on the right electrode sites. Although overall speech LNR did not differ significantly

between groups, there were tendencies of reduced LNR in the autism group evidenced by the waveforms and statistical value (600-800 ms, $F(1,44) = 3.86$, $p = .056$).

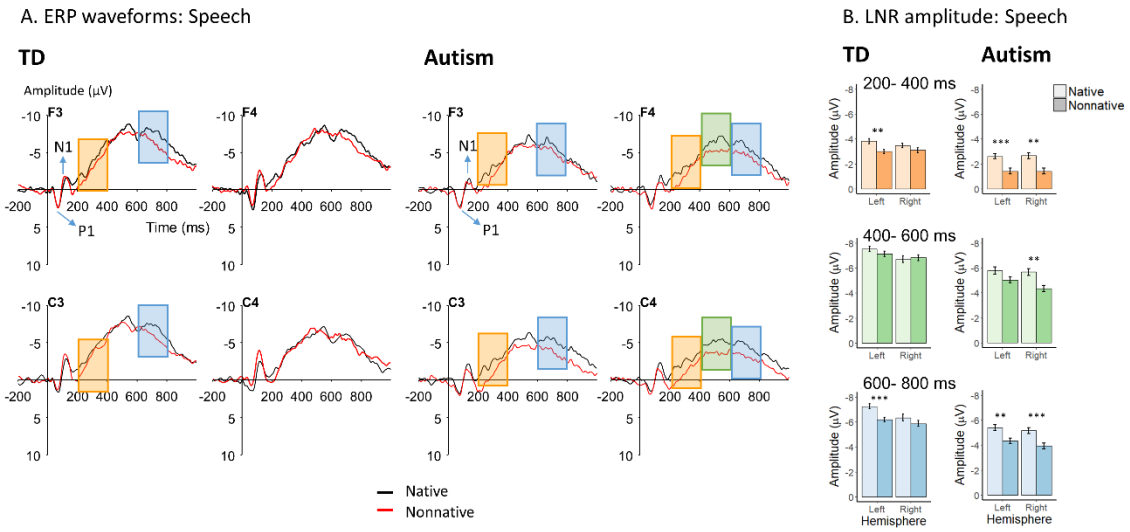

Figure 14. **A.** ERP waveforms for speech at the left (F3/C3) and right (F4/C4) electrode sites in the two groups. In each group, LNR windows with significant difference between native and nonnative speech are shaded (yellow, 200- 400 ms; green, 400-600 ms; blue, 600- 800 ms). **B.** Bar graphs summarizing the LNR results of native vs. nonnative speech distinction (** $p < .01$, *** $p < .001$). The effects in the TD group were restricted to the left scalp, whereas the autism had bilateral distributions with rightward asymmetry.

For the hum condition (Table 12, Fig. 14), LME showed hemisphere effect in the 400- 600 ms window with a leftward dominance ($F(1,316) = 7.81$, $p < .01$). There was a significant language effect in the 600- 800 ms window, with greater LNR for native than nonnative hums ($F(1,316) = 6.45$, $p < .05$). No other effect in the hum condition was found.

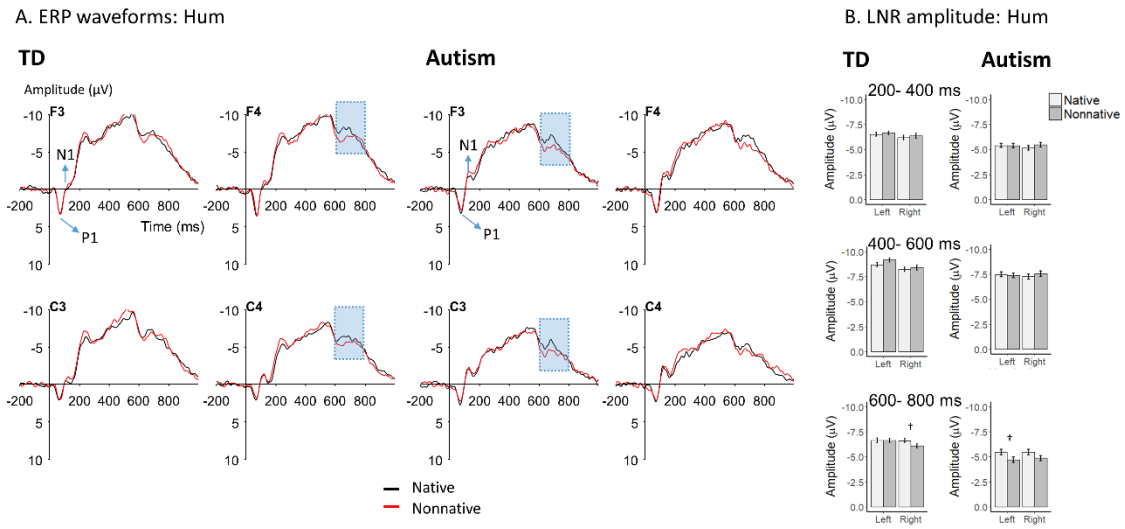

Figure 15. **A.** ERP waveforms for hum at the left (F3/C3) and right (F4/C4) electrode sites in the two groups. **B.** Bar graphs summarizing the LNR results of native vs. nonnative hum distinction († $p < .1$). Neither group had significant LNR difference between native vs. nonnative hum. The only trend was a rightward asymmetry in the TD group and a leftward asymmetry in the autism group.

Table 12.

LNR mean amplitude (μV) in the windows of interest.

Group	Hemi.	Language	200- 400 ms		400- 600 ms		600- 800 ms	
			Speech	Hum	Speech	Hum	Speech	Hum
TD	Left	Native	-3.82	-6.51	-7.51	-8.70	-7.23	-6.64
			(4.43)	(4.37)	(4.16)	(4.06)	(3.61)	(3.78)
		Nonnative	-2.99	-6.63	-7.14	-9.16	-6.18	-6.63
			(3.92)	(3.74)	(3.69)	(3.77)	(2.97)	(3.28)
	Right	Native	-3.48	-6.18	-6.70	-8.24	-6.34	-6.62
			(3.84)	(4.48)	(4.02)	(4.09)	(3.62)	(3.63)
		Nonnative	-3.10	-6.35	-6.82	-8.39	-5.87	-6.09
			(3.87)	(3.38)	(3.72)	(3.50)	(2.90)	(3.27)
Autism	Left	Native	-2.62	-5.39	-5.79	-7.48	-5.40	-5.47

		(3.51)	(3.75)	(4.32)	(3.75)	(3.41)	(3.72)
	Nonnative	-1.43	-5.39	-5.04	-7.37	-4.35	-4.72
		(4.40)	(3.75)	(3.68)	(3.94)	(3.36)	(3.02)
	Native	-2.67	-5.16	-5.66	-7.26	-5.15	-5.46
		(3.48)	(3.84)	(3.84)	(3.75)	(3.26)	(3.54)
Right		-1.43	-5.47	-4.33	-7.54	-3.93	-4.88
	Nonnative	(4.24)	(4.19)	(3.53)	(4.36)	(2.89)	(3.16)

SD in parenthesis

5.3.2 ITPC and Its Relationship with ERP Amplitude

In the speech condition (Table 10, Fig. 16), theta ITPC in the P1 window was affected by group*language interaction ($F(1,316) = 5.20$, $p < .05$). Post-hoc analyses showed a nonsignificant trend of greater theta synchrony for native speech in the autism group (TD, $F(1,172) = 1.87$, $p = .173$; Autism, $F(1,144) = 3.53$, $p = .062$). LME for the N1 window showed significant language effect that nonnative speech produced greater theta synchrony ($F(1,316) = 9.88$, $p < .01$). In both groups, theta ITPC was a significant predictor of P1 amplitude to speech (TD, $F(1,172) = 18.92$, $p < .001$, $\beta = 11.78$; Autism, $F(1,144) = 4.49$, $p < .05$, $\beta = 5.05$). The coefficient β indicates the more synchronized theta oscillation, the greater P1 amplitude. Such relationship was also present in the N1 window (TD, $F(1,172) = 45.32$, $p < .001$, $\beta = -17.97$; Autism, $F(1,144) = 28.14$, $p < .001$, $\beta = -13.73$).

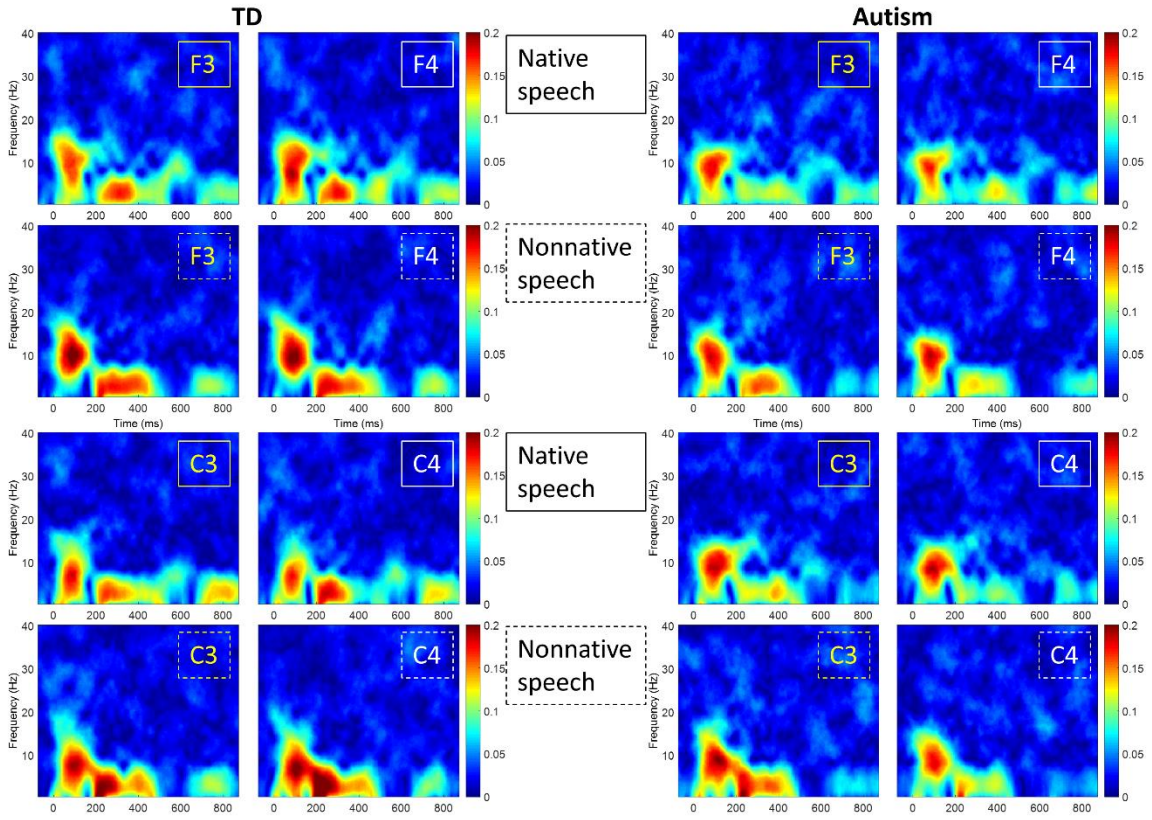

Figure 16. ITPC at 0.5- 40 Hz as a function of time in the **speech** conditions.

For the hum condition (Table 11, Fig. 17), no effect of group, language, hemisphere, or interaction in either P1 or N1 window was found. Nonetheless, in both groups, theta synchrony was associated with P1 amplitude (TD, $F(1,172) = 59.96$, $p < .001$, $\beta = 16.65$; Autism, $F(1,144) = 36.28$, $p < .001$, $\beta = 15.30$). In the N1 window, only the autism group showed ITPC-amplitude relationship (TD, $F(1,172) = 0.53$, $p = .469$, $\beta = -1.89$; Autism, $F(1,144) = 22.78$, $p < .001$, $\beta = -12.59$).

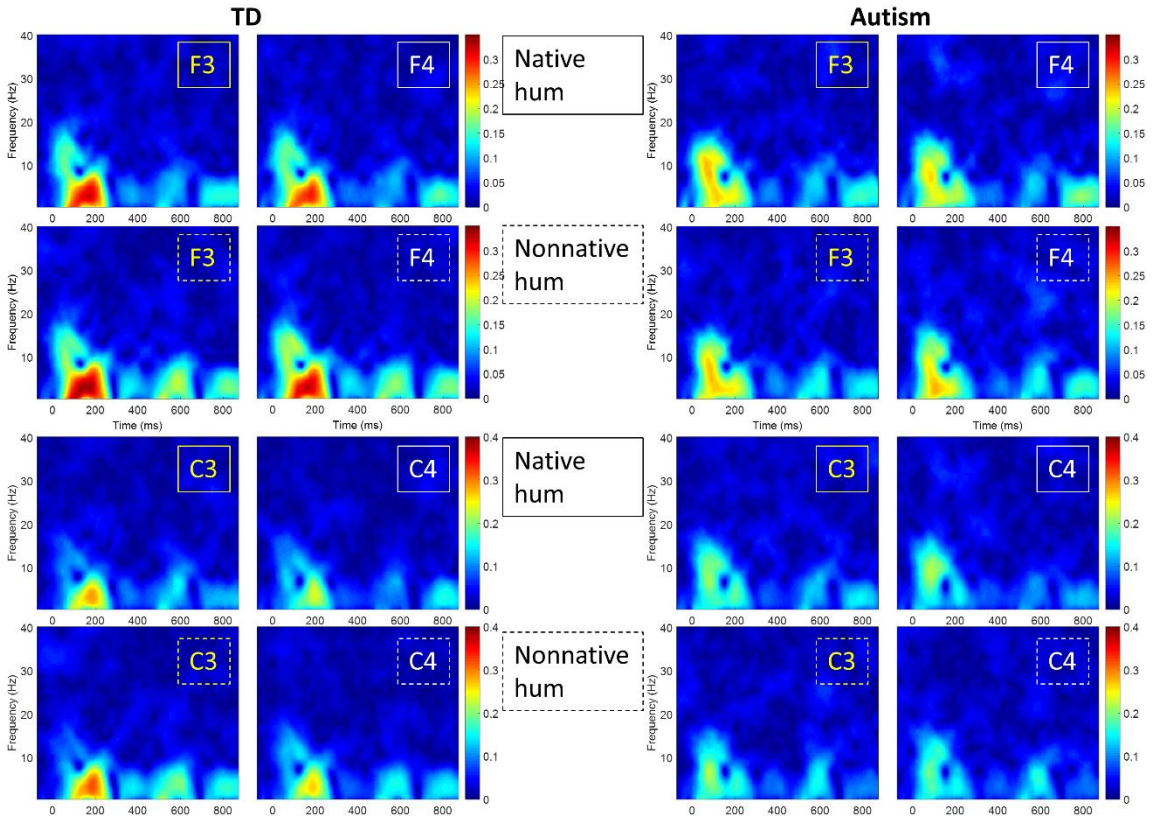

Figure 17. ITPC at 0.5- 40 Hz as a function of time in the **hum** conditions.

5.4 Discussion

This study provided some novel evidence of the neural coding of prosody and the time course in children with autism. The following discussion will be focused on the early auditory responses and the late responses separately.

5.4.1 P1 and N1 findings

5.4.1.1 Morphological Characteristics

The supra-syllabic nonsense speech elicited adult-like N1 in both the TD children and children with autism. The preceding P1 in these children was absent in the adults' waveforms of Study 3 (Chapter 4). This is consistent with ERP maturation data showing that auditory P1 reaches its maximum around the age of 9 years and decreases in

amplitude and latency through adulthood (Čeponienė, Rinne, et al., 2002; Sharma et al., 1997).

From the waveforms, we can see that N1 for the hum stimuli was largely diminished in the TD group but not in the autism group. This TD finding is consistent with the adult data as well. Instead, an N2 or N250 response was visible for hum in the TD group, resembling the child P1-N2 sequence in Study 1 (Chapter 2) with simple syllabic-level stimuli. In other words, sensory ERP data for the TD group indicates differential response patterns to hum and speech stimuli, whereas ERPs for the autism group indicates similar response patterns to the two types of stimuli. Moreover, hum failed to elicit P2 in the adults in Study 3 (Chapter 4), suggesting some overlapping function between child N1 and adult P2.

The N1 latency to speech but not to hum was delayed in the autism group. This speech-specific delay is consistent with the general trend of speech-specific deficits in autism when stimulus complexity is controlled for speech and nonspeech (Wang et al., 2017; Yu et al., 2015). Also, it has been found that children's N1 elicited by longer ISI overlaps in time window with their N2 elicited by shorter ISI (Čeponienė, Rinne, et al., 2002). The visible trend of diminished N1 to speech in the autism group is therefore coherent with the diminished N2 in Study 1, perhaps reflecting a persistent deficit in feature analysis during speech perception beyond initial sound onset detection reflected in P1 (P50) (Diesch, 1996; Tavabi et al., 2007; Tremblay et al., 2001).

5.4.1.2 Early Auditory Distinction of Native and Nonnative Prosody

The TD group displayed later and larger P1 for the native speech than for the nonnative speech, but earlier P1 for the native hum than for the nonnative hum. However,

these modulation effects on P1 were absent in the autism group. P1 (or P50) is associated with pre-attentive arousal and sound onset detection (H. Pratt, Starr, Michalewski, Bleich, & Mittelman, 2008; Rao, Zhang, & Miller, 2010). The P1 enhancement could be attributed to the fact that familiar voice pattern of native speech is more salient for the auditory system (Sidtis & Kreiman, 2012). The differential P1 latency effect for speech and hum might indicate the difference in the thalamocortical propagation speed that sound with sharper acoustic onset is transmitted faster to the auditory cortex, a similar pattern seen in Study 3 (Chapter 4) with adults' N1 amplitude. The autism group's lack of sensitivity in P1 to these features might suggest lack of selective arousal to the familiar voice and neural timing deficits in the ascending auditory pathway.

The TD group had smaller N1 amplitude for the native speech than for the nonnative speech, a pattern consistent with the adult N1 in Study 3 (Chapter 4). Both groups' N1 latency was longer for the native speech than for the nonnative speech but overall the autism group had delayed N1 to speech but not to hum compared to the TD children. Together, P1 and N1 in the autism group were less sensitive to the native and nonnative speech distinction compared to the TD group. In other words, the auditory system in these children with autism might register native and nonnative sounds as equally salient and allocate the same amount of neural resources in their acoustic and initial phonetic analysis. This may indicate a greater reliance on acoustic-level processing rather than linguistic-level processing.

5.4.2 LNR: Phonological Abstraction of Native Prosodic Patterns

The effect of group*language interaction was strikingly different from those in the earlier P1 and N1. In fact, LNR difference between the native and nonnative speech

was consistently greater in the autism group than in the TD group. Moreover, the TD group' LNR effects were confined to the left electrode sites whereas the autism group displayed bilateral effects with a rightward asymmetry. Consistent with the adult data, LNR did not differ between native and nonnative sounds in the hum conditions in either group, further suggesting that the functional significance of LNR is speech-specific.

The LNR difference was enhanced in the autism group for prosodic phonology but not prosodic acoustics compared to the TD group. Such domain-specificity suggests that the enhancement in autism was unlikely merely a result of heightened spectral resolution for pitch varying prosodies, but some other mechanisms involved for speech perception. Because each native Chinese token was composed of two monosyllabic words even though the whole utterance was meaningless, there could be syllable-by-syllable semantic processing involved. One possibility is that children with autism detect the semantic significance of each syllable in a hyper-reactive manner (Yu et al., 2015, pg 9-10). The enhancement for native prosody might be a neural signature of the stereotypical expressive intonation in individuals with autism, whereas the current study is not capable of answering this question.

Another major finding is the spatial distribution of LNR effects with leftward dominance in the TD group but bilateral distribution in autism. Leftward asymmetry in the neural processing of linguistic stimuli has been considered a hallmark of language-specific neural specialization (Zatorre & Gandour, 2008), and is linked with language proficiency in young children (Kuhl et al., 2013; Mills, Plunkett, et al., 2005). Phonological development of TD children is characterized with a substantial lexical advance during the second year of life (McMurray, 2007), which takes place in parallel

with the re-organization of brain activities in a way that familiar words are favorably processed by the left hemisphere (Mills et al., 1997; Mills, Plunkett, et al., 2005). Although at earlier developmental stages, the right hemisphere is preferred for sub-lexical phonetic information before any related linguistic systemization is acquired (Mills et al., 1997; Seery et al., 2013), the leftward migration of cortical activity from bilaterally distributed activity is thought to mark the qualitative change in word experience (Mills, Conboy, et al., 2005).

As introduced earlier in the chapter, lack of leftward asymmetry for language has been well documented with young children with autism (Coffey-Corina et al., 2008; Kuhl et al., 2013). The current study utilizing nonsense speech resulted in a similar pattern as those using lexical stimuli, indicating that disrupted neural specialization for language in autism affects not only the lexical level but also the phonological level of linguistic processing. Taken together, the absence of left hemispheric asymmetry accompanied by diffused cortical activation in children with autism might reflect disrupted or delayed phonological systemization at the supra-syllabic level. It might also reflect an overly bottom-up, compensatory mechanism of "acoustic form" instead of a phonological mode of cortical processing of speech sounds in children with autism.

5.5 Conclusions

Our current observation on native vs. nonnative prosodic phonology distinction provides supporting evidence to the notion of inadequate neural specialization for linguistic pitch pattern. This finding also suggests that the atypical cortical responses exist at school-age for children with formal literacy education and abundant language exposure.

Chapter 6: General Discussion & Conclusions

This dissertation outlines the characteristics of auditory and speech processing of pitch in autism in terms of neural specialization for language. This chapter will provide an overview of findings in the aspect of basic auditory processing and the aspect of language-specific patterns. Limitation and future directions will also be discussed.

6.1 On the Basic Auditory Processing

A consistent observation in Study 1, 2 &4 with children with autism was attenuated ERP amplitude in later but not earlier windows in the autism group compared to the TD group (Table 13). The pattern suggests that in children with autism, increased earlier sensory response was more likely to be followed by later response attenuation. Further exploration in the time-frequency domain suggests that trial-by-trial theta phase synchrony played a role in the ERP amplitude increment or reduction but to a limited extent, which is only partially supportive of the system-wise neural unreliability theory for autism (Haigh, 2018). Indeed, Butler, Molholm, Andrade, and Foxe (2017) examined trial-by-trial visual and somatosensory oscillation synchrony and provided strong counter-evidence for the neural unreliability hypothesis. Nonetheless, it needs to be mentioned that Butler et al tested children with autism with average IQ and only used physically simple stimuli. That said, the intact neural variability and ERP in those children with autism may not apply to the many children on the spectrum with intellectual disabilities when receiving more complex stimulation. Our observation from the various auditory stimuli offers some additional view for inter-trial variability and its relationship with ERP in children with autism that information processing stage must be taken into account. Specifically, network level neural synchronization in children with

autism could be especially vulnerable at later stages during which neuronal populations work together at larger temporal and spatial scales (Hahamy, Behrmann, & Malach, 2015; Simon & Wallace, 2016; Uhlhaas & Singer, 2007). In other words, variable responsivity across neural populations might be causing the overall underpowered ERP in autism despite having normal trial-by-trial synchrony compared with the TD group.

Table 13.

Results Summary of ERP amplitude and theta ITPC in the autism group in comparison with the TD group.

Study	P50		P1		N1		N2 or N250		LNR	
	Amp	ITPC	Amp	ITPC	Amp	ITPC	Amp	ITPC	Amp	ITPC
1	NA	NA	+	+	NA	NA	-	-	NA	NA
2	=	+	-	=	NA	NA	-	=	NA	NA
4	NA	NA	=	=	=	=	NA	NA	-	NA

+ autism > TD, - autism <TD, = no difference, NA not applicable

Another pattern emerged from the data is delayed response latency (P1 in Study 1, N1 in Study 4, and N250 in Study 3) in the autism group compared with the TD group. This finding adds to the ample evidence of delayed auditory response in children with autism reported by previous EEG and MEG investigations (Bruneau et al., 2003; Bruneau, Roux, Adrien, & Barthelemy, 1999; Donkers et al., 2015; Edgar, Khan, et al., 2015; Edgar et al., 2014; Gandal et al., 2010; Jenkins et al., 2016; Oram Cardy et al., 2008; Roberts et al., 2009; Roberts et al., 2010; Roberts et al., 2013). Such abnormality is not restricted to cortical activity but already pronounced in the auditory brainstem level. Wave-V in the auditory brainstem response (ABR) latency has been found significantly prolonged in children with autism and infants at risk for autism and is thought due to deficient myelination and enlarged head circumference in childhood (for a review,

(Miron, Beam, & Kohane, 2018). The current observation provided verification for the neural timing issue in the auditory domain, and that it not only affects the lower hierarchy in the auditory pathway reflected in the P1/N1 but also higher-order perceptual integration reflected in the N2 (or N250). As we mentioned in Chapter 2, AEP latency delay is not unequivocal regardless of stimulus and subject characteristics. Overviewing the literature and current research, it seems that P1/N1 delay is more reliably observed when the sound stimuli have clearer onset features (e.g., pure tone compared to speech) and N2 latency is more prone to complex stimuli.

6.2 On the Neural Specialization for Linguistic Pitch Pattern

Two studies were conducted to characterize the neural coding of linguistic pitch patterns in Chinese-acquiring children with autism. At the syllabic level, Study 2 (Chapter 3) examined whether neural responses in children with autism is tuned to the frequency trajectory that distinguishes a prototypical linguistically relevant pitch and a linguistically meaningless one. The result indicated that although pitch trajectory affected ERP above 250 ms similarly in the two groups, only the TD group but not the ASD group displayed response differentiation in the earlier window. Since the sole information carried by the complex noise stimuli was the pitch trajectory, any group difference can be attributed to the acoustic processing of pitch information. The result suggests delayed auditory separation of the curvilinear and linear tone in autism relative to the TD controls. It also implies that the auditory cortex of children with autism is less tuned for the critical acoustic feature that is linguistically relevant, potentially due to overall heightened responsivity for general pitch contour coding. Coincidentally, the critical window 180-230 ms we observed here overlaps with the previously reported lexical tone

MMN (Yu et al., 2015). If this critical window reflects autism group's greater contour coding ability, it would be consistent with the enhanced MMN for the nonspeech tonal contrast with rising vs. falling contours in autism in the Yu et al. study.

Study 4 explored the neural coding of larger linguistic units—disyllabic prosody. The results indicated that early ERP components in the autism group were less sensitive to native vs. nonnative voice distinction compared to the TD group. However, the LNR in these children with autism responded to the prosodic phonology difference in a hyperreactive manner with much more dispersed effects in the brain. The lack of early native vs. nonnative distinction in autism is in line with Study 2 (Chapter 3). The TD children's left-only distinction mirrors previous adult data of leftward lateralization for the processing of phonological structures such as native phonotactics (Rossi et al., 2011; Wagner et al., 2012). This observation along with the absence of such in for the hum conditions suggests the involvement of language network in these children. In contrast, the autism group' bilateral distribution with a rightward asymmetry is in agreement with the idea of impaired cortical specialization for linguistic processing associated with atypical right hemisphere dominance (Haesen et al., 2011; Kuhl et al., 2013; Sperdin & Schaer, 2016).

Together, the delayed auditory separation of prototypical vs. nonprototypical linguistic pitch patterns along with diffusive neural activity supports the notion of a lesser neural specialization in children with autism.

6.3 Limitations

This dissertation has several outstanding limitations. First, the samples were limited to school-age children. It is impossible to draw any solid conclusion

developmentally with only one age group, especially given the perceptual reorganization for tone in the first year of life (Mattock et al., 2008; Yeung et al., 2013). Although the current research provides evidence for reduced neural specialization for native sound in children with autism who have already received abundant linguistic exposure and literacy education, it remains unknown when the specialization process deviates significantly from that of TD children and how that may interact with the concurrent or subsequent development of core autism symptoms.

Second, the current studies lack the behavioral measures for lexical tone/prosody perception or production. Without the behavioral indicators, it is unknown whether the children's neurobehavioral atypicality align with potential alterations in the behavioral domain. However, behavioral measurements with children with intellectual disabilities can be problematic even for tasks without verbal instruction (Heaton, Hudry, Ludlow, & Hill, 2008; Heaton, Williams, Cummins, & Happe, 2008), as increased task demand for these children may introduce additional confound to the perceptual performance. Clearly, future work is in sore need to develop behavioral protocols that are suitable for children across a broader range of cognitive and adaptive function.

Third, the current studies do not have a non-Chinese sample to verify any potential cross-linguistic differences. For instance, if increased LNR difference to Chinese vs. English prosody in autism was due to enhanced syllable-to-syllable semantic processing of the Chinese prosodic structures, we would see a non-Chinese (e.g., English) group with autism, whose native prosody is not based on monosyllabic word unit, to not show such enhancement. Otherwise, it may be a cue-specific acoustic phenomenon.

6.4 Implications and Future Directions

Using auditory stimulation free from semantic or emotional meaning with passive listening paradigms, the current studies outlined the potentially disrupted neural specialization for language at sub-lexical level. It is important to discuss the impact these lower-level auditory phonological atypicalities may have on everyday listening experience of individuals with autism.

It is common to see children with autism cover their ears with the presence of environmental noise, which has been shown to negatively impact their experiences at school (Kanakri, Shepley, Varni, & Tassinary, 2017). Difficulty understanding speech in background noise and segregating one speech stream from another have also been documented with individuals with autism via behavioral testing (Alcantara et al., 2004; Bhatara, Babikian, Laugeson, Tachdjian, & Sininger, 2013; DePape, Hall, et al., 2012). A recent study discovered that hearing in noise lead to elevated recruitment of neural resources in autism compared to TD controls (Mamashli et al., 2017), a pattern similar to Study 4 (Chapter 5). In TD children, poorer speech-in-noise performance has been associated with delay in neural timing to speech sounds in the brainstem, particularly the processing stage for transcribing onset features such as formant transition (Anderson, Skoe, Chandrasekaran, & Kraus, 2010; Hornickel, Chandrasekaran, Zecker, & Kraus, 2011). Robust speech-in-noise perception by native monolingual speakers compared to nonnative bilingual speakers is reliant on additional top-down modulation from specialized cortical networks for the native language (Bidelman & Dexter, 2015; Hickok & Poeppel, 2007). Together, we have reasons to believe that reduced neural specialization for language accompanied by neural timing issue might be underlying the difficulty coping with complex auditory environment, especially in the extraction of

meaning from multiple information sources. Exhaustion of neural resources in coping with complex auditory information can be detrimental for dealing with concurrent social communicative demand, as normal social brain (right hemisphere) functions might be "squeezed" out, a possibility also suggested by Eyler et al. (2012).

Since optimal and simple listening condition is not always guaranteed for children with autism and for the sake of facilitating language acquisition, it is extremely important to explore interventions to improve their neural specialization for linguistic structures from as young as possible. For this purpose, the search for neural and behavioral markers and a critical period of neural specialization for native speech should be continued.

Autism is not only a cortical level condition but also subcortically affected, as evidenced by delayed transient ABR latency in infants and children with autism (Cohen et al., 2013; Miron et al., 2018; Russo, Nicol, Trommer, Zecker, & Kraus, 2009), although there have been opposing evidence of brainstem involvement as well (Grillon, Courchesne, & Akshoomoff, 1989; Klin, 1991). Brainstem frequency following response (FFR) plays a key role in transcribing sound f0 information and is known to be shaped by language experience (Krishnan & Gandour, 2017). FFR is phase-locked to the sound waveform periodicity, which offers even finer temporal precision than cortical EEG. A recent study has reported a relationship between "clearer" FFR representation of lexical tone f0 and greater speech perception efficiency in native Chinese speakers but not in English speakers, suggesting a subcortical element of language-specific neural specialization (Yu & Zhang, 2018). These findings provide the prerequisite of studying subcortical neural specialization for lexical tone as well as language-nonspecific pitch

encoding and how it may impact subsequent cortical response and behavioral outcome in autism.

To address the outstanding limitations of the current research, future work needs to expand the sample age and language background in order to be more developmentally and linguistically informative. Although Study 3 adult and Study 4 child data suggest that phonological processing at the prosodic level is still not fully maturated at school age, it is necessary to draw a detailed developmental timeline of phonological development for autism which is in extreme need. It has been found that the acquisition of word forms may precede or at least parallel the acquisition of individual phonemes in TD infants (Bergelson & Swingley, 2012), and that perceptual re-organization for tonemes in Chinese-acquiring infants may precede that for vowels and consonants (Yeung et al., 2013). These important findings provide the potential referential timepoints that we may anchor from in future research. For example, Seery et al. (2013) showed counterevidence for delayed perceptual reorganization in the 1st year of life in high-risk infants with autism. However, to the best of our knowledge, no autism research has investigated tone or prosody perception in infancy. Unfortunately, developmental study with only one age section is quite common in autism research, presumably due to limited resources for data collection and access to large samples. One way to bypass this obstacle is to utilize online data repositories, such as the National Institute of Mental Health (NIMH) Data Archive (NDA) and Autism Brain Imaging Database Exchange (ABIDE).

6.5 Conclusions

The current studies utilized auditory stimuli varying in linguistic relevance (native/prototypical vs. nonnative/unprototypical) to reveal specialized neural response

for linguistic structures. Overall, the autism group displayed hyposensitivity for stimulus differentiation at the earlier processing stage followed by hypersensitivity at the later processing stage in comparison with the TD group. The results suggest potentially disrupted neural specialization for linguistic auditory patterns accompanied by neural timing issues. Fine-scale neural indicators of auditory and phonological function may supplement measures of language and social function, as altered sound processing could significantly impact the ability to extract meaning from auditory input especially in complex listening situations. Language-specific neural measures as such can apply to infants and children with severe impairment, thus have the potential to aid diagnostic evaluation and assessment of intervention efficacy.

References

Alcantara, J. I., Cope, T. E., Cope, W., & Weisblatt, E. J. (2012). Auditory temporal-envelope processing in high-functioning children with Autism Spectrum Disorder. *Neuropsychologia, 50*(7), 1235-1251. doi: 10.1016/j.neuropsychologia.2012.01.034

Alcantara, J. I., Weisblatt, E. J., Moore, B. C. J., & Bolton, P. (2004). Speech-in-noise perception in high-functioning individuals with autism orAspergers syndrome. *Journal of Child Psychology and Psychiatry, 45*(6), 1107-1114.

Almeqbel, A., & McMahon, C. (2015). Objective measurement of high-level auditory cortical function in children. *Int J Pediatr Otorhinolaryngol, 79*(7), 1055-1062. doi: 10.1016/j.ijporl.2015.04.026

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (4th ed.)* Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)* Washington, DC: American Psychiatric Association.

Anderson, S., Skoe, E., Chandrasekaran, B., & Kraus, N. (2010). Neural timing is linked to speech perception in noise. *J Neurosci, 30*(14), 4922-4926. doi: 10.1523/JNEUROSCI.0107-10.2010

Arciuli, J., & Slowiaczek, L. M. (2007). The where and when of linguistic word-level prosody. *Neuropsychologia, 45*(11), 2638-2642. doi: 10.1016/j.neuropsychologia.2007.03.010

Barry, R. (2003). Preferred EEG brain states at stimulus onset in a fixed interstimulus interval auditory oddball task, and their effects on ERP components. *International Journal of Psychophysiology, 47*(3), 187-198. doi: 10.1016/s0167-8760(02)00151-4

Bastiaansen, M., Mazaheri, A., & Jensen, O. (2012). Beyond ERPs: Oscillatory Neuronal Dynamics. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 31-50): Oxford University Press.

Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proc Natl Acad Sci U S A, 109*(9), 3253-3258. doi: 10.1073/pnas.1113380109

Bhatara, A., Babikian, T., Laugeson, E., Tachdjian, R., & Sininger, Y. S. (2013). Impaired timing and frequency discrimination in high-functioning autism spectrum disorders. *J Autism Dev Disord, 43*(10), 2312-2328. doi: 10.1007/s10803-013-1778-y

Bidelman, G. M., & Dexter, L. (2015). Bilinguals at the "cocktail party": dissociable neural activity in auditory-linguistic brain regions reveals neurobiological basis for nonnative listeners' speech-in-noise recognition deficits. *Brain Lang, 143*, 32-41. doi: 10.1016/j.bandl.2015.02.002

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. . *Cereb Cortex, 10*, 512-528.

Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat Neurosci, 8*(3), 389-395. doi: 10.1038/nn1409

Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer (Version 5. 3. 79)[Computer program]. Retrieved from http://www.praat.org/.

Bonneh, Y. S., Levanon, Y., Dean-Pardo, O., Lossos, L., & Adini, Y. (2011). Abnormal speech spectrum and increased pitch variability in young autistic children. *Front Hum Neurosci, 4*, 237. doi: 10.3389/fnhum.2010.00237

Bonnel, A., McAdams, S., Smith, B., Berthiaume, C., Bertone, A., Ciocca, V., . . . Mottron, L. (2010). Enhanced pure-tone pitch discrimination among persons with autism but not Asperger syndrome. *Neuropsychologia, 48*(9), 2465-2475. doi: 10.1016/j.neuropsychologia.2010.04.020

Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A. (2003). Enhanced Pitch Sensitivity in Individuals with Autism: A Signal Detection Analysis. *Journal of Cognitive Neuroscience, 15*(2), 226-235.

Brechmann, A., Baumgart, F., & Scheich, H. (2002). Sound-Level-Dependent Representation of Frequency Modulations in Human Auditory Cortex: A Low-Noise fMRI Study. *J Neurophysiol, 87*, 423-433.

Brodeur, D. A., Gordon Green, C., Flores, H., & Burack, J. A. (2014). Time estimation among low-functioning individuals with autism spectrum disorders: evidence of poor sensitivity to variability of short durations. *Autism Res, 7*(2), 237-244. doi: 10.1002/aur.1364

Bruneau, N., Bonnet-Brilhault, F., Gomot, M., Adrien, J. L., & Barthelemy, C. (2003). Cortical auditory processing and communication in children with autism: electrophysiological/behavioral relations. *Int J Psychophysiol, 51*(1), 17-25. doi: 10.1016/s0167-8760(03)00149-1

Bruneau, N., Roux, S., Adrien, J. L., & Barthelemy, C. (1999). Auditory associative cortex dysfunction in children with autism: evidence from late auditory evoked potentials (N1 wave-T complex). *Clinical Neurophysiology, 110*(11), 1927-1934. doi: Doi 10.1016/S1388-2457(99)00149-2

Buard, I., Rogers, S. J., Hepburn, S., Kronberg, E., & Rojas, D. C. (2013). Altered oscillation patterns and connectivity during picture naming in autism. *Front Hum Neurosci, 7*, 742. doi: 10.3389/fnhum.2013.00742

Buchwald, J. S., Erwin, R. D., Van Lancker, D., Guthrie, D., Schwafel, J., & Tanguay, P. (1992). Midlatency auditory evoked responses: PI abnormalities in adult autistic subjects *Electroencephalography and clinical Neurophysiology, 84*, 164-171.

Buchwald, J. S., Rubinstein, E. H., Schwafel, J., & Strandburg, R. J. (1991). Midlatency auditory evoked responses: differential effects of a cholinergic agonist and antagonist. *Electroencephalogr Clin Neurophysiol, 80*(4), 303-309.

Butler, J. S., Molholm, S., Andrade, G. N., & Foxe, J. J. (2017). An Examination of the Neural Unreliability Thesis of Autism. *Cereb Cortex, 27*(1), 185-200. doi: 10.1093/cercor/bhw375

Caron, K. G., Schaaf, R. C., Benevides, T. W., & Gal, E. (2012). Cross-cultural comparison of sensory behaviors in children with autism. *Am J Occup Ther, 66*(5), e77-80. doi: 10.5014/ajot.2012.004226

Čeponienė, R., Alku, P., Westerfield, M., Torki, M., & Townsend, J. (2005). ERPs differentiate syllable and nonphonetic sound processing in children and adults. *Psychophysiology, 42*(4), 391-406. doi: 10.1111/j.1469-8986.2005.00305.x

Čeponienė, R., Cheour, M., & Näätänen, R. (1998). Interstimulus interval and auditory event-related potentials in children: evidence for multiple generators. *Electroencephalogr Clin Neurophysiol, 108*(4), 345-354.

Čeponienė, R., Cummings, A., Wulfeck, B., Ballantyne, A., & Townsend, J. (2009). Spectral vs. temporal auditory processing in specific language impairment: a developmental ERP study. *Brain Lang, 110*(3), 107-120. doi: 10.1016/j.bandl.2009.04.003

Čeponienė, R., Lepistö, T., Shestakova, A., Vanhala, R., Alku, P., Näätänen, R., & Yaguchi, K. (2003). Speech-sound-selective auditory impairment in children with autism: they can perceive but do not attend. *Proc Natl Acad Sci U S A, 100*(9), 5567-5572. doi: 10.1073/pnas.0835631100

Čeponienė, R., Rinne, T., & Näätänen, R. (2002). Maturation of cortical sound processing as indexed by event-related potentials. *Clinical Neurophysiology, 113*(6), 870-882. doi: 10.1016/s1388-2457(02)00078-0

Čeponienė, R., Shestakova, A., Balan, P., Alku, P., Yiaguchi, K., & Näätänen, R. (2001). Children's auditory event-related potentials index sound complexity and "speechness". *Int J Neurosci, 109*(3-4), 245-260.

Čeponienė, R., Torki, M., Alku, P., Koyama, A., & Townsend, J. (2008). Event-related potentials reflect spectral differences in speech and non-speech stimuli in children and adults. *Clin Neurophysiol, 119*(7), 1560-1577. doi: 10.1016/j.clinph.2008.03.005

Čeponienė, R., Yaguchi, K., Shestakova, A., Alku, P., Suominen, K., & Näätänen, R. (2002). Sound complexity and 'speechness' effects on pre-attentive auditory discrimination in children. *International Journal of Psychophysiology, 43*(3), 199-211.

Chan, K. K., & To, C. K. (2016). Do Individuals with High-Functioning Autism Who Speak a Tone Language Show Intonation Deficits? *J Autism Dev Disord, 46*(5), 1784-1792. doi: 10.1007/s10803-016-2709-5

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007a). Experience-dependent neural plasticity is sensitive to shape of pitch contours. *Neuroreport, 18*(18), 1963-1967.

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007b). Mismatch negativity to pitch contours is influenced by language experience. *Brain Res, 1128*(1), 148-156. doi: 10.1016/j.brainres.2006.10.064

Chrabaszcz, A., Winn, M., Lin, C. Y., & Idsardi, W. J. (2014). Acoustic cues to perception of word stress by English, Mandarin, and Russian speakers. *J Speech Lang Hear Res, 57*(4), 1468-1479. doi: 10.1044/2014_JSLHR-L-13-0279

Coffey-Corina, S., Padden, D., & Kuhl, P. K. (2008). ERPs to words correlate with behavioral measures in children with Autism Spectrum Disorder. *The Journal of the Acoustical Society of America, 123*(5).

Cohen, I. L., Gardner, J. M., Karmel, B. Z., Phan, H. T., Kittler, P., Gomez, T. R., . . . Barone, A. (2013). Neonatal brainstem function and 4-month arousal-modulated attention are jointly associated with autism. *Autism Res, 6*(1), 11-22. doi: 10.1002/aur.1259

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology, 1*(1), 42-45. doi: 10.20982/tqmp.01.1.p042

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clin Neurophysiol, 115*(4), 732-744. doi: 10.1016/j.clinph.2003.11.021

Cuccaro, M. L., Brinkley, J., Abramson, R. K., Hall, A., Wright, H. H., Hussman, J. P., . . . Pericak-Vance, M. A. (2007). Autism in African American families: clinical-phenotypic findings. *Am J Med Genet B Neuropsychiatr Genet, 144B*(8), 1022-1026. doi: 10.1002/ajmg.b.30535

Cunningham, J., Nicol, T., Zecker, S., & Kraus, N. (2000). Speech-evoked neurophysiologic responses in children with learning problems: development and behavioral correlates of perception. *Ear Hear, 21*(6), 554-568.

Dawson, M., Soulieres, I., Gernsbacher, M. A., & Mottron, L. (2007). The level and nature of autistic intelligence. *Psychol Sci, 18*(8), 657-662. doi: 10.1111/j.1467-9280.2007.01954.x

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208*(4448), 1174-1176.

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behav Dev, 9*(2), 133-150.

Dehaene-Lambertz, G., Dupoux, E., & Gout, A. (2000). Electrophysiological correlates of phonological processing: A cross-linguistic study.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods, 134*(1), 9-21. doi: 10.1016/j.jneumeth.2003.10.009

Depape, A. M., Chen, A., Hall, G. B., & Trainor, L. J. (2012). Use of prosody and information structure in high functioning adults with autism in relation to language ability. *Front Psychol, 3*, 72. doi: 10.3389/fpsyg.2012.00072

DePape, A. M., Hall, G. B., Tillmann, B., & Trainor, L. J. (2012). Auditory processing in high-functioning adolescents with Autism Spectrum Disorder. *PLoS ONE, 7*(9), e44084. doi: 10.1371/journal.pone.0044084

Diamond, E., & Zhang, Y. (2016). Cortical processing of phonetic and emotional information in speech: A cross-modal priming study. *Neuropsychologia, 82*, 110-122. doi: 10.1016/j.neuropsychologia.2016.01.019

Diehl, J. J., Watson, D., Bennetto, L., McDonough, J., & Gunlogson, C. (2009). An acoustic analysis of prosody in high-functioning autism. *Applied Psycholinguistics, 30*(03), 385. doi: 10.1017/s0142716409090201

Diesch, E. (1996). The Neurotopography of Vowels as Mirrored by Evoked Magnetic Field Measurements. *Brain and Language, 53*, 143-168.

Digester, F. M., Wohlberedt, T., & Hoppe, U. (2009). Contribution of Spectrotemporal Features on Auditory Event-Related Potentials Elicited by Consonant-Vowel Syllables. *Ear and Hearing, 30*, 704-712.

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage, 85 Pt 2*, 761-768. doi: 10.1016/j.neuroimage.2013.06.035

Donkers, F. C., Schipul, S. E., Baranek, G. T., Cleary, K. M., Willoughby, M. T., Evans, A. M., . . . Belger, A. (2015). Attenuated auditory event-related potentials and associations with atypical sensory response patterns in children with autism. *J Autism Dev Disord, 45*(2), 506-523. doi: 10.1007/s10803-013-1948-y

Edgar, J. C., Fisk Iv, C. L., Berman, J. I., Chudnovskaya, D., Liu, S., Pandey, J., . . . Roberts, T. P. (2015). Auditory encoding abnormalities in children with autism spectrum disorder suggest delayed development of auditory cortex. *Mol Autism, 6*, 69. doi: 10.1186/s13229-015-0065-5

Edgar, J. C., Khan, S. Y., Blaskey, L., Chow, V. Y., Rey, M., Gaetz, W., . . . Roberts, T. P. (2015). Neuromagnetic oscillations predict evoked-response latency delays and core language deficits in autism spectrum disorders. *J Autism Dev Disord, 45*(2), 395-405. doi: 10.1007/s10803-013-1904-x

Edgar, J. C., Lanza, M. R., Daina, A. B., Monroe, J. F., Khan, S. Y., Blaskey, L., . . . Roberts, T. P. (2014). Missing and delayed auditory responses in young and older children with autism spectrum disorders. *Front Hum Neurosci, 8*, 417. doi: 10.3389/fnhum.2014.00417

Edwards, E., Soltani, M., Kim, W., Dalal, S. S., Nagarajan, S. S., Berger, M. S., & Knight, R. T. (2009). Comparison of time-frequency responses and the event-related potential to auditory speech stimuli in human cortex. *J Neurophysiol, 102*(1), 377-386. doi: 10.1152/jn.90954.2008

Eggermont, J. J., Ponton, C. W., Don, M., Waring, M. D., & Kwong, B. (1997). Maturational delays in cortical evoked potentials in cochlear implant users. *Acta Otolaryngol, 117*(2), 161-163. doi: 10.3109/00016489709117760

Eigsti, I. M., Schuh, J., Mencl, E., Schultz, R. T., & Paul, R. (2012). The neural underpinnings of prosody in autism. *Child Neuropsychol, 18*(6), 600-617. doi: 10.1080/09297049.2011.639757

Elangovan, S., & Stuart, A. (2011). A cross-linguistic examination of cortical auditory evoked potentials for a categorical voicing contrast. *Neurosci Lett, 490*(2), 140-144. doi: 10.1016/j.neulet.2010.12.044

Eyler, L. T., Pierce, K., & Courchesne, E. (2012). A failure of left temporal cortex to specialize for language is an early emerging and fundamental property of autism. *Brain, 135*(Pt 3), 949-960. doi: 10.1093/brain/awr364

Ferri, R., Elia, M., Agarwal, N., Lanuzza, B., Musumeci, S. A., & Pennisi, G. (2003). The mismatch negativity and the P3a components of the auditory event-related potentials in autistic low-functioning subjects. *Clin Neurophysiol, 114*(9), 1671-1680. doi: 10.1016/s1388-2457(03)00153-6

Filipe, M. G., Frota, S., Castro, S. L., & Vicente, S. G. (2014). Atypical prosody in Asperger syndrome: perceptual and acoustic measurements. *J Autism Dev Disord, 44*(8), 1972-1981. doi: 10.1007/s10803-014-2073-2

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., & Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Exp Brain Res, 203*(2), 381-389. doi: 10.1007/s00221-010-2240-4

Freeth, M., Sheppard, E., Ramachandran, R., & Milne, E. (2013). A cross-cultural comparison of autistic traits in the UK, India and Malaysia. *J Autism Dev Disord, 43*(11), 2569-2583. doi: 10.1007/s10803-013-1808-9

Friedrich, M., & Friederici, A. D. (2005). Phonotactic knowledge and lexical-semantic processing in one-year-olds: brain responses to words and nonsense words in picture contexts. *J Cogn Neurosci, 17*(11), 1785-1802. doi: 10.1162/089892905774589172

Friedrich, M., & Friederici, A. D. (2015). The origins of word learning: Brain responses of 3-month-olds indicate their rapid association of objects and words. *Dev Sci*. doi: 10.1111/desc.12357

Fuentemilla, L., Marco-Pallares, J., & Grau, C. (2006). Modulation of spectral power and of phase resetting of EEG contributes differentially to the generation of auditory event-related potentials. *NeuroImage, 30*(3), 909-916. doi: 10.1016/j.neuroimage.2005.10.036

Fujioka, T., Ross, B., Kakigi, R., Pantev, C., & Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain, 129*(Pt 10), 2593-2608. doi: 10.1093/brain/awl247

Gage, N. M., Siegel, B., Callen, M., & Roberts, T. P. (2003). Cortical sound processing in children with autism disorder: an MEG investigation. *Neuroreport, 14*(16), 2047-2051. doi: 10.1097/01.wnr.0000090030.46087.4a

Gandal, M. J., Edgar, J. C., Ehrlichman, R. S., Mehta, M., Roberts, T. P., & Siegel, S. J. (2010). Validating gamma oscillations and delayed auditory responses as translational biomarkers of autism. *Biol Psychiatry, 68*(12), 1100-1106. doi: 10.1016/j.biopsych.2010.09.031

Gandour, J., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y., . . . Lowe, M. (2004). Hemispheric roles in the perception of speech prosody. *NeuroImage, 23*(1), 344-357. doi: 10.1016/j.neuroimage.2004.06.004

Garcia-Larrea, L., Lukaszewicz, A., & Mauguiére, F. (1992). Revisiting the oddball paradigm. Non-target vs neutral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia, 30*(8), 723-741.

Gebauer, L., Skewes, J., Horlyck, L., & Vuust, P. (2014). Atypical perception of affective prosody in Autism Spectrum Disorder. *Neuroimage Clin, 6*, 370-378. doi: 10.1016/j.nicl.2014.08.025

Gervain, J., & Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annu Rev Psychol, 61*, 191-218. doi: 10.1146/annurev.psych.093008.100408

Gillberg, C. (1991). Outcome in Autism and Autistic-like Conditions. *J Am Acad Child Adolesc Psychiatry, 30*(3), 375-382.

Gillespie-Lynch, K., Sepeta, L., Wang, Y., Marshall, S., Gomez, L., Sigman, M., & Hutman, T. (2012). Early childhood predictors of the social competence of adults with autism. *J Autism Dev Disord, 42*(2), 161-174. doi: 10.1007/s10803-011-1222-0

Gilley, P. M., Sharma, A., Dorman, M., & Martin, K. (2005). Developmental changes in refractoriness of the cortical auditory evoked potential. *Clin Neurophysiol, 116*(3), 648-657. doi: 10.1016/j.clinph.2004.09.009

Gilliam, J. E. (2006). *Gilliam Autism Rating Scale - Second Edition*. Austin, TX: PRO-ED.

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

Gomot, M., Blanc, R., Clery, H., Roux, S., Barthelemy, C., & Bruneau, N. (2011). Candidate electrophysiological endophenotypes of hyper-reactivity to change in autism. *J Autism Dev Disord, 41*(6), 705-714. doi: 10.1007/s10803-010-1091-y

Gomot, M., Giard, M. H., Adrien, J. L., Barthelemy, C., & Bruneau, N. (2002). Hypersensitivity to acoustic change in children with autism: electrophysiological evidence of left frontal cortex dysfunctioning. *Psychophysiology, 39*(5), 577-584. doi: 10.1017.S0048577202394058

Green, H., & Tobin, Y. (2009). Prosodic analysis is difficult … but worth it: A study in high functioning autism. *Int J Speech Lang Pathol, 11*(4), 308-315. doi: 10.1080/17549500903003060

Grillon, C., Courchesne, E., & Akshoomoff, N. (1989). Brainstem and middle latency auditory evoked potentials in autism and developmental language disorder. *J Autism Dev Disord, 19*(2), 255-269.

Groen, W. B., van Orsouw, L., Huurne, N., Swinkels, S., van der Gaag, R. J., Buitelaar, J. K., & Zwiers, M. P. (2009). Intact spectral but abnormal temporal processing of auditory stimuli in autism. *J Autism Dev Disord, 39*(5), 742-750. doi: 10.1007/s10803-008-0682-3

Grunwald, T., Boutros, N. N., Pezer, N., von Oertzen, J., Fernandez, G., Schaller, C., & Elger, C. E. (2003). Neuronal substrates of sensory gating within the human brain. *Biological Psychiatry, 53*(6), 511-519. doi: 10.1016/S0002-3223(03)01673-2

Gueorguieva, R., & Krystal, J. H. (2004a). Move over anova: progress in analyzing repeated-measures data andits reflection in papers published in the archives of general psychiatry. *Archives of general psychiatry, 61*(3), 310-317.

Gueorguieva, R., & Krystal, J. H. (2004b). Move over anova: progress in analyzing repeated-measures data andits reflection in papers published in the archives of general psychiatry. *Archives of general psychiatry, 61*(3), 310-317.

Haesen, B., Boets, B., & Wagemans, J. (2011). A review of behavioural and electrophysiological studies on auditory processing and speech perception in autism spectrum disorders. *Res Autism Spect Dis, 5*(2), 701-714. doi: 10.1016/j.rasd.2010.11.006

Hahamy, A., Behrmann, M., & Malach, R. (2015). The idiosyncratic brain: distortion of spontaneous connectivity patterns in autism spectrum disorder. *Nat Neurosci, 18*(2), 302-309. doi: 10.1038/nn.3919

Haigh, S. M. (2018). Variable sensory perception in autism. *Eur J Neurosci, 47*(6), 602-609. doi: 10.1111/ejn.13601

Hall, D., Johnsrude, I. S., Haggard, M. P., Palmer, A. R., Akeroyd, M. A., & Summerfield, A. Q. (2002). Spectral and Temporal Processing in Human Auditory Corte. *Cereb Cortex, 12,* 140-149.

Heaton, P. (2003). Pitch memory, labelling and disembedding in autism. *J Child Psychol Psychiatry, 44*(4), 543-551.

Heaton, P. (2005). Interval and Contour Processing in Autism. *Journal of Autism and Developmental Disorders, 35*(6), 787-793. doi: 10.1007/s10803-005-0024-7

Heaton, P., Hermelin, B., & Pring, L. (1998). Autism and Pitch Processing: A Precursor for Savant Musical Ability? . *Music Perception, 15(3)*, 291-305

Heaton, P., Hudry, K., Ludlow, A., & Hill, E. (2008). Superior discrimination of speech pitch and its relationship to verbal ability in autism spectrum disorders. *Cogn Neuropsychol, 25*(6), 771-782. doi: 10.1080/02643290802336277

Heaton, P., Williams, K., Cummins, O., & Happe, F. (2008). Autism and pitch processing splinter skills: a group and subgroup analysis. *Autism, 12*(2), 203-219. doi: 10.1177/1362361307085270

Hegarty, J. P., 2nd, Gu, M., Spielman, D. M., Cleveland, S. C., Hallmayer, J. F., Lazzeroni, L. C., . . . Hardan, A. Y. (2018). A proton MR spectroscopy study of the thalamus in twins with autism spectrum disorder. *Prog Neuropsychopharmacol Biol Psychiatry, 81*, 153-160. doi: 10.1016/j.pnpbp.2017.09.016

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nat Rev Neurosci, 8*(5), 393-402. doi: 10.1038/nrn2113

Hoekert, M., Bais, L., Kahn, R. S., & Aleman, A. (2008). Time course of the involvement of the right anterior superior temporal gyrus and the right fronto-parietal operculum in emotional prosody perception. *PLoS ONE, 3*(5), e2244. doi: 10.1371/journal.pone.0002244

Hornickel, J., Chandrasekaran, B., Zecker, S., & Kraus, N. (2011). Auditory brainstem measures predict reading and speech-in-noise perception in school-aged children. *Behav Brain Res, 216*(2), 597-605. doi: 10.1016/j.bbr.2010.08.051

Howard, M. A., & Volkov, L. O. (2000). Auditory Cortex on the Human Posterior Superior Temporal Gyrus.

Howlin, P. (2005). Outcomes in Autism Spectrum Disorders. In F. R. Volkmar, R. Paul, A. Klin & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders*.

Huang, D., Yu, L., Wang, X., Fan, Y., Wang, S., & Zhang, Y. (2018). Distinct patterns of discrimination and orienting for temporal processing of speech and nonspeech in Chinese children with autism: an event-related potential study. *Eur J Neurosci, 47*(6), 662-668. doi: 10.1111/ejn.13657

Hubbard, D. J., Faso, D. J., Assmann, P. F., & Sasson, N. J. (2017). Production and perception of emotional prosody by adults with autism spectrum disorder. *Autism Res, 10*(12), 1991-2001. doi: 10.1002/aur.1847

Hyde, K. L., Peretz, I., & Zatorre, R. J. (2008). Evidence for the role of the right auditory cortex in fine pitch resolution. *Neuropsychologia, 46*(2), 632-639. doi: 10.1016/j.neuropsychologia.2007.09.004

Isler, J. R., Martien, K. M., Grieve, P. G., Stark, R. I., & Herbert, M. R. (2010). Reduced functional connectivity in visual evoked potentials in children with autism spectrum disorder. *Clin Neurophysiol, 121*(12), 2035-2043. doi: 10.1016/j.clinph.2010.05.004

Jarvinen-Pasley, A., & Heaton, P. (2007). Evidence for reduced domain-specificity in auditory processing in autism. *Dev Sci, 10*(6), 786-793. doi: 10.1111/j.1467-7687.2007.00637.x

Jarvinen-Pasley, A., Pasley, J., & Heaton, P. (2008). Is the linguistic content of speech less salient than its perceptual features in autism? *J Autism Dev Disord, 38*(2), 239-248. doi: 10.1007/s10803-007-0386-0

Jarvinen-Pasley, A., Peppe, S., King-Smith, G., & Heaton, P. (2008). The relationship between form and function level receptive prosodic abilities in autism. *J Autism Dev Disord, 38*(7), 1328-1340. doi: 10.1007/s10803-007-0520-z

Jarvinen-Pasley, A., Wallace, G. L., Ramus, F., Happe, F., & Heaton, P. (2008). Enhanced perceptual processing of speech in autism. *Dev Sci, 11*(1), 109-121. doi: 10.1111/j.1467-7687.2007.00644.x

Jenkins, J., 3rd, Chow, V., Blaskey, L., Kuschner, E., Qasmieh, S., Gaetz, L., . . . Roberts, T. P. (2016). Auditory Evoked M100 Response Latency is Delayed in Children with 16p11.2 Deletion but not 16p11.2 Duplication. *Cereb Cortex, 26*(5), 1957-1964. doi: 10.1093/cercor/bhv008

Jiang, J., Liu, F., Wan, X., & Jiang, C. (2015). Perception of Melodic Contour and Intonation in Autism Spectrum Disorder: Evidence From Mandarin Speakers. *J Autism Dev Disord, 45*(7), 2067-2075. doi: 10.1007/s10803-015-2370-4

Jochaut, D., Lehongre, K., Saitovitch, A., Devauchelle, A. D., Olasagasti, I., Chabane, N., . . . Giraud, A. L. (2015). Atypical coordination of cortical oscillations in response to speech in autism. *Front Hum Neurosci, 9*, 171. doi: 10.3389/fnhum.2015.00171

Jones, E. J., Gliga, T., Bedford, R., Charman, T., & Johnson, M. H. (2014). Developmental pathways to autism: a review of prospective studies of infants at risk. *Neurosci Biobehav Rev, 39*, 1-33. doi: 10.1016/j.neubiorev.2013.12.001

Jones, S. J., & Perez, N. (2001). The auditory 'C-process': analyzing the spectral envelope of complex sounds. *Clinical Neurophysiology, 112*(6), 965- 975.

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 Neurosci, 9*, 53. doi: 10.1186/1471-2202-9-53

Kanakri, S. M., Shepley, M., Varni, J. W., & Tassinary, L. G. (2017). Noise and autism spectrum disorder in children: An exploratory survey. *Res Dev Disabil, 63*, 85-94. doi: 10.1016/j.ridd.2017.02.004

Karhu, J., Herrgård, E., Pääkkönen, A., Luoma, L., Airaksinen, E., & Partanen, J. (1997). Dual cerebral processing of elementary auditory input in children. *Neuroreport, 8*(6), 1327-1330.

Kasai, K., Hashimoto, O., Kawakubo, Y., Yumoto, M., Kamio, S., Itoh, K., . . . Kato, N. (2005). Delayed automatic detection of change in speech sounds in adults with autism: a magnetoencephalographic study. *Clin Neurophysiol, 116*(7), 1655-1664. doi: 10.1016/j.clinph.2005.03.007

Keating, P., & Kuo, G. (2012). Comparison of speaking fundamental frequency in English and Mandarin. *J Acoust Soc Am, 132*(2), 1050-1060. doi: 10.1121/1.4730893

Khan, S., Hashmi, J. A., Mamashli, F., Bharadwaj, H. M., Ganesan, S., Michmizos, K. P., . . . Kenet, T. (2016). Altered Onset Response Dynamics in Somatosensory Processing in Autism Spectrum Disorder. *Front Neurosci, 10*, 255. doi: 10.3389/fnins.2016.00255

King, K. A., Campbell, J., Sharma, A., Martin, K., Dorman, M., & Langran, J. (2008). The representation of voice onset time in the cortical auditory evoked potentials of young children. *Clin Neurophysiol, 119*(12), 2855-2861. doi: 10.1016/j.clinph.2008.09.015

Klimesch, W., Sauseng, P., Hanslmayr, S., Gruber, W., & Freunberger, R. (2007). Event-related phase reorganization may explain evoked neural dynamics. *Neurosci Biobehav Rev, 31*(7), 1003-1016. doi: 10.1016/j.neubiorev.2007.03.005

Klin, A. (1991). Auditory Brainstem Responses in Autism: Brainstem Dysfunction or Peripheral Hearing Loss? *J Autism Dev Disord*.

Koerner, T. K., & Zhang, Y. (2015). Effects of background noise on inter-trial phase coherence and auditory N1-P2 responses to speech stimuli. *Hear Res, 328*, 113-119. doi: 10.1016/j.heares.2015.08.002

Koerner, T. K., & Zhang, Y. (2017). Application of Linear Mixed-Effects Models in Human Neuroscience Research: A Comparison with Pearson Correlation in Two Auditory Electrophysiology Studies. *Brain Sci, 7*(3), 26. doi: 10.3390/brainsci7030026

Koerner, T. K., Zhang, Y., Nelson, P. B., Wang, B., & Zou, H. (2016). Neural indices of phonemic discrimination and sentence-level speech intelligibility in quiet and noise: A mismatch negativity study. *Hear Res, 339*, 40-49. doi: 10.1016/j.heares.2016.06.001

Koh, H. C., & Milne, E. (2012). Evidence for a cultural influence on field-independence in autism spectrum disorder. *J Autism Dev Disord, 42*(2), 181-190. doi: 10.1007/s10803-011-1232-y

Korpilahti, P., Jansson-Verkasalo, E., Mattila, M.-L., Kuusikko, S., Suominen, K., Rytky, S., . . . Moilanen, I. (2006). Processing of Affective Speech Prosody is Impaired in Asperger Syndrome. *Journal of Autism and Developmental Disorders, 37*(8), 1539-1549. doi: 10.1007/s10803-006-0271-2

Korzyukov, O., Pflieger, M. E., Wagner, M., Bowyer, S. M., Rosburg, T., Sundaresan, K., . . . Boutros, N. N. (2007). Generators of the intracranial P50 response in auditory sensory gating. *NeuroImage, 35*(2), 814-826. doi: 10.1016/j.neuroimage.2006.12.011

Kreitewolf, J., Friederici, A. D., & von Kriegstein, K. (2014). Hemispheric lateralization of linguistic prosody recognition in comparison to speech and speaker recognition. *NeuroImage, 102 Pt 2*, 332-344. doi: 10.1016/j.neuroimage.2014.07.038

Krishnan, A., & Gandour, J. T. (2017). Shaping Brainstem Representation of Pitch-Relevant Information by Language Experience. In N. Kraus, S. Anderson, T. White-Schwoch, R. R. Fay & A. N. Popper (Eds.), *The Frequency-Following Response: A Window into Human Communication* (pp. 45-73). Cham: Springer International Publishing.

Krishnan, A., Gandour, J. T., Smalt, C. J., & Bidelman, G. M. (2010). Language-dependent pitch encoding advantage in the brainstem is not limited to acceleration rates that occur in natural speech. *Brain Lang, 114*(3), 193-198. doi: 10.1016/j.bandl.2010.05.004

Krishnan, A., Gandour, J. T., & Suresh, C. H. (2014). Cortical pitch response components show differential sensitivity to native and nonnative pitch contours. *Brain Lang, 138*, 51-60. doi: 10.1016/j.bandl.2014.09.005

Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron, 67*(5), 713-727. doi: 10.1016/j.neuron.2010.08.038

Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: behavioral and electrophysiological measures. *Dev Sci, 8*(1), F1-F12.

Kuhl, P. K., Coffey-Corina, S., Padden, D., Munson, J., Estes, A., & Dawson, G. (2013). Brain responses to words in 2-year-olds with autism predict developmental outcomes at age 6. *PLoS ONE, 8*(5), e64967. doi: 10.1371/journal.pone.0064967

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.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants at 6 months of age. *Science, 255*.

Kujala, T. (2007). The role of early auditory discrimination deficits in language disorders. *Journal of Psychophysiology, 21*(3-4), 239-250. doi: 10.1027/0269-8803.21.34.239

Kujala, T., Lepistö, T., & Näätänen, R. (2013). The neural basis of aberrant speech and audition in autism spectrum disorders. *Neurosci Biobehav Rev, 37*(4), 697-704. doi: 10.1016/j.neubiorev.2013.01.006

Kujala, T., Lepisto, T., Nieminen-von Wendt, T., Näätänen, P., & Näätänen, R. (2005). Neurophysiological evidence for cortical discrimination impairment of prosody in Asperger syndrome. *Neurosci Lett, 383*(3), 260-265. doi: 10.1016/j.neulet.2005.04.048

Kwakye, L. D., Foss-Feig, J. H., Cascio, C. J., Stone, W. L., & Wallace, M. T. (2011). Altered auditory and multisensory temporal processing in autism spectrum disorders. *Front Integr Neurosci, 4*, 129. doi: 10.3389/fnint.2010.00129

Langers, D. R. M., Backes, W. H., & Dijk, P. v. (2003). Spectrotemporal features of the auditory cortex: the activation in response to dynamic ripples. *NeuroImage, 20*(1), 265-275. doi: 10.1016/s1053-8119(03)00258-1

Lehmann, D., & Skrandies, W. (1984). Spatial analysis of evoked potentials in man--a review. *Progress in Neurobiology, 23*(3), 227-250.

Lenth, R. V. (2016). Least-Squares Means: TheR Package lsmeans. *Journal of Statistical Software, 69*(1). doi: 10.18637/jss.v069.i01

Lepisto, T., Kajander, M., Vanhala, R., Alku, P., Huotilainen, M., Näätänen, R., & Kujala, T. (2008). The perception of invariant speech features in children with autism. *Biol Psychol, 77*(1), 25-31. doi: 10.1016/j.biopsycho.2007.08.010

Lepistö, T., Kujala, T., Vanhala, R., Alku, P., Huotilainen, M., & Näätänen, R. (2005). The discrimination of and orienting to speech and non-speech sounds in children with autism. *Brain Res, 1066*(1-2), 147-157. doi: 10.1016/j.brainres.2005.10.052

Lepistö, T., Silokallio, S., Nieminen-von Wendt, T., Alku, P., Näätänen, R., & Kujala, T. (2006). Auditory perception and attention as reflected by the brain event-related

potentials in children with Asperger syndrome. *Clin Neurophysiol, 117*(10), 2161-2171. doi: 10.1016/j.clinph.2006.06.709

Liegeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalogr Clin Neurophysiol, 92*(3), 204-214.

Lindell, A. K., & Hudry, K. (2013). Atypicalities in cortical structure, handedness, and functional lateralization for language in autism spectrum disorders. *Neuropsychol Rev, 23*(3), 257-270. doi: 10.1007/s11065-013-9234-5

Lindstrom, R., Lepisto-Paisley, T., Vanhala, R., Alen, R., & Kujala, T. (2016). Impaired neural discrimination of emotional speech prosody in children with autism spectrum disorder and language impairment. *Neurosci Lett, 628*, 47-51. doi: 10.1016/j.neulet.2016.06.016

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front Hum Neurosci, 8*, 213. doi: 10.3389/fnhum.2014.00213

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2001). *Autism diagnostic observation schedule*. Los Angeles, CA: Western Psychological Services.

Luck, S. J. (2014). *An introduction to the event-related potential technique*: MIT Press.

Malekmohammadi, M., Elias, W. J., & Pouratian, N. (2015). Human thalamus regulates cortical activity via spatially specific and structurally constrained phase-amplitude coupling. *Cereb Cortex, 25*(6), 1618-1628. doi: 10.1093/cercor/bht358

Mamashli, F., Khan, S., Bharadwaj, H., Michmizos, K., Ganesan, S., Garel, K. A., . . . Kenet, T. (2017). Auditory processing in noise is associated with complex patterns of disrupted functional connectivity in autism spectrum disorder. *Autism Res, 10*(4), 631-647. doi: 10.1002/aur.1714

Mampe, B., Friederici, A. D., Christophe, A., & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Curr Biol, 19*(23), 1994-1997. doi: 10.1016/j.cub.2009.09.064

Martin, B. A., Kurtzberg, D., & Stapells, D. R. (1999). The Effects of Decreased Audibility Produced by High-Pass Noise Masking on N1 and the Mismatch Negativity to Speech Sounds /ba/ and /da/. *Journal of Speech, Language, and Hearing Research, 42*(271-286).

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. doi: 10.1016/j.cognition.2007.07.002

McCann, J., & Peppe, S. (2003). Prosody in autism spectrum disorders: a critical review. *Int J Lang Commun Disord, 38*(4), 325-350. doi: 10.1080/1368282031000154204

McCann, J., Peppe, S., Gibbon, F. E., O'Hare, A., & Rutherford, M. (2007). Prosody and its relationship to language in school-aged children with high-functioning autism. *Int J Lang Commun Disord, 42*(6), 682-702. doi: 10.1080/13682820601170102

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science, 317*(5838), 631. doi: 10.1126/science.1144073

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29*(2), 143-178.

Meng, X., Sai, X., Wang, C., Wang, J., Sha, S., & Zhou, X. (2005). Auditory and speech processing and reading development in Chinese school children: behavioural and ERP evidence. *Dyslexia, 11*(4), 292-310.

Mills, D. L., Coffey-Corina, S. A., & Neville, H. J. (1993). Language Acquisition and Cerebral Specialization in 20-Month-Old Inhts. *Journal of Cognitive Neuroscience, 5*(3), 317-334.

Mills, D. L., CoffeyCorina, S., & Neville, H. J. (1997). Language comprehension and cerebral specialization from 13 to 20 months. *Developmental Neuropsychology, 13*(3), 397-445. doi: 10.1080/87565649709540685

Mills, D. L., Conboy, B. T., & Paton, C. (2005). Do Changes in Brain Organization Reflect Shifts in Symbolic Functioning? *Symbol use and symbolic representation* (pp. 123-153).

Mills, D. L., Plunkett, K., Prat, C., & Schafer, G. (2005). Watching the infant brain learn words: effects of vocabulary size and experience. *Cognitive Development, 20*(1), 19-31. doi: 10.1016/j.cogdev.2004.07.001

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

Milne, E. (2011). Increased intra-participant variability in children with autistic spectrum disorders: evidence from single-trial analysis of evoked EEG. *Front Psychol, 2*, 51. doi: 10.3389/fpsyg.2011.00051

Miron, O., Beam, A. L., & Kohane, I. S. (2018). Auditory brainstem response in infants and children with autism spectrum disorder: A meta-analysis of wave V. *Autism Res, 11*(2), 355-363. doi: 10.1002/aur.1886

Morgan, J. L., & Demuth, K. (1996). Signal to syntax: An overview. *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, 1-22.

Morillon, B., Lehongre, K., Frackowiak, R. S. J., Ducorps, A., Kleinschmidt, A., Poeppel, D., & Giraud, A. L. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proc Natl Acad Sci U S A, 107*(43), 18688–18693.

Mottron, L., Bouvet, L., Bonnel, A., Samson, F., Burack, J. A., Dawson, M., & Heaton, P. (2013). Veridical mapping in the development of exceptional autistic abilities. *Neurosci Biobehav Rev, 37*(2), 209-228. doi: 10.1016/j.neubiorev.2012.11.016

Mottron, L., Dawson, M., Soulieres, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception. *J Autism Dev Disord, 36*(1), 27-43. doi: 10.1007/s10803-005-0040-7

Munson, B., Edwards, J., & Beckman, M. E. (2012). Phonological Representationsin Language Acquisition: Climbing The Ladder of Abstraction. In A. Cohn, C. Fougeron & M. Huffman (Eds.), *The Oxford Handbook of Laboratory Phonology*: Oxford University Press.

Munson, B., Edwards, J., Beckman, M. E., Cohn, A., Fougeron, C., & Huffman, M. (2011). Phonological representations in language acquisition: Climbing the ladder of abstraction. *Handbook of laboratory phonology*, 288-309.

Näätäneiv, R., Lehtokoski, A., Lennest, M., Luuki, A., Alliki, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385*, 432-434.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and brain sciences, 13*(02), 201-233. doi: 10.1017/s0140525x00078407

Näätänen, R., Kujala, T., Escera, C., Baldeweg, T., Kreegipuu, K., Carlson, S., & Ponton, C. (2012). The mismatch negativity (MMN)--a unique window to disturbed central auditory processing in ageing and different clinical conditions. *Clin Neurophysiol, 123*(3), 424-458. doi: 10.1016/j.clinph.2011.09.020

Näätänen, R., Kujala, T., & Winkler, I. (2011). Auditory processing that leads to conscious perception: a unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology, 48*(1), 4-22. doi: 10.1111/j.1469-8986.2010.01114.x

Näätänen, R., & Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic Response to Sound: A Review and an Analysis of the Component Structure. *Psychophysiology, 24*(4), 375-425. doi: 10.1111/j.1469-8986.1987.tb00311.x

Nadig, A., & Shaw, H. (2012). Acoustic and perceptual measurement of expressive prosody in high-functioning autism: increased pitch range and what it means to listeners. *J Autism Dev Disord, 42*(4), 499-511. doi: 10.1007/s10803-011-1264-3

Oram Cardy, J. E., Flagg, E. J., Roberts, W., & Roberts, T. P. (2008). Auditory evoked fields predict language ability and impairment in children. *Int J Psychophysiol, 68*(2), 170-175. doi: 10.1016/j.ijpsycho.2007.10.015

Orekhova, E. V., Stroganova, T. A., Prokofiev, A. O., Nygren, G., Gillberg, C., & Elam, M. (2009). The right hemisphere fails to respond to temporal novelty in autism: evidence from an ERP study. *Clin Neurophysiol, 120*(3), 520-529. doi: 10.1016/j.clinph.2008.12.034

Orekhova, E. V., Stroganova, T. A., Prokofyev, A. O., Nygren, G., Gillberg, C., & Elam, M. (2008). Sensory gating in young children with autism: relation to age, IQ, and EEG gamma oscillations. *Neurosci Lett, 434*(2), 218-223. doi: 10.1016/j.neulet.2008.01.066

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

Peppe, S., McCann, J., Gibbon, F., O'Hare, A., & Rutherford, M. (2007). Receptive and Expressive Prosodic Ability in Children With High-Functioning Autism. *Journal of Speech, Language, and Hearing Research, 50*, 1015-1028.

Pereira, D. R., Cardoso, S., Ferreira-Santos, F., Fernandes, C., Cunha-Reis, C., Paiva, T. O., . . . Marques-Teixeira, J. (2014). Effects of inter-stimulus interval (ISI) duration on the N1 and P2 components of the auditory event-related potential. *Int J Psychophysiol, 94*(3), 311-318. doi: 10.1016/j.ijpsycho.2014.09.012

Phillips, C. (2001). Levels of representation in the electrophysiology of speech perception. *Cognitive Science, 25*, 711-731.

Ponton, C., Eggermont, J. J., Khosla, D., Kwong, B., & Don, M. (2002). Maturation of human central auditory system activity: separating auditory evoked potentials by dipole source modeling. *Clinical Neurophysiology, 113*(3), 407-420.

Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clin Neurophysiol, 111*(2), 220-236.

Pratt, H. (2012). Sensory ERP Components. In S. J. Luck & E. S. Kappenman (Eds.), *Oxford handbook of event-related potentials* (pp. 89-114).

Pratt, H., Starr, A., Michalewski, H. J., Bleich, N., & Mittelman, N. (2008). The auditory P50 component to onset and offset of sound. *Clin Neurophysiol, 119*(2), 376-387. doi: 10.1016/j.clinph.2007.10.016

Radicevic, Z., Vujovic, M., Jelicic, L., & Sovilj, M. (2008). Comparative findings of voice and speech: language processing at an early ontogenetic age in quantitative EEG mapping. *Exp Brain Res, 184*(4), 529-532. doi: 10.1007/s00221-007-1119-5

Ramus, F. (2014). Neuroimaging sheds new light on the phonological deficit in dyslexia. *Trends Cogn Sci, 18*(6), 274-275. doi: 10.1016/j.tics.2014.01.009

Rao, A., Zhang, Y., & Miller, S. (2010). Selective listening of concurrent auditory stimuli: an event-related potential study. *Hear Res, 268*(1-2), 123-132. doi: 10.1016/j.heares.2010.05.013

Raven, J., & Court, J. (1998). Raven manual: Section 3. standard progressive matrices: Oxford, England: Oxford Psychologists Press.

Roberts, T. P., Heiken, K., Zarnow, D., Dell, J., Nagae, L., Blaskey, L., . . . Edgar, J. C. (2014). Left hemisphere diffusivity of the arcuate fasciculus: influences of autism spectrum disorder and language impairment. *AJNR Am J Neuroradiol, 35*(3), 587-592. doi: 10.3174/ajnr.A3754

Roberts, T. P., Khan, S. Y., Blaskey, L., Dell, J., Levy, S. E., Zarnow, D. M., & Edgar, J. C. (2009). Developmental correlation of diffusion anisotropy with auditory-evoked response. *Neuroreport, 20*(18), 1586-1591. doi: 10.1097/WNR.0b013e3283306854

Roberts, T. P., Khan, S. Y., Rey, M., Monroe, J. F., Cannon, K., Blaskey, L., . . . Edgar, J. C. (2010). MEG detection of delayed auditory evoked responses in autism spectrum disorders: towards an imaging biomarker for autism. *Autism Res, 3*(1), 8-18. doi: 10.1002/aur.111

Roberts, T. P., Lanza, M. R., Dell, J., Qasmieh, S., Hines, K., Blaskey, L., . . . Berman, J. I. (2013). Maturational differences in thalamocortical white matter microstructure and auditory evoked response latencies in autism spectrum disorders. *Brain Res, 1537*, 79-85. doi: 10.1016/j.brainres.2013.09.011

Rossi, S., Jurgenson, I. B., Hanulikova, A., Telkemeyer, S., Wartenburger, I., & Obrig, H. (2011). Implicit processing of phonotactic cues: evidence from electrophysiological and vascular responses. *J Cogn Neurosci, 23*(7), 1752-1764. doi: 10.1162/jocn.2010.21547

Russo, N., Nicol, T., Trommer, B., Zecker, S., & Kraus, N. (2009). Brainstem transcription of speech is disrupted in children with autism spectrum disorders. *Dev Sci, 12*(4), 557-567. doi: 10.1111/j.1467-7687.2008.00790.x

Rutter, M., Bailey, A., & Lord, C. (2003). *The social communication questionnaire: Manual*: Western Psychological Services.

Samson, F., Hyde, K. L., Bertone, A., Soulieres, I., Mendrek, A., Ahad, P., . . . Zeffiro, T. A. (2011). Atypical processing of auditory temporal complexity in autistics. *Neuropsychologia, 49*(3), 546-555. doi: 10.1016/j.neuropsychologia.2010.12.033

Samson, F., Mottron, L., Jemel, B., Belin, P., & Ciocca, V. (2006). Can spectro-temporal complexity explain the autistic pattern of performance on auditory tasks? *J Autism Dev Disord, 36*(1), 65-76. doi: 10.1007/s10803-005-0043-4

Schonwiesner, M., Rubsamen, R., & von Cramon, D. Y. (2005). Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *Eur J Neurosci, 22*(6), 1521-1528. doi: 10.1111/j.1460-9568.2005.04315.x

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

Seery, A. M., Vogel-Farley, V., Tager-Flusberg, H., & Nelson, C. A. (2013). Atypical lateralization of ERP response to native and non-native speech in infants at risk for autism spectrum disorder. *Dev Cogn Neurosci, 5*, 10-24. doi: 10.1016/j.dcn.2012.11.007

Shahin, A. J., Roberts, L. E., Miller, L. M., McDonald, K. L., & Alain, C. (2007). Sensitivity of EEG and MEG to the N1 and P2 auditory evoked responses modulated by spectral complexity of sounds. *Brain Topogr, 20*(2), 55-61. doi: 10.1007/s10548-007-0031-4

Sharma, A., & Dorman, M. F. (1999). Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *J Acoust Soc Am, 106*(2), 1078-1083.

Sharma, A., & Dorman, M. F. (2000). Neurophysiologic correlates of cross-language phonetic perception. *The Journal of the Acoustical Society of America, 107*(5), 2697-2703. doi: 10.1121/1.428655

Sharma, A., Kraus, N., McGee, T. J., & Nicol, T. G. (1997). Developmental changes in P1 and N1 central auditory responses elicited by consonant-vowel syllables. *Electroencephalogr Clin Neurophysiol, 104*(6), 540-545.

Shen, X. S., Lin, M., & Yan, J. (1993). F0 turning point as an F0 cue to tonal contrast: A case study of Mandarin tones 2 and 3. *The Journal of the Acoustical Society of America, 93*(4), 2241-2243. doi: 10.1121/1.406688

Shu, H., Peng, H., & McBride-Chang, C. (2008). Phonological awareness in young Chinese children. *Developmental Science, 11*(1), 171-181. doi: 10.1111/j.1467-7687.2007.00654.x

Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proc Natl Acad Sci U S A, 108*(15), 6038-6043. doi: 10.1073/pnas.1017617108

Sidtis, D., & Kreiman, J. (2012). In the beginning was the familiar voice: personally familiar voices in the evolutionary and contemporary biology of communication. *Integr Psychol Behav Sci, 46*(2), 146-159. doi: 10.1007/s12124-011-9177-4

Siegal, M., & Blades, M. (2003). Language and auditory processing in autism. *Trends Cogn Sci, 7*(9), 378-380.

Simon, D. M., & Wallace, M. T. (2016). Dysfunction of sensory oscillations in Autism Spectrum Disorder. *Neurosci Biobehav Rev, 68*, 848-861. doi: 10.1016/j.neubiorev.2016.07.016

Singh, L., & Fu, C. S. (2016). A New View of Language Development: The Acquisition of Lexical Tone. *Child Dev, 87*(3), 834-854. doi: 10.1111/cdev.12512

Singh, L., Hui, T. J., Chan, C., & Golinkoff, R. M. (2014). Influences of vowel and tone variation on emergent word knowledge: a cross-linguistic investigation. *Dev Sci, 17*(1), 94-109. doi: 10.1111/desc.12097

Sittiprapaporn, W., Chindaduangratn, C., Tervaniemi, M., & Khotchabhakdi, N. (2003). Preattentive processing of lexical tone perception by the human brain as indexed by the mismatch negativity paradigm. *Annals of the New York Academy of Sciences, 999*(9), 199-203. doi: 10.1196/annals.1284.029

Smith, I. C., Reichow, B., & Volkmar, F. R. (2015). The Effects of DSM-5 Criteria on Number of Individuals Diagnosed with Autism Spectrum Disorder: A Systematic Review. *J Autism Dev Disord, 45*(8), 2541-2552. doi: 10.1007/s10803-015-2423-8

Sperdin, H. F., & Schaer, M. (2016). Aberrant Development of Speech Processing in Young Children with Autism: New Insights from Neuroimaging Biomarkers. *Front Neurosci, 10*, 393. doi: 10.3389/fnins.2016.00393

Stronach, S. T., & Wetherby, A. M. (2017). Observed and Parent-Report Measures of Social Communication in Toddlers With and Without Autism Spectrum Disorder Across Race/Ethnicity. *Am J Speech Lang Pathol, 26*(2), 355-368. doi: 10.1044/2016_AJSLP-15-0089

Sun, L., Grutzner, C., Bolte, S., Wibral, M., Tozman, T., Schlitt, S., . . . Uhlhaas, P. J. (2012). Impaired gamma-band activity during perceptual organization in adults with autism spectrum disorders: evidence for dysfunctional network activity in frontal-posterior cortices. *J Neurosci, 32*(28), 9563-9573. doi: 10.1523/JNEUROSCI.1073-12.2012

Sussman, E., Steinschneider, M., Gumenyuk, V., Grushko, J., & Lawson, K. (2008). The maturation of human evoked brain potentials to sounds presented at different stimulus rates. *Hear Res, 236*(1-2), 61-79. doi: 10.1016/j.heares.2007.12.001

Swaab, T. Y., Ledoux, K., Camblin, C. C., & Boudewyn, M. A. (2012). Language-Related ERP Components *Oxford handbook of event related potentials*.

Swaminathan, J., Krishnan, A., Gandour, J. T., & Xu, Y. (2008). Applications of static and dynamic iterated rippled noise to evaluate pitch encoding in the human auditory brainstem. *IEEE Trans Biomed Eng, 55*(1), 281-287. doi: 10.1109/TBME.2007.896592

Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. *Handbook of autism and pervasive developmental disorders, 1*, 335-364.

Tavabi, K., Obleser, J., Dobel, C., & Pantev, C. (2007). Auditory evoked fields differentially encode speech features: an MEG investigation of the P50m and N100m time courses during syllable processing. *Eur J Neurosci, 25*(10), 3155-3162. doi: 10.1111/j.1460-9568.2007.05572.x

Tong, Y., Gandour, J., Talavage, T., Wong, D., Dzemidzic, M., Xu, Y., . . . Lowe, M. (2005). Neural circuitry underlying sentence-level linguistic prosody. *NeuroImage, 28*(2), 417-428. doi: 10.1016/j.neuroimage.2005.06.002

Tong, Y., Melara, R. D., & Rao, A. (2009). P2 enhancement from auditory discrimination training is associated with improved reaction times. *Brain Res, 1297*, 80-88. doi: 10.1016/j.brainres.2009.07.089

Tremblay, K., & Kraus, N. (2002). Auditory Training Induces Asymmetrical Changes in Cortical Neural Activity. *Journal of Speech, Language, and Hearing Research, 45*, 564-572.

Tremblay, K., Kraus, N., Mcgee, T., Ponton, C., & Otis, B. (2001). Central Auditory Plasticity: Changes in the N1-P2 Complex after Speech Sound Training. *Ear and Hearing, 22*, 79-90.

Uhlhaas, P. J., & Singer, W. (2007). What do disturbances in neural synchrony tell us about autism? *Biol Psychiatry, 62*(3), 190-191. doi: 10.1016/j.biopsych.2007.05.023

van Noordt, S., Wu, J., Venkataraman, A., Larson, M. J., South, M., & Crowley, M. J. (2017). Inter-trial Coherence of Medial Frontal Theta Oscillations Linked to Differential Feedback Processing in Youth and Young Adults with Autism. *Res Autism Spectr Disord, 37*, 1-10. doi: 10.1016/j.rasd.2017.01.011

Wagner, M., Shafer, V. L., Martin, B., & Steinschneider, M. (2012). The phonotactic influence on the perception of a consonant cluster /pt/ by native English and native Polish listeners: a behavioral and event related potential (ERP) study. *Brain Lang, 123*(1), 30-41. doi: 10.1016/j.bandl.2012.06.002

Wagner, M., Shafer, V. L., Martin, B., & Steinschneider, M. (2013). The effect of native-language experience on the sensory-obligatory components, the P1-N1-P2 and the T-complex. *Brain Res, 1522*, 31-37. doi: 10.1016/j.brainres.2013.04.045

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., & Tojo, Y. (2006). The Autism-Spectrum Quotient (AQ) in Japan: A cross-cultural comparison. *J Autism Dev Disord, 36*(2), 263-270. doi: 10.1007/s10803-005-0061-2

Wang, X., Wang, S., Fan, Y., Huang, D., & Zhang, Y. (2017). Speech-specific categorical perception deficit in autism: An Event-Related Potential study of lexical tone processing in Mandarin-speaking children. *Sci Rep, 7*, 43254. doi: 10.1038/srep43254

Wechsler, D. (2003). *Wechsler intelligence scale for children-WISC-IV*: Psychological Corporation.

Werker, J. F., & Curtin, S. (2011). PRIMIR: A Developmental Framework of Infant Speech Processing. *Language Learning and Development, 1*(2), 197-234. doi: 10.1080/15475441.2005.9684216

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development, 7*(1), 49-63.

Whitehouse, A. J., & Bishop, D. V. (2008). Do children with autism 'switch off' to speech sounds? An investigation using event-related potentials. *Dev Sci, 11*(4), 516-524. doi: 10.1111/j.1467-7687.2008.00697.x

Winer, J. A., Miller, L. M., Lee, C. C., & Schreiner, C. E. (2005). Auditory thalamocortical transformation: structure and function. *Trends Neurosci, 28*(5), 255-263. doi: 10.1016/j.tins.2005.03.009

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics, 25*(1), 61-83. doi: 10.1006/jpho.1996.0034

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. doi: 10.1016/j.jml.2012.09.004

Yip, M. (2002). *Tone*: Cambridge University Press.
You, R. S., Serniclaes, W., Rider, D., & Chabane, N. (2017). On the nature of the speech perception deficits in children with autism spectrum disorders. *Res Dev Disabil, 61*, 158-171. doi: 10.1016/j.ridd.2016.12.009
Yu, L., Fan, Y., Deng, Z., Huang, D., Wang, S., & Zhang, Y. (2015). Pitch Processing in Tonal-Language-Speaking Children with Autism: An Event-Related Potential Study. *J Autism Dev Disord, 45*(11), 3656-3667. doi: 10.1007/s10803-015-2510-x
Yu, L., Rao, A., Zhang, Y., Burton, P. C., Rishiq, D., & Abrams, H. (2017). Neuromodulatory Effects of Auditory Training and Hearing Aid Use on Audiovisual Speech Perception in Elderly Individuals. *Front Aging Neurosci, 9*(30), 30. doi: 10.3389/fnagi.2017.00030
Yu, L., Wang, S., Huang, D., Wu, X., & Zhang, Y. (2018). Role of inter-trial phase coherence in atypical auditory evoked potentials to speech and nonspeech stimuli in children with autism. *Clin Neurophysiol, 129*(7), 1374-1382. doi: 10.1016/j.clinph.2018.04.599
Yu, L., & Zhang, Y. (2018). Testing native language neural commitment at the brainstem level: A cross-linguistic investigation of the association between frequency-following response and speech perception. *Neuropsychologia, 109*, 140-148. doi: 10.1016/j.neuropsychologia.2017.12.022
Zaehle, T., Jancke, L., & Meyer, M. (2007). Electrical brain imaging evidences left auditory cortex involvement in speech and non-speech discrimination based on temporal features. *Behav Brain Funct, 3*, 63. doi: 10.1186/1744-9081-3-63
Zaehle, T., Wustenberg, T., Meyer, M., & Jancke, L. (2004). Evidence for rapid auditory perception as the foundation of speech processing: a sparse temporal sampling fMRI study. *Eur J Neurosci, 20*(9), 2447-2456. doi: 10.1111/j.1460-9568.2004.03687.x
Zatorre, R. J., & Gandour, J. T. (2008). Neural specializations for speech and pitch: moving beyond the dichotomies. *Philos Trans R Soc Lond B Biol Sci, 363*(1493), 1087-1104. doi: 10.1098/rstb.2007.2161
Zhang, Y., Koerner, T., Miller, S., Grice-Patil, Z., Svec, A., Akbari, D., . . . Carney, E. (2011). Neural coding of formant-exaggerated speech in the infant brain. *Dev Sci, 14*(3), 566-581. doi: 10.1111/j.1467-7687.2010.01004.x
Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. (2005). Effects of language experience: neural commitment to language-specific auditory patterns. *NeuroImage, 26*(3), 703-720. doi: 10.1016/j.neuroimage.2005.02.040
Zhang, Y., & Wang, Y. (2007). Neural plasticity in speech acquisition and learning. *Bilingualism: Language and Cognition, 10*(2), 147-160. doi: 10.1017/S1366728907002908
Zhang, Y., Zhang, L., Shu, H., Xi, J., Wu, H., & Li, P. (2012). Universality of categorical perception deficit in developmental dyslexia: an investigation of Mandarin Chinese tones. *J Child Psychol Psychiatry, 53*(8), 874-882. doi: 10.1111/j.1469-7610.2012.02528.x
Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *Int J Dev Neurosci, 23*(2-3), 143-152. doi: 10.1016/j.ijdevneu.2004.05.001