

Chapter 13

Behavioral and Neurophysiological Evidence of Speech Processing in Chinese-Speaking Individuals with Autism Spectrum Disorder: A Review and Future Directions



Yan H. Yu and Valerie L. Shafer

Abstract Autism spectrum disorder (ASD) is a neurodevelopmental disorder that presents with core deficits in language and social communication areas. Past decades have witnessed a growing number of studies concerning this population's language and communication skills. However, studies focusing on Chinese-speaking individuals with ASD are rare and have just begun to accumulate. This review focuses on prosody and lexical tone perception and production in Chinese-speaking individuals with ASD. We also briefly review the evidence from general ASD literature for cross-language comparisons. Similar to patterns seen in many non-tonal language speakers with ASD, Chinese-speaking individuals with ASD generally demonstrate atypical pitch in terms of both average and range of values in verbal productions. Behavioral and neurophysiological evidence suggest atypicality, such as enhanced lower-level auditory processing and reduced higher-level linguistic processing in Chinese-speaking individuals with ASD. We also report some preliminary neural intervention data on bilingual English–Mandarin-learning children with ASD. Future directions on advancing theory and practice are discussed.

13.1 Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder with core deficits in social interaction, language, and communication, as defined by the International Classification of Diseases and Related Health Problems, tenth edition (WHO, 1992),

Y. H. Yu (✉)

Department of Communication Sciences and Disorders, St. John's University, Queens, NY, USA

e-mail: yuy1@stjohns.edu; yanhyu@gmail.com

V. L. Shafer

Speech-Language-Hearing Science, The Graduate Center, City University of New York, New York, NY, USA

e-mail: vshafer@gc.cuny.edu

© Springer Nature Singapore Pte Ltd. 2020

H.-M. Liu et al. (eds.), *Speech Perception, Production and Acquisition*,

Chinese Language Learning Sciences,

https://doi.org/10.1007/978-981-15-7606-5_13

and the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5, American Psychiatric Association, 2013). As the name suggests, individuals with such a “spectrum disorder” show a wide range of symptoms, ranging from mild to severe. The DSM-5 has abandoned the use of subcategorized diagnoses; however, researchers sometimes use subcategorical terms, such as “high-functioning autism,” “low-functioning autism,” and “Asperger’s syndrome,” especially in publications prior to the DSM-5. High-functioning autism (HFA) is an unofficial term used to describe the milder forms of ASD. Individuals with HFA usually have an intelligence quotient of 70 or above, and are generally able to use words to communicate in daily life. In contrast, low-functioning autism (LFA) is an unofficial term referring to individuals on the severe end of the autism spectrum, often with an intelligence quotient of below 70 and very limited language production and comprehension. Since the first report by Kanner in 1943, the prevalence of autism has been increasing. One of the most recent reports from the US Centers for Disease Control and Prevention (CDC) (2014) estimated that ASD occurs in 1 out of 68 children (or 14.7 per 1000 8-year-old children). The prevalence of ASD in the Chinese population is at least 1.18 per 1000 in China (Sun et al., 2013; Zhang & Ji, 2005), with at least 1.3–2 million children under 13 years old affected by ASD nationwide (Huang, Jia, & Wheeler, 2013). A brief report by Tao (1987) on four cases of infantile autism marked the first published study of ASD on the Chinese population. Chinese languages, such as Mandarin and Cantonese, are tonal languages. It is currently unclear whether theories and findings based on non-tonal languages such as English can be applied to the Chinese-speaking individuals with ASD due to a paucity of cross-language studies.

The landscape of research and scholarly work pertaining to ASD on the Chinese population is still relatively sparse, despite an exponential increase in the number of studies on English-speaking individuals with ASD in the last 15 years. Chinese (e.g., Mandarin or Cantonese) phonology differs from non-tonal languages, such as English, in terms of its use of pitch at the phonemic level. Specifically, the pitch pattern (e.g., high fundamental frequency (F0) vs. low-rising F0) serves to differentiate meanings of words that share the same phonetic segments (e.g., Tone 1 *bi* means *force* vs. Tone 2 *bi* means *nose*). Whether individuals with ASD whose native language is tonal show similar speech-processing deficits to those with a non-tonal native language (e.g., English or Finnish) is an important experimental and theoretical question, because, as reviewed in Tsao’s chapter (Chap. 10) and Singh’s chapter (Chap. 11) in this book, children with tonal language backgrounds may follow a somewhat divergent developmental trajectory from that of children with non-tonal language backgrounds (also see a recent review by Curtin & Werker, 2018).

Our goal in this review is to provide an overview and an evaluation of the behavioral and neurological evidence examining Chinese prosody and lexical tone processing in individuals with ASD. Research on speech perception and production abilities in Chinese individuals with ASD is a recently emerging area, but significant progress has been made. Systematically summarizing our understanding of prosody and lexical tone processing in tonal language speakers, as well as how these factors

associate with ASD at both behavioral and brain levels, will further the research and clinical practice in these areas.

We searched the following databases: Cochrane, ERIC, Google Scholar, NCBI/PubMed, PsycINFO, and Web of Science, using the keywords: {"lexical tone," "prosody," "intonation," "pitch," or "fundamental frequency"} and {"Mandarin," "Cantonese," or "Chinese"} and {"Autism" or "Asperger"} in March and August 2017. We checked the bibliographies of the relevant articles and found only nine relevant studies on Chinese prosody and lexical tone perception and production across the field of behavioral and neurophysiological research. Two of the studies focused on Cantonese and seven focused on Mandarin. In this review, we examine these studies in relation to the extensively researched area of prosody, pitch production, and perception in individuals with ASD from English and other non-tonal language backgrounds.

We first discuss findings from the behavioral literature and then describe recent data obtained using brain measures. This review is followed by a case study on the neuroplasticity of children with ASD. Lastly, we discuss the theoretical and clinical significance of the evidence that has accumulated thus far, and we point out gaps and challenges in understanding prosody, lexical tone perception, and production of Chinese-speaking individuals with ASD.

13.2 Prosody in Individuals with ASD: General Background

The terms "prosody" and "intonation" are often used interchangeably in the literature. In this paper, we use "prosody" as a superordinate term to describe changes in pitch, intensity, duration, and voice quality (Cummins et al., 2015; Titze, 1994). In other words, prosody is a suprasegmental feature of speech that is expressed via variations in pitch (fundamental frequency), loudness (intensity), duration, stress, and rhythm (Culter & Isard, 1980). Pitch is the perceptual correlate to the frequency of vocal fold vibration (i.e., F0). During vocal production, individuals often modulate their pitch to convey different emotions and pragmatic connotations (e.g., posing a question, delivering a statement, making an imperative order, expressing surprise). In tonal languages such as Mandarin, Cantonese, and Thai, pitch (which serves as a phonemic contrast) is also referred to as "lexical tone."

Two of the most widely used diagnostic tools for ASD in English-speaking culture, the Autism Diagnostic Observation Schedule-2 (ADOS-2; Lord et al., 2012) and the Autism Diagnostic Interview-Revised (ADI-R; Rutter, Le Couteur, & Lord, 2003), include atypical prosody production as a diagnostic criterion. Atypical prosodic production and perception in individuals with ASD have been reported from the onset of research within this group (Asperger, 1944; Kanner, 1943; Simmons & Baltaxe, 1975). Some researchers proposed that distinctive and atypical vocal characteristics such as monotonous and robot- or machine-like speech may serve as

one of the earliest-appearing biological markers of later ASD diagnosis. However, there is some evidence that children with ASD may have intact prosody perception (Grossman & Tager-Flusberg, 2012; Paul, Augustyn, Klin, & Volkmar, 2005) or superior pitch processing skills (e.g., Bonnel Mottron, Peretz, Trudel, Gallun, Bonnel, 2003; Stanutz, Wapnick, & Burack, 2014). An understanding of prosodic production and perception in ASD is critical since prosody often serves as a major cue in conveying both linguistic and paralinguistic functions (Crystal & Quirk, 1964). Moreover, attention and sensitivity to prosody play a critical role in early language development (Juszyk, 1997; Mehler et al., 1988). Children with ASD are known to have difficulties detecting vocal prosodic cues that convey irony and sarcasm (Wang, Dapretto, Hariri, Sigman, & Brookheimer, 2004; Wang, Lee, Sigman, & Dapretto, 2006, 2007). Children and adolescents with high-functioning autism (HFA) sometimes have difficulty making use of vocal cues to make inferences about a speaker's intentions in tasks designed to probe theory of mind (ToM) (e.g., Chevallier, Noveck, Happé, & Wislon, 2011). ToM is a theoretical position that individuals, such as children with ASD, have difficulty understanding that other people have separate thoughts, intentions, and feelings that are different from one's own (Baron-Cohen, Leslie, & Frith, 1985). In an fMRI study, Eigsti and colleagues found that, compared to typically developing peers, high-functioning children with ASD showed broader recruitment of brain areas while processing both affective and grammatical prosodic cues (Eigsti, Schuh, Mencl, Schultz, & Paul, 2012). The authors suggested that for a fairly simple language-processing task, greater recruitment of areas involved in executive functions and what they refer to as "mind-reading" functions can be interpreted as less automaticity in processing language.

13.3 Atypical Prosody Production in ASD

13.3.1 Atypical Prosody Production in ASD with Non-tonal Language Background

The speech of individuals with ASD has been described as both "monotone" and "exaggerated" (Baltaxe & Simmons, 1985). Atypical prosody production has been suggested as a "bellwether" of the cognitive profiles of individuals with ASD, as well as a behavioral indicator of subtypes of ASD (e.g., hypersensitive vs. hyposensitive to auditory input) (Diehl, Berkovits, & Harrison, 2010, p. 167). However, the precise features of prosody in the speech of individuals with ASD are only recently becoming evident. A review of 16 earlier studies on prosody production in ASD by McCann and Peppé (2003) revealed many contradictory findings; these contradictions may be due to inadequate research, small sample sizes in many studies, and great variability in methodology across studies.

A number of more recent studies with larger sample sizes indicate that higher mean pitch and/or wider pitch range in speech of participants with ASD is the primary

prosodic difference in comparison with the speech of controls when using tasks such as lexical elicitation (e.g., Bonnel et al., 2003), sentence elicitation (Diehl & Paul, 2013) and spontaneous prosodic production (Diehl, Watson, Bennetto, McDonough, & Gunlogson, 2009; but see Quigley et al., 2016 for contrary findings). A recent systematic review and meta-analysis on 34 empirical studies of vocal production in ASD calculated a moderate effect size (Cohen's d of 0.4–0.5), but with a discriminatory accuracy of only about 61–64% (Fusaroli, Bang, Bowler, & Gaigg, 2017). Acoustic measures (e.g., duration and intensity) other than mean and variance of pitch have also been examined, but they were not found to be stable predictors of speech produced by individuals with ASD (Fusaroli et al., 2017).

Endeavors to link pitch production with severity of ASD have led to highly inconclusive results. A computerized task, the Profiling Elements of Prosodic Systems in Children (PEPS-C), developed to assess prosody perception and production in children aged 4–16 years, has been used in a number of studies. The general finding was that children with ASD have atypical prosodic output (e.g., Diehl et al., 2009; McCann, Peppé, Gibbon, O'Hare, & Rutherford, 2007), but that they also demonstrated a considerable amount of similarity in their prosodic output/production relative to their typically developing (TD) peers and children with other types of disorders. This was especially the case with simple tasks, such as when models or prompts were provided and/or no face-to-face spontaneous interaction was required (e.g., Diehl & Paul, 2013).

As Fusaroli et al. (2017) pointed out in their review paper, among the five studies that have examined the correlation between pitch measures and severity of ASD symptoms, both strength and direction of the relation varied across studies and pitch parameters. For example, Bone and colleagues analyzed speech segments of spontaneous interaction between a child and a psychologist during a standard observational evaluation session using the ADOS and found that autism severity score was negatively correlated with the median pitch slope of the turn end, but no correlation with the pitch center or pitch slope variability (Bone et al., 2014). Nadig and Shaw (2012) found no correlation between pitch range and behavioral characteristics (IQ, language, and autism severity scores) of children that were examined based on analyzing spontaneous conversation samples. Nakai and colleagues have also reported no correlation between pitch coefficient of variation and total score from the Autism Screening Questionnaire (Nakai, Takashima, Takiguchi, & Takada, 2014). Diehl et al. (2009) tested two groups of children using similar narrative elicitation tasks. There was a positive correlation between clinician's judgment of ASD severity and the variance of ASD pitch production in the older children and teenagers (Study 1, 10–18 years old), but no such correlation in the younger children (Study 2, 6–14 years old). It is possible that the difference in measurement and methods (e.g., unprompted test condition) among these studies led to different outcomes.

So far, pitch measures show promise to distinguish individuals with and without ASD, but there is no evidence that they reflect severity of the disorder. We are certainly not ready to use prosodic measures as a tool to measure the severity of ASD. Several steps need to be taken first. In particular, it will be important to develop tasks that consistently and robustly result in prosodic differences between individuals with

ASD and controls. Additional research needs to be undertaken with individuals who have language impairment, as well as ASD.

13.3.2 Atypical Prosody Production in Chinese Speakers with ASD

Tonal languages, such as Mandarin and Cantonese, use pitch both as a phonemic contrast at the lexical level and as a suprasegmental cue for intonation changes (as with non-tonal languages). The “contour interaction” theory (Thorsen, 1980; Vaissiere, 1983) and the “tone sequence” theory (Pierrehumbert, 1980) posited that lexical tone, stress, and intonation are closely knitted into the final suprasegmental pitch movement output. Both prosodic pitch patterns and lexical tone are realized via F₀. F₀ modulation for lexical tone, however, is used to distinguish meaning at the level of the morpheme, whereas F₀ modulation for prosody is used for meaning at sentence/discourse level (e.g., pragmatics). F₀ can also be used in lexical stress, which signals the prominence of a syllable in a multisyllabic word or phrase. In tonal languages, lexical stress is superimposed on top of the lexical tone. The sentence-level prosodic patterns (“sentence-level tunes”), including contrastive stress and sentential intonation, can convey modality (i.e., illocutionary force, e.g., assertive vs. interrogative). These are considered “larger waves,” while lexical tone at the syllable level and lexical stress of multisyllabic morphemes are “small ripples.” Chao (1968) described the competition for F₀ space between lexical tone and prosody as “small ripples riding on larger waves” (p. 39). See Yu, Wang, and Li’s chapter (Chap. 5) of this book for the neural mechanisms of lexical tone processing in healthy adults, and Lee and Cheng’s chapter (Chap. 6) of this book for the neural development of lexical tone processing in early childhood. Tsao and Liu (Chap. 10) reviewed lexical tone perception in infancy. The question relevant to this review is how tonal language speakers with ASD convey multiple levels of pitch information, and whether they demonstrate similar pitch output at the syllable, morpheme, word, phrase, and sentential levels compared to that of healthy controls.

To the best of our knowledge, Chan and To (2016) is the only study that has examined the acoustic features of prosodic output in Chinese speakers with ASD. Chan and To (2016) focused on the use of sentence-final particles (SFPs) and the expressive intonation in Cantonese-speaking adults. Cantonese SFPs are bound morphemes that play a similar role to that of prosodic patterns in other languages. In Cantonese, the SFPs convey grammatical, pragmatic, and affective meaning. The authors proposed that individuals with high-functioning ASD (HFA) might show difficulties in mastering the use of SFPs and intonation, due to their known deficits in decoding pragmatic and affective cues. These two skills (the use of SFPs and the use of intonation) might interact with each other and work in a compensatory fashion. Speech samples were generated from 38 young adults (HFA group: $n = 19$; control group: $n = 19$) using spontaneous story retelling. The pitch variance

was measured using sentence as a unit. Higher average pitch and larger pitch variations were found in the HFA group than the control group. The HFA groups and the controls are comparable in terms of the total frequency of SFP use, but HFA groups produced slightly fewer SFP types on average than the control groups ($p = 0.072$). The correlations between pitch and SFP measures were all nonsignificant, with the exception of a moderate positive correlation between the type of SFPs and the pitch variability in the HFA group only. At the individual level, some individuals with HFA showed similar pitch patterns to the healthy control counterparts.

It appears that the general patterns of atypical prosody production in Cantonese speakers with HFA were very similar to those of non-tonal language speakers with HFA, which suggests that prosody impairment may be language-independent.

13.4 Prosody Perception in Children with ASD

13.4.1 *Infant Development*

Prosodic cues are critical in assisting infants with the segmentation of running speech input into linguistically meaningful units (e.g., syllables, words, phrases). Language-specific prosody perception and processing develop during early infancy (Bosch & Sebastián-Gallés, 1997; Friedrich, Herold, & Friederici, 2009; Jusczyk & Aslin, 1995; Sambeth, Ruohio, Alku, Fellman, & Huotilainen, 2008; Shafer, Jaeger, & Shucard, 1999; Stefanics et al., 2009; see Chap. 10 by Tsao & Liu in this book for a review on lexical tone development). For example, a language-specific preference for words with trochaic structure, which is a predominant English stress pattern, was observed in English-learning infants between 6 and 9 months of age (Jusczyk, Cutler, & Redanz, 1993), in German-learning infants between 4 and 6 months of age, but not in French-learning 6-month-olds (Höhle, Bijeljac-Babic, Herold, Weisenborn, & Nazzi, 2009). In addition, neural responses in English-exposed three-month-old infants showed that they process Dutch, but not Italian in a similar fashion to English stories (Shafer et al., 1999). These findings are presumably due to the fact that Germanic languages (e.g., English, German, and Dutch) are stress-timed with trochaic predominance at the syllable level. In contrast, Romance languages (e.g., French and Italian) are syllable-timed and favor iambic stress patterns. The infants in Shafer et al. (1999) also showed evidence that they distinguished the greater pitch range of the Dutch compared to the English stories. This finding indicated early sensitivity to small differences in the melody of speech. Typically developing infants demonstrated an intrinsic preference for the prosodically rich child-directed speech (e.g., Vouloumanos & Werker, 2007). Infants' early preference for higher pitch and exaggerated prosody as in child-directed speech assists early socio-communicative learning (Kuhl, Coffrey-Corina, Padden, & Dawson, 2005), and facilitates and predicts later language development (Thiessen, Hill, & Saffran,

2005). Furthermore, preference for speech correlated positively with general cognitive ability at 12 months, and weaker preference for speech over non-speech sounds correlated with more autistic-like behavior in infants who had siblings with diagnosed ASD (Curtin & Vouloumanos, 2013). Multiple studies have reported that young children with ASD have less robust preference for child-directed speech in comparison with their age-matched TD peers (e.g., Paul, Chawarska, Fowler, Cicchetti, & Volkmar, 2007; see Filipe, Watson, Vicenta, & Frota, 2017 for a review).

Thus, deviant patterns of prosodic perception and processing in the first few years of life could serve as a risk factor of ASD. More prospective studies of children at risk for ASD will need to be carried out to further explore this possibility.

13.4.2 Prosodic Perception in Older Children and Adults with ASD

Mixed evidence has been reported regarding acoustic tone perception and speech (word, phrase, and sentence) prosodic perception abilities in older children and adults with ASD. The majority of studies on acoustic tone perception have reported intact prosody perception in these children with ASD (Grossman & Tager-Flusberg, 2012; Paul et al., 2005). For example, individuals with ASD have superior pitch processing skills based on evidence from a variety of psychophysical measures (e.g., Bonnel et al., 2003, 2010; Stanutz et al., 2014), superior pitch direction detection in small intervals (Heaton, 2003, 2005; Heaton et al., 1999), or superior melodic contour identification (Järvinen-Pasley, Peppé, King-Smith, & Heaton, 2008). Contradictory findings included inferior pure tone discrimination when the reference tone was varied across trials (e.g., Boets, Verhoeven, Wouters, & Steyaert, 2015). See Haesen, Boets, and Wagemans (2011) for an extensive recent review. Many studies on sentence-level intonation revealed that children with ASD have deficits in intonation perception and/or production, especially at the sentence level. For example, Järvinen-Pasley and colleagues reported that children with ASD have unimpaired perception of word-level intonation, but deficits in understanding sentence-level intonation (e.g., Järvinen-Pasley et al., 2008). McCann and colleagues reported that most children with HFA have either expressive or receptive prosody deficits as measured by tasks that assess effect-related prosody at the single word-level, phrase-level stress, and sentence intonation (McCann et al., 2007). Prosodic production deficits have also been reported in other studies in children with HFA (e.g., Diehl & Paul, 2013); furthermore, there is also evidence that adults with HFA have difficulties using prosodic cues to extract information about mood and emotion (e.g., Rutherford, Baron-Cohen, & Wheelwright, 2002). A brain imaging study by Wang et al. (2006) showed that unlike TD controls, school-aged children with ASD were not only less accurate at interpreting prosodic cues for irony, but also showed aberrant brain activation patterns including absent activity in the medial prefrontal cortex, greater activation in the right inferior frontal gyrus, and bilaterally superior temporal sulcus (STS) regions.

The evidence so far suggests that the function of the pitch information (linguistic and non-linguistic) and the size of the prosodic unit (e.g., word level vs. sentence level) influence the performance level in these populations.

13.4.3 Prosody and Lexical Tone Perception in Chinese-Speaking Individuals with ASD: Behavioral Research

Crosslinguistic studies have suggested that extensive experience with a native tonal language attunes perception of pitch contour in language processing (Burnham & Francis, 1997; Gandour, 1983; Gandour & Harshman, 1978; Stevens, Keller, & Tyler, 2013; Wayland & Guion, 2004; Xu, Krishnan, & Gandour, 2006b) and music processing (Alexander, Bradlow, Ashley, & Wong, 2011; Bidelman, Hutka, & Moreno, 2013; Stevens et al., 2013). See Chap. 8 of this book for a review by Ong and colleagues). For example, Mandarin listeners can better perceive the subtle acoustic differences of Mandarin tonal categories compared to non-Mandarin speakers (Hallé, Chang, & Best, 2004; Leather, 1983; Lee, Vokoch, & Wurm, 1996). Cantonese speakers can discriminate Mandarin lexical tones better than English speakers (Lee et al., 1996). Native tonal (i.e., Thai) language speakers were faster and more accurate at discriminating pitch contour in both natural speech and musical contour tasks (Stevens et al., 2013). Mandarin listeners outperformed English listeners when discriminating Thai lexical tones post-training under both short and long interstimulus interval conditions (Wayland & Guion, 2004). Furthermore, using a multidimensional scaling method, Gandour and colleagues found that there was a perceptual dimension weighting difference between Chinese listeners and native English listeners when processing Mandarin tones. Specifically, native English listeners tended to rely on the pitch height while Chinese (Cantonese and Mandarin) listeners focused on both pitch height and pitch direction when processing Mandarin lexical tones (Gandour, 1984; Gandour & Harshman, 1978).

Whether Chinese individuals with ASD demonstrate the same superior attunement to lexical tone and musical prosody in relation to non-tonal language speakers with and without ASD is an open question due to lack of cross-language research. There are only five behavioral speech perception studies on Chinese children with ASD. One study focused on lexical tone, one on intonation cues, and one on musical perception, while the remaining two studies focused on sentential prosody. These studies revealed differences between children with ASD and children with typical development in processing linguistic and non-linguistic stimuli. The findings were generally consistent with studies of prosodic perception in children of non-tonal languages. Below are the highlights of each study.

Chen et al. (2016) examined lexical tone (Tone 1 and Tone 2) identification and discrimination of the Mandarin syllable /i/ in eleven 6- to 8-year-old boys with ASD.

The stimuli were from an 11-step lexical tone continuum ranging between a high-level tone (Tone 1) and a low-rising tone (Tone 2). These children had an average language level of 3 years and 6 months. Compared to age-matched typically developing controls, children with ASD exhibited much lower discrimination accuracy and a broader category identification boundary. Furthermore, a strong negative correlation between the boundary width and the developmental language age was found in children with ASD.

Li, Law, Lam, and To (2013) examined how Cantonese-speaking children with ASD implemented sentential prosodic cues and sentence-final particles (SFPs) in ironic stories to judge speaker's belief and intent. Ironic expression is usually achieved via slow speaking rate, larger pitch variation, and greater intensity (Cutler & Bruck, 1974). As discussed above, some believe that SFPs and intonation have a trading relation with each other in presenting sentential connotation (Kwok, 1984; Yau, 1980). Li et al. (2013) tested children with and without ASD between the ages of 8.3 and 12.9 years by using 16 ironic stories and 5 complementary stories. The two groups demonstrated similar levels of comprehension when sentences did not contain prosodic or SFP cues. Participants in both groups answered the questions about the factual content of the stories with similar accuracy. A large group difference was observed for sentences with prosodic cues only, sentences with SFP cues only, and sentences with both prosodic and SFP cues. These Cantonese-speaking children with ASD failed to exploit either prosodic cues or SFP cues, similarly to English-speaking children with ASD (Happé, 1993). Note that the Cantonese-speaking TD children answered the questions with only slightly above chance accuracy under the prosody-only condition, suggesting that the prosody-only condition is challenging even for TD children.

Jiang, Liu, Wan, and Jiang (2015) investigated discrimination and identification of music and linguistic pitch contours in 17 Mandarin-speaking individuals (age: 6.0–16.2) with high-functioning autism and 17 control children and adolescents with matching age, nonverbal IQ, and years of music training. They used a five-tone sequence for the melodic contour discrimination and identification, and disyllabic verb-object constructions as stimuli for the speech intonation discrimination and identification tasks. Participants were asked to match the auditory sequence with the visual display of the melodic contour of the music or to identify whether the disyllabic verb-object construction was a statement or question. They found that the ASD group performed worse than the control group in terms of Mandarin intonation discrimination and identification, but the ASD group performed better than the control in the melodic contour identification task. The two groups showed similar performance in the melodic contour discrimination task. Jiang and colleagues suggested that linguistic pitch may not be processed the same way as musical pitch in Chinese-speaking individuals with ASD.

The Mandarin question words such as “what” (什么) and “who” (谁) can convey both statements and questions depending on their prosodic features (rising vs. falling sentence-final intonation) and/or their semantic structures (e.g., by adding the Mandarin universal quantifier “dou”/都, which means “all” in English). Su, Jin, Wan, Zhang, and Su (2014) tested 28 children (14 four- to eight-year-olds; 14 nine-

to fourteen-year-olds) with high-functioning ASD and 28 age-matched TD controls using a computerized sentence comprehension task to identify statements and question sentences with either prosodic cues or semantic cues. They found that older children with ASD performed on par with their TD peers, by using either prosodic cues in ambiguous sentences or semantic cues in unambiguous sentences. However, younger children with ASD performed more poorly than the TD controls in terms of statement sentence processing under all structure conditions (prosody, semantic, and control structures). This highlighted a developmental delay in children with ASD in comprehending statement sentences containing wh-words. The findings in this study, together with several other studies on English-speaking individuals with ASD, support the claim that grammatical prosody is relatively spared in children with ASD compared to the affective or pragmatic prosody.

Wang and Tsao (2015) did not aim to explore the tone language-specific processing in Mandarin-speaking ASD children; rather, they aimed to examine “emotional prosody” perception in this group of children. Given that the data were collected from tonal language-learning children, it can potentially include an interaction between tonal language experience and emotional prosody processing. Moreover, this is the only study so far that has examined emotional prosody processing in tonal language-learning children with ASD. For both reasons, we included the study in this review. Wang and Tsao (2015) used three emotion tones (happy, sad, and angry) and a neutral tone presented in words and short sentences. Twenty-five boys with high-functioning autism (HFA) and 25 TD boys between 6 and 11 years of age were tested using an emotional prosody identification task. The study found that children with HFA performed more poorly than TD children in identifying prosodic patterns associated with “happy,” but that they did not differ in identifying prosodic patterns associated with “sad” or “angry.” This was true regardless of whether the semantic condition of the stimulus (words or sentences) was neutral or emotionally relevant. Correlation analyses revealed a strong positive association between perception accuracy of happy prosody and the pragmatic language skills and social adaptation skills of children with ASD.

13.4.4 Neural Indices of Lexical Tone Processing in Non-tonal ASD

Event-related potential (ERP), recorded using electroencephalogram (EEG), and event-related field (ERF), recorded via magnetoencephalography (MEG), are most often used to examine auditory processing of pure tone and speech. Many such ERP/ERF studies have adopted an oddball paradigm in which repetition of one sound pattern is interspersed with an infrequent sound pattern. The long-latency ERP/ERF obligatory sequence of peaks, P1-N1-P2, is thought to reflect the brain’s response to the physical features of the stimuli at the scalp level with multiple neural generators in the primary and secondary auditory cortex (Näätänen & Picton, 1987; Ponton,

Eggermont, Khosla, Kwong, & Don, 2002; Scherg & Von Cramon, 1986). P1 to auditory events is observed at frontocentral sites in early childhood. P1 latency shifts earlier as the brain matures (Čeponienė, Rinne, & Näätänen 2002; Choudhury & Benasich, 2011; Kushnerenko, Čeponienė, Balan, Fellman, & Näätänen, 2002; Morr, Shafer, Kreuzer, & Kurtzberg, 2002; Shafer, Yu, & Datta, 2010; see Sharma, Glick, Deeves, & Duncan, 2015 for a review). N1 and P2 are not always apparent in young children to auditory stimuli presented at rates less than about 1 per second (Ponton, Eggermont, Kwong, & Don, 2000). The P1-N1-P2 complex does not reach full maturity until later adolescent years (Ponton et al., 2000). Bishop, Hardiman, Uwer, and von Suchodoletz (2007) reported that P1 amplitude appears to be larger and that it also peaks earlier for speech than for pure tones in children under 11 years of age.

Mismatch negativity (MMN) serves as an index of automatic preattentive cortical discrimination of auditory contrast. MMN is largest at frontal sites and is best seen by subtracting the response to a frequent stimulus/pattern from the response to an infrequent stimulus pattern. MMN is larger for greater physical (acoustic) differences between two stimuli and is often larger for speech sounds that cross a phoneme boundary than for speech sounds that fall within the same phoneme category (Näätänen, Paavilainen, Rinne, & Alho, 2007). Significant maturational changes have also been evidenced in the presence, amplitude, and latency of MMN in TD children with non-tonal language backgrounds (Friederici, Friedrich, & Weber, 2002; Kushnerenko et al., 2007; Leppänen et al., 2002; Morr et al., 2002; Shafer et al., 2010; Shafer, Yu, & Datta, 2011) and tonal language backgrounds (Cheng et al., 2013, 2015; Liu, Chen, & Tsao, 2014). In particular, infants and young children often show a positive mismatch response (pMMR), rather than MMN or in addition to the MMN (Shafer et al., 2010). The pMMR may reflect greater recovery from refractoriness (because the deviant stimulus is less frequent).

The P3a response is an index of involuntary attention switch elicited by a salient stimulus change or a rare stimulus change. Its latency is later than the MMN and shifts earlier progressively starting from early toddlerhood, reaching stabilization at around 12 years of age (Fuchigami et al., 1995). However, its scalp distribution matches with those of the adults only until late adolescence (Määttä et al., 2005).

The first few ERP studies to examine auditory processing in children with ASD suggested superior processing of non-speech stimuli compared to typically developing controls. Oades, Walker, Geffen, and Stern (1988) tested seven children with ASD and found shorter N1 latency and larger N1 amplitude to a pure tone contrast in the ASD group than the TD controls. Ferri et al. (2003) also found that the children with low-functioning autism (LFA) showed earlier and enhanced N1 peaks to pure tone contrasts (1000 Hz vs. 1300 Hz). Enhanced MMN and/or P3a have also been evident in children with ASD for auditory tones (Ferri et al., 2003; Kujala et al., 2007). Gomot et al. (2008) used a combination of behavioral discrimination and fMRI measures and found that children with Asperger's syndrome were hyperactive to sound, as indicated by faster discrimination reaction times and similar response accuracy. They also showed stronger activation in the prefrontal and inferior parietal cortices to complex tone discrimination.

Čeponienė et al. (2003) found that pitch encoding and discrimination were similar for children with ASD and a TD group as measured using P1-N2-P2-N4, MMN, and P3a. They did observe a marginally smaller P1 amplitude in the ASD group. In contrast, other studies have found delayed and diminished cortical responses to auditory and speech stimuli in children with ASD. School-aged children with ASD demonstrated delayed N1 latency to pure tones, compared to age-matched healthy controls (longer N1c latencies; Bruneau, Bonnet-Brilhault, Gomot, Adrien, & Barthélémy, 2003; Bruneau, Roux, Adrien, & Barthélémy, 1999). Another study revealed diminished and delayed P1 responses to pure tone contrasts (Jansson-Verkasalo et al., 2003). Diminished P1, N2, P3, and N4 have also been reported for vowel contrasts in children with ASD (Whitehouse & Bishop, 2008). Quite a number of studies have reported that individuals with ASD have delayed and/or diminished MMN and P3a responses to phonemic change (Lepistö et al., 2005, 2006). We have also observed a delayed latency of the P1 peak to auditory words in a picture–word priming paradigm in minimally verbal 3–7-year-old children with ASD compared to age-matched controls (Cantiani et al., 2016). The controversial findings among studies may be due to experimental factors such as stimuli and tasks used, and the heterogeneous nature of the ASD population in general.

13.4.5 Neurophysiological Measures of Pitch Processing in Chinese-Speaking Individuals with ASD

The pursuit of understanding how the autistic brain processes native lexical tone has only just begun. Currently, only three careful studies from the same research group have been undertaken (Huang et al., 2017; Wang, Wang, Fan, Huang, & Zhang, 2017; Yu et al., 2015). These studies suggest that children with ASD have greater difficulty in processing speech than non-speech information, but that certain types of cues (e.g., duration) may be spared. Below are the highlights of the three studies.

Yu et al. (2015) were the first to examine the question of whether enhanced lower-level perceptual features such as pitch variation would hinder the processing of higher-level phonemic units of the lexical tone categories. In their first oddball experiment, three types of stimuli were used: simple pure tone contrast (standard 216 Hz vs. deviant 299 Hz), lexical tone contrast (standard /bai2/ vs. deviant /bai4/), and a nonword condition (standard /rai2/ vs. deviant /rai4/). The ASD group had larger MMN amplitudes than the TD control group at the vertex site (Cz) and had smaller MMNs than the TD group at the frontal site (Fz) for the lexical tone contrast /ba2/-/ba4/. MMN was present in the TD group for all contrast types, but for the ASD group, MMN was absent to the nonword /rai2/-/rai4/ contrast at both Fz and Cz sites. In contrast to their TD peers, enhanced P3a amplitude was observed for the pure tone contrast in the ASD group. In order to further understand the influence of lexicality, children were also tested using hummed speech. Results demonstrated that the ASD group had larger MMN amplitude than the controls at Cz but not at

Fz. The ASD group also showed larger P3a amplitude at Cz; the TD group showed a tendency toward shorter latencies for the P3a peak at Fz compared to the ASD group. The authors speculated that the reduced neural sensitivity in lexical tone processing was probably due to inadequate suppression of the irrelevant within-category pitch differences.

The account of speech-specific deficits in autism for lexical tone processing proposed in Yu et al. (2015) was further supported by the findings in Wang et al. (2017). Wang et al. (2017) used vigorously controlled synthetic speech and non-speech contrasts. The speech condition consisted of three acoustically equidistant stimuli from a nine-point continuum /ba2/ (step 1)-/ba4/ (step 5)-/ba4/ (step 9) for between-category (steps 1 and 5) and within-category (steps 5 and 9) contrasts. The non-speech condition consisted of three complex stimuli that matched with speech stimuli on all acoustic parameters, except for harmonic composition. The study used a passive listening oddball paradigm and found that the TD controls showed larger MMN to between-category than within-category, whereas the ASD group had equal MMN for the between- and within-category comparisons; this finding indicated a lack of categorical perception in children with ASD. No significant P3a was observed under any condition for either group. Results from time–frequency analysis provided further evidence for group differences. The two groups demonstrated similar phase locking to harmonic speech stimuli, but for the lexical tone condition only the TD group showed a significant inter-trial phase coherence (ITPC) difference in the theta band for the MMNs of within- versus cross-category contrasts. This evidence further suggested that children with ASD do not have categorical perception in the lexical tone condition.

Duration of speech segments can also serve to distinguish meaning in some languages (e.g., Finnish and Japanese). Behavioral and neurological evidence indicates that individuals with ASD have deficits processing small durational differences in auditory and speech contrasts (Brodeur, Gordon Green, Flores, & Burack, 2014; Falter, Noreika, Wearden, & Bailey, 2012; Lambrechts, Falter-Wagner, & van Wassenhove, 2017; Lepistö et al., 2005; Maister & Plaisted-Grant, 2011; Szelag, Kowalska, Galkowski, & Pöppel, 2004). To answer the question of whether individuals with tonal language backgrounds also show deficits in temporal processing, Huang et al. (2017) used both pure tones (295 Hz) and nonsense syllables (/tý/) of two durations (250 ms vs. 350 ms) in a passive oddball paradigm. They compared the neurophysiological responses to these duration changes in school-aged children with ASD and TD peers. A delayed and diminished MMN peak was found in the ASD group in comparison with the TD control group for the pure tone stimulus condition only. In contrast, a delayed and diminished P3a peak was evidenced in the ASD group for the speech condition only. It is not entirely clear how to interpret this finding, considering that the results from the pure tone differed from that of the vowel for within-category comparisons. Clearly, additional research is necessary to fully understand how native language experience modulates auditory processing in children with ASD.

13.5 Treatment Study of Children with ASD Using Transcranial Direct Current Stimulation: A Feasibility, Pilot Study

Transcranial direct current stimulation (tDCS) is a noninvasive technique of applying constant low-intensity electrical currents to the scalp. This method has been extensively used in animal studies. It is well established in animal models that this type of stimulation can alter the threshold, rate, and balance of excitation and inhibition of neurons, and therefore can modulate brain functions both *in vivo* and *in vitro* (Bikson et al., 2004; Bindman, Lippold, & Redfearn, 1964; Chan, Hounsgaard, & Nicholson, 1988; Purpura & Mcmurtry, 1965; Rahman, Toshev, & Bikson, 2014; see Reato, Rahman, Bikson, & Parra, 2013 for a review). In the past decade, numerous studies have reported that noninvasive brain stimulation (NIBS) techniques such as transcranial magnetic stimulation (TMS) and transcranial direct current stimulation (tDCS) can facilitate language recovery in patients with aphasia (see Norise & Hamilton, 2016 for a review).

Whether language-related plasticity in the brains of children with ASD can be modulated by noninvasive brain stimulation is an emerging area of research. The available literature on the use of tDCS in ASD is preliminary, consisting of studies with methodological limitations (see Jacobson, Koslowsky, & Lavidor, 2012 for a review), but some of the results are promising. A treatment of 20 NIBS sessions was found to improve social and behavioral scales in children with ASD with a lasting effect of six months (Gómez et al., 2017), and a single session of anodal tDCS was found to increase peak alpha frequency at the stimulation site and to decrease autistic-like behavioral symptoms (1 mA, 20 min; Amatachaya et al., 2015), and to increase syntax comprehension in children with ASD (2 mA, 30 min; Schneider & Hopp, 2011). In a randomized controlled trial, after a single session of tDCS treatment (1.5 mA, 15 min) children with attention deficit and hyperactivity disorder showed increased inhibition accuracy (Soltaninejad, Nejati, & Ekhtiari, 2015).

Mandarin-learning children with ASD have shown atypical cortical oscillations as reviewed above (Wang et al., 2017). As per Wang et al. (2017), Mandarin-learning children with ASD may have deficits in inhibiting neural sensitivity to within-category lexical tone variation. We are interested in whether the sensory and cognitive functions associated with aberrant excitatory and inhibitory neuronal activities can be modulated by tDCS stimulation. If we accept the hypothesis that brain stimulation via tDCS technique can enhance the inhibitory neural activity associated with within-category speech processing, then this treatment could possibly enhance the categorization processing of speech that varies at the within-category level. The following pilot study was designed to test this hypothesis.

13.5.1 *Materials and Methods*

Participants Data from two children with ASD (male, 8.1 years old and 9.8 years old) and two typically developing control children (male, 9.11 years old and 10.7 years old) were obtained. A 10.6-year-old nonverbal child with ASD was recruited but could not be tested due to lack of adequate compliance and was therefore excluded from the study. All four children are from the same neighborhood in the metropolitan New York City area, and all were from families in which both parents were native Mandarin speakers. Language background questionnaires from the parents indicated that all four children have had consistent Mandarin exposure from both parents since birth. According to parental reports and background questionnaires, all four children understood daily conversations in Mandarin, and all could carry out simple conversations about daily routines in Mandarin. All four children also had consistent English exposure via school since English was the language of instruction for all four children. According to parental reports, the two children with ASD were more dominant in Mandarin, while the two typically developing children were more dominant in English at the time of testing. Both children with ASD were diagnosed by certified developmental psychologists around 3 years of age, and both children with ASD had individual educational plans (IEPs) and were receiving special education via the public school system due to their ASD diagnoses. Their diagnoses were also validated by the school's special education teacher, as well as by an experienced speech language pathologist (the first author). Informed consent from each parent and verbal assent from each child were obtained following the approved protocol by the local institutional review board.

The Event-Related Potential Procedures

Stimuli Disyllabic nonword stimuli with Tone 2 and Tone 3 contrast were used in an oddball paradigm. The frequent/standard stimuli were three tokens of /gu3pa1/, and the infrequent/deviant stimuli were two tokens of /gu2pa1/. The use of multiple tokens of the same lexical category was to facilitate between-category processing rather than within-category processing. Specifically, the tokens varied in non-relevant acoustic information and only the relevant tone difference could be used to correctly categorize the stimuli. These stimuli were used in Yu, Shafer, and Sussman (2017). In this study, native Mandarin adult speakers showed larger MMN responses compared to English speakers to these lexical tone differences. A total of 165 deviant (20%) and 645 standard (80%) stimuli were presented in 15 blocks with an average interstimulus interval of 675 ms (645–709 ms). Each block was separated by a 10-s break.

ERP recording The electroencephalogram (EEG) was time-locked to the onset of stimuli and recorded using 65-channel sensor nets at the sampling rate of 500 Hz with a band-pass filter of 0.1–100 Hz. Two sessions of ERP recordings were collected, one occurring before the tDCS procedure and another shortly after the tDCS procedure during the same laboratory visit. The data were filtered with a band-pass filter of 0.3–15 Hz and segmented 200 ms before the onset of stimuli and 700 ms post-stimulus onset. Artifact rejection, baseline correction, and average re-reference were performed in BESA 6.1. All children had at least 100 trials from the deviant condition.

The amplitudes of the subtraction waves (deviant minus standard) were compared across the four participants.

High-Definition-tDCS Procedure

We applied high-definition (HD)-tDCS stimulation using the Soterix 1×1 tDCS Low-Intensity Stimulator with a Soterix 4×1 adaptor (Soterix Medical, New York, NY). We placed the stimulating ring electrodes around the frontocentral scalp region (C3, C4, F3, F4 as cathodes and FCz as anode). The sintered Ag/AgCl ring electrodes were fixated with an EEG cap with HD-tDCS electrode holders (Soterix Medical, New York, NY). The impedances of all five electrodes were in the range of 0.50–0.80 quality value. The anode (FCz) was set to deliver a total current of 1 mA, and the return electrodes shared the same current intensity of 1 mA (0.25 mA each). The duration of stimulation was set to 10 min. We used the intensity of 1 mA because HD-tDCS is known to deliver more focalized stimulation (Edwards, Cortes, Datta, Minhas, Wassermann, Bikson 2013), as well as to reduce an unnecessary tingling sensation that many children with ASD might not tolerate. Furthermore, an intensity of 1 mA single session of 20-min stimulation using conventional tDCS was found to elicit significant behavioral improvement in children with ASD (Amatachaya et al., 2015).

13.5.2 Preliminary Results

P1-N1-P2 Results

Figure 13.1 shows the amplitude of the standard condition for all four participants before and after tDCS stimulation. The two children with TD showed a P1 (80–100 ms), followed by N1 (around 150 ms) and P2 (220 ms). The N1 was attenuated compared to adults, but it was clearly emerging for these 10-year-old children (note that N1 is attenuated at short interstimulus intervals for children under 10 years of age). The two ASD children showed only a broad P1 peak followed by an N2 around 250 ms. This pattern was observed in younger typically developing children. There was a negative shift in the response at Fz after about 100 ms for all participants after tDCS. The response at the left mastoid (LM) showed the inverse pattern. The children with ASD showed a somewhat different pattern. Specifically, ASD01 showed a similar pattern to the children with TD from about 200 to 350 ms. ASD02 appeared to show an increased response (i.e., greater Fz negativity and greater LM positivity) from 200 to 600 ms.

MMN Results

Figure 13.2 demonstrates the topography of the subtraction (deviant minus standard) waves. Before tDCS stimulation, in the two TD participants, there was a negative response (blue) between 150 and 250 ms at the frontocentral scalp region. Within the same time window, a predominantly positive response (red) was evidenced in both ASD participants. A robust negativity at the superior central scalp region was

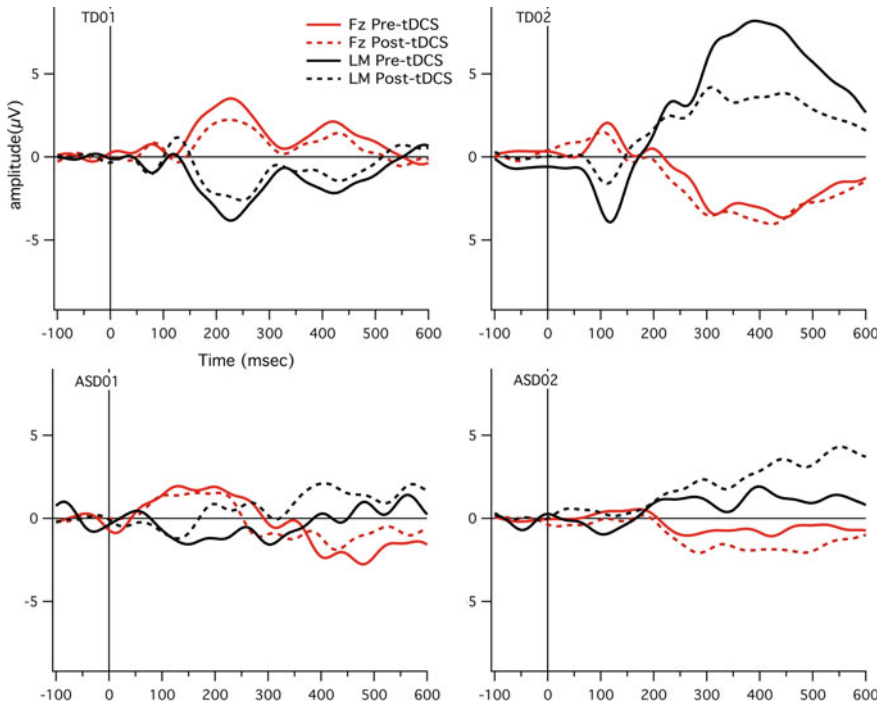


Fig. 13.1 Grand average ERPs to the standard stimulus waveforms for the Fz and left mastoid. The top panel shows the two control participants, and the bottom panel shows the two ASD participants

seen in both children with ASD between 300 and 400 ms. Post-tDCS, the peak of the negativity shifted earlier in both TD children and in one of the children with ASD (Participant D in Fig. 13.2). Due to the small sample size, we can only speculate at this time what these differences indicate: The first, superior negative peak was possibly the MMN and was observed in both the children with ASD and the TD children; however, it appeared in the much later time window before and after HD-tDCS for children with ASD than found for children with TD. The latency of the negativity in the two ASD children was very similar to the four-to-seven-year-old TD children in Shafer et al. (2010). Younger children (infant to seven years of age) often show a positive mismatch response sometimes alone (infants) and sometimes preceding the MMN (Morr et al., 2002; Shafer, Yu, & Garrido-Nag, 2012). In addition, the timing of the negativity in the two children with ASD may be the late negativity (LN) observed in children with specific language impairment in Shafer, Morr, Datta, Kurtzberg, Schwartz (2005). This pattern might indicate a developmental delay in processing complex speech sounds. The ERP amplitude changes post-HD-tDCS suggest that both individuals with ASD and TD respond to the HD-tDCS treatment. In addition, these preliminary data reveal that this is a promising approach for examining the effect of HD-tDCS treatment on lexical tone processing.

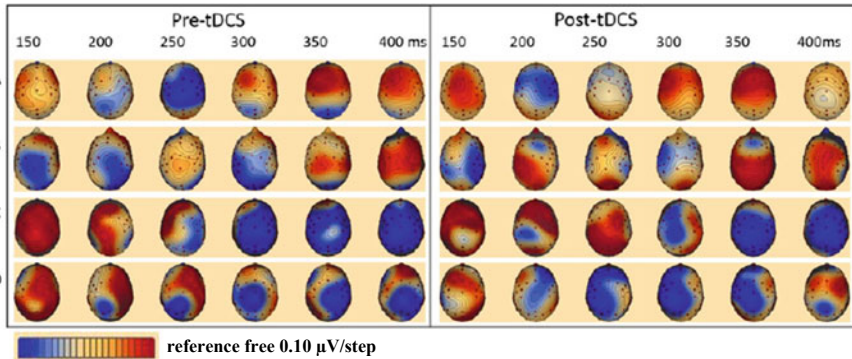


Fig. 13.2 Topographical voltage maps of the subtraction wave (deviant standard) between 150 and 400 ms after the stimulus onset. Red portion shows positivity, and blue shows negativity. A and B are typical control children, and C and D are participants with ASD

13.6 General Discussion and Future Directions

The literature concerning auditory processing in Chinese/tonal language-speaking individuals with ASD is sparse with only a few studies examining perception or production and with very few studies attempting to address both non-linguistic basic auditory pitch processing and higher-level linguistic pitch and prosody processing. Further research needs first to replicate findings of these few studies and then to allow synthesis of the evidence. Even so, some preliminary comparisons of the findings can be accomplished between the general literature of auditory processing in children with ASD and these studies of Chinese-speaking children with ASD.

13.6.1 *Developmental Prosody and Lexical Tone Processing in Children with ASD*

Phonetic features that serve to distinguish lexical tone categories can be differentiated by multiple parameters along several spectral and temporal dimensions (e.g., onset F0, offset F0, contour of F0, and duration of the contour). TD children make few production errors for Mandarin lexical tones as early as 1.6 years of age in picture naming tasks (Hua & Dodd, 2000). However, based on results from perceptual judgment tasks, 6-year-old Mandarin-speaking children could not produce adultlike lexical tones. They also did not reach adultlike perception levels (Wong, 2013; Wong, Schwartz, & Jenkins, 2005). MMN responses to lexical tone change were not adultlike in school-aged children (Liu et al., 2014). Developmental changes in lexical tone perception in Chinese-learning children with ASD have only been addressed in a few studies. Certain prosodic cues (e.g., prosodic cues for irony) and lexical tones (e.g., difference between Tone 2 and Tone 3 in Mandarin) are intrinsically more

challenging to learn, even for TD children. As shown in Li et al. (2013), TD children between 8 and 12 years of age do not demonstrate high comprehension accuracy for prosodic cues when encoding irony. In Su et al. (2014), the older children with HFA performed at or near ceiling, equally as well as the TD controls for all types of sentence processing, and performance gaps were only evident in the younger children with and without ASD. Such findings highlight a developmental delay instead of a persistent deficit in children with ASD when processing statement sentences that contain wh-words (e.g., “what” or “shenme” in Mandarin).

Future studies are needed to elaborate on the developmental trajectories of various prosodic cues and lexical categories in tonal language-learning children, both with and without ASD. This step is necessary before determining how to use this information to enhance clinical practice.

13.6.2 Lexical Tone and Music Processing in ASD

Mounting evidence suggests that brain plasticity governing language processing can be adapted for musical processing (Koelsch et al., 2002; Maess, Koelsch, Gunter, & Friederici, 2001; Patel, Gibson, Ratner, Besson, & Holcomb, 1998). We would like to suggest that future studies explore this relationship in children from tonal language backgrounds. Frameworks such as the shared syntactic integration resource hypothesis (SSIRH) proposed that the shared neural resources between music and language processing are located in the frontal brain regions, and that these resources are recruited “when structural integration of incoming elements in a sequence is costly.” That is, these shared networks are the “processing regions” for structural integration in linguistic syntax and tonal harmony (Patel, 2003, 2014).

For healthy individuals, language experience fine-tunes the production and processing abilities of critical auditory elements in both speech and non-speech domains. For example, Mandarin speakers are better at imitating and discriminating musical pitch than English speakers (Pfordresher & Brown, 2009), and they are more sensitive to both lexical pitch and non-speech (harmonic) pitch category boundaries (Xu, Gandour, & Francis, 2006a). The mutual enhancement between language and music, as seen in musicians and tonal language speakers, has been widely reported (Bidelman et al., 2013). Theoretically, such mutual enhancement should benefit music and speech processing in individuals with ASD and in other communication disorders as well.

Accumulating evidence suggests that individuals with ASD who come from tonal language backgrounds share the common characteristics of superior melodic contour processing and inferior linguistic prosody processing, as observed in individuals with ASD from non-tonal language backgrounds (Järvinen-Pasley & Heaton, 2007; Jiang et al., 2015; Heaton, 2005). The disparity between music and speech-processing skills leads researchers to believe that linguistic pitch and musical pitch are processed differently in individuals with ASD, and that tonal language experience does not compensate for the linguistic prosody processing deficit in individuals with ASD

(Jiang et al., 2015). Currently, there is no direct cross-language comparison. Further cross-language studies are needed to directly test whether individuals with ASD from a tonal language background have an advantage over their counterparts from a non-tonal language background on musical and linguistic prosody processing.

Positive transfer from music training to speech processing has been evidenced in TD children with non-tonal language backgrounds (Moreno et al., 2009). The OPERA model hypothesized by Patel (2011) posited that the shared acoustic features in speech and music, such as F0 changes over time, are processed in an anatomically overlapping network. Patel further argued that music perception “place[s] higher demands on the encoding of certain acoustic features than does speech perception” for adequate communication. However, there is no evidence to support whether such a claim can be applied to all types of languages, especially tonal languages, in which high precision of pitch perception and production is a constant demand. Further studies should examine whether there is a positive transfer from lexical tone learning to music processing, or from music training to lexical tone learning in individuals with ASD. Answers to such questions will provide evidence for testing theory and for designing a therapeutic framework for treating tonal language speakers with ASD.

13.6.3 Effects of Experimental Variables on Chinese Individuals with ASD

So far, all nine studies on Chinese-speaking individuals with ASD have recruited individuals with HFA and no studies have yet focused on minimally verbal individuals with ASD, which counts for about one-third of the ASD population. Our study (which is one of the few existing studies) suggests that some English-exposed children with ASD and with minimal verbal skills have intact (although slightly delayed) lower-level visual and auditory processing skills (Cantiani et al., 2016). Unfortunately, the tasks employed in studies examining prosody and lexical tone production and processing often require functional communication skills (e.g., answering questions) as well as competency in social interaction (e.g., initiating conversation). Such demands preclude the participation of children with poor verbal skills who have LFA. Neurophysiological methods, including EEG-ERP methods, can use tasks that do not require a response (e.g., the passive oddball paradigm used by Zhang’s research group reviewed above provides a good alternative to test individuals with LFA). However, individuals with LFA often have sensory sensitivities and some cannot tolerate wearing the sensor net. On the other hand, it is possible to desensitize some of these children so that they will be able to tolerate electrodes (see Roesler et al., 2013). Multiple studies have conducted correlational analyses to seek the relationship between prosody characteristics and severity of autistic symptoms; some did not find a correlation between prosody and the severity of ASD, but given that few studies have included children with LFA, it remains unclear whether a relationship exists. This shortcoming is particularly important given that the definition of autism

does not differentiate HFA from LHA under the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5).

Recently, Eigsti and Fein (2013) compared pitch discrimination sensitivity of teenagers with optimal outcomes and teenagers with HFA and age-matched TD controls, and found that superior pitch perceptual skill is correlated with ASD symptomatology. Specifically, teenagers who were diagnosed with ASD before 5 years of age and who at the time of testing did not have any autistic symptoms performed the same as the controls. In contrast, teenagers who still maintained the ASD diagnosis continued to demonstrate heightened pitch discrimination. The heterogeneous nature of this disorder calls for further studies that examine the features of subgroups on the spectrum.

As the world is becoming more plural, the proportion of bilingual individuals with ASD will also increase. Language development in bilingual children who are exposed to two languages from birth follows a different developmental trajectory starting from the first year of life (Bosch & Sebastián-Gallés, 2003; Shafer et al., 2012). There is evidence of both positive and negative transfers between the first and second language phonologies (Hambly, Wren, McLeod, & Roulstone, 2013). Wider pitch range and higher average pitch have been found in English-speaking, German-speaking, Mandarin-speaking monolingual individuals with ASD and Hindi–English bilinguals with ASD, but the opposite patterns were evidenced in Japanese-speaking children with ASD (Baltaxe & Simmons, 1985; Chan & To, 2016; Green & Tobin, 2009; Sharda et al., 2010; Nakai et al., 2014). It will be of theoretical and clinical interest to compare the pitch perception and lexical tone production of bilingual children with those of monolingual Chinese-speaking individuals such as in Chan and To (2016). The different language backgrounds of bilingual children with ASD (e.g., two tonal languages vs. one tonal language plus one non-tonal language; more balanced bilingual vs. one language dominant bilingual) would presumably lead to different hypotheses regarding the association between pitch perception, music perception, and lexical tone production.

Stimulus complexity and task demand are both known to influence auditory processing. Differences between a pair of acoustically less-salient lexical pairs such as Tone 2 versus Tone 3 can be more challenging than an acoustically more salient pair such as Tone 1 versus Tone 3 for non-native listeners (Chandrasekaran, Krishnan, & Gandour, 2007; Yu et al., 2017). Tone sandhi (e.g., when there are two 3rd tones in a row, the first one becomes 2nd tone) also further complicates the production and perception of Tone 2 versus Tone 3 processing (see Chap. 7 of this book by Chang & Kuo). Speech versus non-speech comparisons are routinely used to measure domain-general versus language-specific auditory processing. Lexical status of the stimuli (e.g., word vs. nonword) seems to play a subtle role in the nature of the cortical response to such contrasts in individuals with ASD (Wang et al., 2017; Yu et al., 2015). Stimulus complexity often interacts with attention and/or memory demands in children with ASD and other learning disorders (Čeponienė et al., 2003; Whitehouse & Bishop, 2008). Systematic investigation using carefully controlled tasks and varying stimulus complexity will provide important implications for theories and clinical application.

13.6.4 Theoretical Implications

Higher perception accuracy, faster response time, and larger brain response to relatively simple, domain-general perceptual contrasts in ASD than in controls are taken as evidence for enhanced lower-level perceptual processing (Mottron, Dawson, Soulieres, Hubert, & Burack, 2006; Mottron et al., 2013), along with evidence that sometimes less robust brain responses and less accurate behavioral responses to domain-specific complex stimuli such as vocal pitch processing and phonemic/semantic discrimination processing in some individuals with ASD are taken as evidence for deficits in information integration for global processing or deficits in higher-level linguistic processing (Cantiani et al., 2016). The neural complexity hypothesis suggested that superior performance for simple tone processing in the primary auditory cortex, along with impaired complex perceptual performance in the associative cortex, is autism-specific (Bertone, Mottron, Jelenic, & Faubert, 2005). Happé and Frith (2006) proposed the “weak central coherence” theory referring to the detail-focused processing bias in individuals with ASD. The weak central coherence (WCC) theory pointed out that the processing bias for the local/detail-focused/lower-level information over global/meaning-focused/higher-level information may impose a negative impact on higher-level integrated processing (Happé & Frith, 2006). It is necessary for research to determine whether there is an impact of such a bias. A deficit in theory of mind (ToM) is a hallmark of ASD as mentioned above. A meta-analysis on ToM development showed that there is a two-year or greater developmental timing difference in false belief performance between Chinese-speaking children and children in North American culture (Liu, Wellman, Tardif, & Sabbagh, 2008). This significant timing difference suggests that learning a tonal language such as Cantonese or Mandarin, or growing up in Chinese culture, enhances the development of ToM. However, we do not have direct evidence regarding whether learning a tonal language will enhance the development of ToM in children with ASD. Future studies need to test ToM in Chinese-speaking individuals with ASD directly.

The three neurophysiological studies on Mandarin-speaking children with ASD are consistent with the WCC theory mentioned above. You and colleagues proposed that different from children with other common developmental disorders, such as children with dyslexia and developmental language disorder (DLD), children with ASD do not have categorical perception deficits (CP deficit), but instead have a categorical precision deficit (CPR deficit). CPR allows perception of allophonic differences (You, Serniclaes, Rider, & Chabane, 2017). This proposal was based on the evidence that children with ASD made more categorical judgments than TD controls did on a natural vowel continuum, yet did not show categorical perception deficits of vowels and consonants. One important issue here is whether the vowel versus consonant difference in terms of categorical precision would be applied to lexical tone and consonant differences, since lexical tone is largely superimposed on the vowel part of the syllable. You et al. (2017) used a four-parameter logistic model including perceptual boundary, slope, and asymptotes of the identification

function to assess categorical precision. Due to differences in analysis, it is unclear if the children with ASD in Chen et al. (2016) had a CPR deficit to lexical tone processing as well. But it is clear that children with and without ASD were both near chance level when discriminating within-category lexical tone contrasts (Fig. 4 in Chen et al., 2016), and children with ASD had a shallower identification slopes, suggesting continuous perception rather than CP.

Recently, there has been a lively discussion on spectral versus temporal auditory processing in the ASD literature (e.g., Huang et al., 2017; Kasai, Hashimoto, Kawakubo, Yumoto, Kamio, Itoh, Koshida, Iwanami, Nakagome, Fukuda, Yamasue, Yamada, Abe, Aoki, Kato, 2005; Lambrechts et al., 2017; Lepistö et al., 2006; see Haesen et al., 2011 for a review). Spectral information is generally processed in the right hemisphere, and rapid temporal dynamics for speech processing is more dominant in the left hemisphere (Zatorre, Evan, Meyer, & Gjedde, 1992). However, the acoustic properties of auditory cues interact with the function of the cues in a complex way. As Zatorre and Gandour (2008) pointed out, the right hemisphere dominance for pitch processing may be altered to the left when the pitch is linguistically relevant, as in lexical tone. Indeed, Mandarin speakers have shown larger left hemisphere responses to between-category lexical tone contrasts and marginally larger right hemisphere responses for within-category lexical tone contrasts (e.g., Xi, Zhang, Shu, Zhang, & Li, 2010).

Frequency differences are the primary cues to distinguish the four lexical tones in Mandarin. These differ from the spectral differences of vowels and consonants in serving as the carrier frequency (fundamental). There are some intrinsic durational differences among the four lexical tones. For example, as reported in Shen (1990) and other studies, Tone 3 is consistently longer than Tone 2, while Tone 4 has the shortest duration among the four (Lin, 1965; Shen, 1990). Future studies should examine how the intrinsic durational differences among lexical tone categories in Chinese will influence the developmental patterns of speech perception and production in the ASD individuals. Future studies comparing the neural responses to spectral versus durational cues for lexical categories would provide further evidence about the neural specialization related to atypical lexical tone processing in Chinese-speaking individuals with ASD.

We currently have no data on neuroplasticity in response to intervention. Children who are diagnosed with ASD usually receive language and behavioral intervention. A two-day perceptual training on Thai lexical tone alters the neural responses to the training tones in native speakers of a tonal language (i.e., Mandarin Chinese) and in those of a non-tonal language (i.e., English) (Kaan et al., 2008). Furthermore, research shows that there is an association between the range of brain activation and degree of lexical tone learning in speech and word training paradigms (Wong, Perrachione, & Parrish, 2007). Novel neural stimulation methods, such as the HD-tDCS method presented in the pilot study above, can also provide a valuable way to investigate intervention-related neuroplasticity in Chinese-speaking individuals with ASD.

13.7 Conclusion

The behavioral and electrophysiological evidence accumulated thus far on ASD in Chinese/tonal languages suggests that brain activation patterns and behavioral measures of lexical tone perception are not typical in Chinese-speaking individuals with ASD, and that these individuals with ASD also show atypical pitch production that is similar to non-tonal language speakers with ASD. Future studies need to systematically examine the effect of the stimuli, the relationship between lexical tone and music, and brain plasticity in response to lexical tone learning in order to fill the vast gap in the literature on the topic of Chinese-speaking individuals with ASD.

Appendix 1. Behavioral Studies on Prosody and Lexical Tone in Chinese ASD

Study	Participants	Method	Results
Chan and To (2016)	19 Cantonese adults with HFA 19 controls	Stimuli: Frog books Task: Narrative production	Larger pitch variations in HFA; no difference in the amount and diversity of SFPs
Chen et al. (2016)	11 boys with ASD and 14 TD 6–8 years old, Mandarin-speaking	Stimuli: /i/ in Tone 1 and Tone 2 Tasks: Pitch contour identification and discrimination 11 steps from 230 to 290 Hz	Wider ID boundary and poor between-category discrimination accuracy in children with ASD
Jiang et al. (2015)	17 Mandarin-speaking individuals (age: 6.0–16.2) with HFA and 17 age-matched control	Five music tone sequences DISC and ID Disyllabic verb–object linguistic pitch contours DISC and ID	Melodic contour ID: ASD > TD Melodic contour DISC: ASD = TD Intonation DISC: ASD < TD Intonation ID: ASD < TD
Su et al. (2014)	Younger group: 14 ASD, 14 TD Older group: 14 ASD, 14 TD	Computer-based question/statement task 4 conditions: 2 (prosody/semantics) by 2 (question/statement)	Statement reading with wh-words: ASD-young < TD; ASD-old = TD-old; TD-old = TD young Question reading with wh-words: no group differences

(continued)

(continued)

Study	Participants	Method	Results
Li et al. (2013)	13 Cantonese children with ASD and 13 age-matched TD (8.3–12.9)	16 stories to include 4 conditions: prosody-only, SPF-only, both, and neither	Judgment of speaker’s belief: ASD = TD Judgment of speaker’s intention: TD: well above chance in “both” and “SPF-only” condition, slightly above chance for “prosody-only,” below chance for neither ASD: below chance in all conditions
Wang and Tsao (2015)	25 TD, 25 HFA ASD; 6–11 years	Emotional prosody identification; pragmatic and social adaptive abilities; language abilities	Neutral semantic condition: Word context: ASD = TD Sentence context: ASD < TD identifying happy prosody Emotionally relevant semantic condition: Word context: ASD < TD identifying happy prosody Sentence context: ASD = TD Correlation between prosody identification skills and pragmatic function: positive

Appendix 2. Neurophysiological Studies on Mandarin Lexical Tone Processing in Children with ASD

Study	Participants	Method	Results
Yu et al. (2015)	EXP1: 17 ASD (6.9–12.4 years) 15 TD (7.7–11.8 years) EXP2: 16 ASD (7.9–12 years) 18 TD (6.9–12.4 years)	EXP1: Tone: 216–299 Hz Real word: /bai2/-/bai4/ Nonword: /rai2/-rai4/ EXP2: Hummed speech /bai2/-/bai4/	EXP1: MMN: Pure tone: enhanced at Cz but not Fz; marginally shorter latency Speech (real word): diminished Speech (nonword): MMN only for TD at Fz EXP1: P3a: Pure tone: enhanced at Cz Real word: enhanced at Cz, delayed EXP2: MMN Enhanced at Cz EXP2: P3a Reduced at Cz, marginally delayed
Wang et al. (2017)	16 ASD; 15 TD, 8–13 years old	/ba2/-/ba4/ continuum: steps 1, 5, and 9	Speech MMN: TD: between-category > within-category ASD: similar MMN latency and amplitude for between and within Harmonic MMN: MMN amplitude: ASD > TD No P3a differences of group or stimulus conditions Inter-trial phase coherence (ITPC): Theta band speech: ITPC ASD > TD TD: between-category > within-category ASD: between-category = within-category Theta band non-speech: ITPC ASD > TD Beta band speech: ITPC ASD > TD Beta band non-speech: ITPC ASD > TD
Huang et al. (2017)	Pure tone 22 ASD, 9.6 years old 20 TD, 9.4 years old Vowel 18 ASD, 9.8 years old 17 TD, 9.4 years old	Pure tone duration contrast 350 ms versus 250 ms Vowel duration contrast	MMN Pure tone: diminished and delayed Vowel: no group difference P3a Pure tone: no group difference Vowel: diminished

References

- Alexander, J. A., Bradlow, A. R., Ashley, R. D., & Wong, P. (2011). Music-melody perception in tone-language and non-tone-language speakers. In *Proceedings of the Psycholinguistic Representation of Tone Conference* (pp. 1–4), Hong Kong.
- Amatachaya, A., Jensen, M. P., Patjanasoonorn, N., Auvichayapat, N., Suphakunpinyo, C., Janjarasjitt, S., ... Auvichayapat, P. (2015). The short-term effects of transcranial direct current stimulation on electroencephalography in children with autism: A randomized crossover controlled trial. *Behavioural Neurology*, *2015*, 928631.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*, 5th ed.. <https://doi.org/10.1176/appi.books.9780890425596>.
- Asperger, H. (1944). Autistic psychopathy in childhood. In U. Frith (Ed.), *Autism and asperger syndrome* (pp. 37–92). Cambridge, UK: Cambridge University Press.
- Baltaxe, C., & Simmons, J. (1985). Prosodic development in normal and autistic children. In E. Schopler & G. Mesibov (Eds.), *Communication problems in autism* (pp. 95–125). New York: Plenum.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, *21*(1), 37–46.
- Bertone, A., Mottron, L., Jelenic, P., & Faubert, J. (2005). Enhanced and diminished visuo-spatial information processing in autism depends on stimulus complexity. *Brain*, *128*(10), 2430–2441. <https://doi.org/10.1093/brain/awh561>.
- Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: Evidence for bidirectionality between the domains of language and music. *PLoS ONE*, *8*(4), e60676. <https://doi.org/10.1371/journal.pone.0060676>.
- Bikson, M., Inoue, M., Akiyama, H., Deans, J. K., Fox, J. E., Miyakawa, H., & Jefferys, J. G. R. (2004). Effects of uniform extracellular DC electric fields on excitability in rat hippocampal slices in vitro. *The Journal of Physiology*, *557*(Pt 1), 175–190.
- Bindman, L. J., Lippold, O. C., & Redfearn, J. W. (1964). The action of brief polarizing currents on the cerebral cortex of the rat (1) during current flow and (2) in the production of long-lasting after-effects. *The Journal of Physiology*, *172*, 369–382.
- Bishop, D. V. M., Hardiman, M., Uwer, R., & von Suchodoletz, W. (2007). Maturation of the long-latency auditory ERP: Step function changes at start and end of adolescence. *Developmental Science*, *10*(5), 565–575.
- Boets, B., Verhoeven, J., Wouters, J., & Steyaert, J. (2015). Fragile spectral and temporal auditory processing in adolescents with autism spectrum disorder and early language delay. *Journal of Autism and Developmental Disorders*, *45*(6), 1845–1857.
- Bone, D., Lee, C.-C., Black, M. P., Williams, M. E., Lee, S., Levitt, P., & Narayanan, S. (2014). The psychologist as an interlocutor in autism spectrum disorder assessment: Insights from a study of spontaneous prosody. *Journal of Speech, Language, and Hearing Research*, *57*(4), 1162–1177.
- Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A. M. (2003). Enhanced pitch sensitivity in individuals with autism: A signal detection analysis. *Journal of Cognitive Neuroscience*, *15*(2), 226–235. <https://doi.org/10.1162/089892903321208169>.
- Bonnel, A., McAdams, S., Smith, B., Berthiaume, C., Bertone, A., Ciocca, V., et al. (2010). Enhanced pure-tone pitch discrimination among persons with autism but not Asperger syndrome. *Neuropsychologia*, *48*, 2465–2475.
- Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*(1), 33–69.
- Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, *46*(Pt 2–3), 217–243. <https://doi.org/10.1177/00238309030460020801>.
- 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 Research: Official Journal of the International Society for Autism Research*, 7(2), 237–244.
- Bruneau, N., Bonnet-Brilhault, F., Gomot, M., Adrien, J.-L., & Barthélémy, C. (2003). Cortical auditory processing and communication in children with autism: Electrophysiological/behavioral relations. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 51(1), 17–25.
- Bruneau, N., Roux, S., Adrien, J. L., & Barthélémy, C. (1999). Auditory associative cortex dysfunction in children with autism: Evidence from late auditory evoked potentials (N1 wave-T complex). *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 110(11), 1927–1934.
- Burnham, D., & Francis, E. (1997). The role of linguistic experience in the perception of Thai tones. In A. Abramson (Ed.). Bangkok: Chulalongkorn University Press.
- Cantiani, C., Choudhury, N. A., Yu, Y. H., Shafer, V. L., Schwartz, R. G., & Benasich, A. A. (2016). From sensory perception to lexical-semantic processing: An ERP study in non-verbal children with autism. *PLoS ONE*, 11(8), e0161637. <https://doi.org/10.1371/journal.pone.0161637>.
- 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. *Proceedings of the National Academy of Sciences of the United States of America*, 100(9), 5567–5572. <https://doi.org/10.1073/pnas.0835631100>.
- Čeponienė, R., Rinne, T., & Näätänen, R. (2002). Maturation of cortical sound processing as indexed by event-related potentials. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 113(6), 870–882.
- Chan, C. Y., Hounsgaard, J., & Nicholson, C. (1988). Effects of electric fields on transmembrane potential and excitability of turtle cerebellar Purkinje cells in vitro. *The Journal of Physiology*, 402, 751–771.
- Chan, K. K. L., & To, C. K. S. (2016). Do individuals with high-functioning autism who speak a tone language show intonation deficits? *Journal of Autism and Developmental Disorders*, 46(5), 1784–1792. <https://doi.org/10.1007/s10803-016-2709-5>.
- Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007). Mismatch negativity to pitch contours is influenced by language experience. *Brain Research*, 1128(1), 148–156. <https://doi.org/10.1016/j.brainres.2006.10.064>.
- Chang, C. H. C., & Kuo, W.-J. (2020). Neural processing of tone sandhi in production and perception: The case of Mandarin tone 3 sandhi. In H. M. Liu, F.M. Tsao, & P. Li (Eds.), *Speech perception, production and acquisition—Multidisciplinary approaches in Chinese language*. Berlin: Springer.
- Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley, CA.: University of California Press.
- Chen, F., Yan, N., Pan, X., Yang, F., Ji, Z., Wang, L., & Peng, G. (2016). Impaired categorical perception of Mandarin tones and its relationship to language ability in autism spectrum disorders (pp. 233–237). *Interspeech*.
- Cheng, Y.-Y., Wu, H.-C., Tzeng, Y.-L., Yang, M.-T., Zhao, L.-L., & Lee, C.-Y. (2013). The development of mismatch responses to Mandarin lexical tones in early infancy. *Developmental Neuropsychology*, 38(5), 281–300.
- Cheng, Y.-Y., Wu, H.-C., Tzeng, Y.-L., Yang, M.-T., Zhao, L.-L., & Lee, C.-Y. (2015). Feature-specific transition from positive mismatch response to mismatch negativity in early infancy: Mismatch responses to vowels and initial consonants. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 96(2), 84–94.
- Chevallier, C., Noveck, I., Happé, F., & Wilson, D. (2011). What's in a voice? Prosody as a test case for the theory of mind account of autism. *Neuropsychologia*, 49(3), 507–517. <https://doi.org/10.1016/j.neuropsychologia.2010.11.042>.
- Choudhury, N., & Benasich, A. A. (2011). Maturation of auditory evoked potentials from 6 to 48 months: Prediction to 3 and 4 year language and cognitive abilities. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 122(2), 320–338.
- Crystal, D., & Quirk, R. (1964). Systems of prosodic and paralinguistic features in English (pp. 10–12). Mouton.

- Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication, 71*, 10–49.
- Curtin, S., & Vouloumanos, A. (2013). Speech preference is associated with autistic-like behavior in 18-month-olds at risk for autism spectrum disorder. *Journal of Autism and Developmental Disorders, 43*(9), 2114–2120.
- Curtin, S., & Werker, J. F. (2018). PRIMIR on tone. *Frontiers in Psychology, 9*, 1007. <https://doi.org/10.3389/fpsyg.2018.01007>.
- Cutler, A., & Bruck, A. (1974). On saying what you mean without meaning what you say. In M. Galy & R. Fox (Eds.) (pp. 117–127). Chicago, Illinois.
- Culter, A., & Isard, S. D. (1980). The production of prosody (Vol. 1, pp. 245–269). London: Academic Press.
- Diehl, J. J., Berkovits, L., & Harrison, A. (2010). Is prosody a diagnostic and cognitive bellwether of autism spectrum disorders. In *Speech disorders: Causes, treatments, and social effects* (pp. 159–176). New York: Nova Science.
- Diehl, J. J., & Paul, R. (2013). Acoustic and perceptual measurements of prosody production on the profiling elements of prosodic systems in children by children with autism spectrum disorders. *Applied Psycholinguistics, 34*(01), 135–161.
- Diehl, J. J., Watson, D. G., Bennetto, L., McDonough, J., & Gunlogson, C. (2009). An acoustic analysis of prosody in high-functioning autism. *Applied Psycholinguistics, 30*, 385–404.
- Edwards, D., Cortes, M., Datta, A., Minhas, P., Wassermann, E. M., & Bikson, M. (2013). Physiological and modeling evidence for focal transcranial electrical brain stimulation in humans: a basis for high-definition tDCS. *NeuroImage, 74*, 266–275. <https://doi.org/10.1016/j.neuroimage.2013.01.042>
- Eigsti, I. M., & Fein, D. A. (2013). More is less: Pitch discrimination and language delays in children with optimal outcomes from autism: Pitch perception, language delay, and outcomes. *Autism Research, 6*(6), 605–613. <https://doi.org/10.1002/aur.1324>.
- Eigsti, I. M., Schuh, J., Mencl, E., Schultz, R. T., & Paul, R. (2012). The neural underpinnings of prosody in autism. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence, 18*(6), 600–617. <https://doi.org/10.1080/09297049.2011.639757>.
- Falter, C. M., Noreika, V., Wearden, J. H., & Bailey, A. J. (2012). More consistent, yet less sensitive: Interval timing in autism spectrum disorders. *Quarterly Journal of Experimental Psychology, 65*(11), 2093–2107. <https://doi.org/10.1080/17470218.2012.690770>.
- 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. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology, 114*(9), 1671–1680.
- Filipe, M. G., Watson, L., Vicente, S. G., & Frota, S. (2017). Atypical preference for infant-directed speech as an early marker of autism spectrum disorders? A literature review and directions for further research. *Clinical Linguistics & Phonetics, 1*–19.
- Friederici, A. D., Friedrich, M., & Weber, C. (2002). Neural manifestation of cognitive and precognitive mismatch detection in early infancy. *NeuroReport, 13*(10), 1251–1254.
- Friedrich, M., Herold, B., & Friederici, A. D. (2009). ERP correlates of processing native and non-native language word stress in infants with different language outcomes. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior, 45*(5), 662–676.
- Fuchigami, T., Okubo, O., Ejiri, K., Fujita, Y., Kohira, R., Noguchi, Y., ... Harada, K. (1995). Developmental changes in P300 wave elicited during two different experimental conditions. *Pediatric Neurology, 13*(1), 25–28.
- Fusaroli, R., Lambrechts, A., Bang, D., Bowler, D. M., & Gaigg, S. B. (2017). Is voice a marker for autism spectrum disorder? A systematic review and meta-analysis. *Autism Research: Official Journal of the International Society for Autism Research, 10*(3), 384–407.
- Gandour, J. T. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics, 11*, 149–175.

- Gandour, J. T. (1984). Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics*, 12, 235–261.
- Gandour, J. T., & Harshman, R. A. (1978). Cross-language difference in tone perception: A multidimensional scaling investigation. *Language and Speech*, 21, 1–33.
- Gómez, L., Vidal, B., Maragoto, C., Morales, L. M., Berrillo, S., Vera Cuesta, H., ... Robinson, M. (2017). Non-invasive brain stimulation for children with autism spectrum disorders: A short-term outcome study. *Behavioral Sciences (Basel, Switzerland)*, 7(3).
- Gomot, M., Belmonte, M. K., Bullmore, E. T., Bernard, F. A., & Baron-Cohen, S. (2008). Brain hyper-reactivity to auditory novel targets in children with high-functioning autism. *Brain: A Journal of Neurology*, 131(Pt 9), 2479–2488. <https://doi.org/10.1093/brain/awn172>.
- Green, H., & Tobin, Y. (2009). Prosodic analysis is difficult ... but worth it: A study in high functioning autism. *International Journal of Speech-Language Pathology*, 11(4), 308–315.
- Grossman, R. B., & Tager-Flusberg, H. (2012). “Who said that?” Matching of low- and high-intensity emotional prosody to facial expressions by adolescents with ASD. *Journal of Autism and Developmental Disorders*, 42(12), 2546–2557.
- Haesen, B., Boets, B., & Wagemans, J. (2011). A review of behavioural and electrophysiological studies on auditory processing and speech perception in autism spectrum disorders. *Research in Autism Spectrum Disorders*, 5(2), 701–714.
- Hallé, P. A., Chang, Y.-C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3), 395–421.
- Hambly, H., Wren, Y., McLeod, S., & Roulstone, S. (2013). The influence of bilingualism on speech production: A systematic review. *International Journal of Language & Communication Disorders*, 48(1), 1–24.
- Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 36(1), 5–25. <https://doi.org/10.1007/s10803-005-0039-0>.
- Happé, F. G. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, 48, 101–119.
- Heaton, P. (2003). Pitch memory, labeling and disembedding in autism. *Journal of Child Psychology and Psychiatry*, 44(4), 543–551.
- Heaton, P. (2005). Interval and contour processing in autism. *Journal of Autism and Developmental Disorders*, 35, 787–793.
- Heaton, P., Hermelin, B., & Pring, L. (1999). Can children with autistic spectrum disorders perceive affect in music? An experimental investigation. *Psychological Medicine*, 29(6), 1405–1410.
- Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazim, T. (2009). Language specific prosodic preferences during the first half year of life: Evidence from German and French infants. *Infant Behavior & Development*, 32(3), 262–274.
- Hua, Z., & Dodd, B. (2000). The phonological acquisition of Putonghua (Modern Standard Chinese). *Journal of Child Language*, 27(1), 3–42.
- Huang, A. X., Jia, M., & Wheeler, J. J. (2013). Children with autism in the people’s Republic of China: Diagnosis, legal issues, and educational services. *Journal of Autism and Developmental Disorders*, 43(9), 1991–2001. <https://doi.org/10.1007/s10803-012-1722-6>.
- Huang, D., Yu, L., Wang, X., Fan, Y., Wang, S., & Zhang, Y. (2017). Distinct patterns of discrimination and orienting for temporal processing of speech and nonspeech in Chinese children with autism: An event-related potential study. *The European Journal of Neuroscience*. <https://doi.org/10.1111/ejn.13657>.
- Jacobson, L., Koslowsky, M., & Lavidor, M. (2012). tDCS polarity effects in motor and cognitive domains: A meta-analytical review. *Experimental Brain Research*, 216(1), 1–10. <https://doi.org/10.1007/s00221-011-2891-9>.
- Jansson-Verkasalo, E., Čeponienė, R., Kielinen, M., Suominen, K., Jäntti, V., Linna, S. L., ... Näätänen, R. (2003). Deficient auditory processing in children with Asperger syndrome, as indexed by event-related potentials. *Neuroscience Letters*, 338(3), 197–200.

- Järvinen-Pasley, A., & Heaton, P. (2007). Evidence for reduced domain-specificity in auditory processing in autism. *Developmental Science*, *10*(6), 786–793.
- Järvinen-Pasley, A., Peppé, S., King-Smith, G., & Heaton, P. (2008). The relationship between form and function level receptive prosodic abilities in autism. *Journal of Autism and Developmental Disorders*, *38*, 1328–1340.
- Jiang, J., Liu, F., Wan, X., & Jiang, C. (2015). Perception of melodic contour and intonation in autism spectrum disorder: Evidence from Mandarin speakers. *Journal of Autism and Developmental Disorders*, *45*(7), 2067–2075. <https://doi.org/10.1007/s10803-015-2370-4>.
- Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge: MIT Press.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*(1), 1–23. <https://doi.org/10.1006/cogp.1995.1010>.
- Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*(3), 675–687.
- Kaan, E., Barkley, C. M., Bao, M., & Wayland, R. (2008). Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: An event-related potentials training study. *BMC Neuroscience*, *9*, 53. <https://doi.org/10.1186/1471-2202-9-53>.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, *2*, 217–250.
- Kasai, K., Hashimoto, O., Kawakubo, Y., Yumoto, M., Kamio, S., Itoh, K., Koshida, I., Iwanami, A., Nakagome, K., Fukuda, M., Yamasue, H., Yamada, H., Abe, O., Aoki, S., & Kato, N. (2005). Delayed automatic detection of change in speech sounds in adults with autism: a magnetoencephalographic study. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *116*(7), 1655–1664. <https://doi.org/10.1016/j.clinph.2005.03.007>.
- Koelsch, S., Gunter, T. C., Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical “language-network” serves the processing of music. *NeuroImage*, *17*(2), 956–966.
- Kuhl, P. K., Coffrey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological measures. *Developmental Science*, *8*(1), F1–F12.
- Kujala, T., Aho, E., Lepistö, T., Jansson-Verkasalo, E., Nieminen-von Wendt, T., von Wendt, L., & Näätänen, R. (2007). Atypical pattern of discriminating sound features in adults with Asperger syndrome as reflected by the mismatch negativity. *Biological Psychology*, *75*(1), 109–114.
- Kushnerenko, E., Čepionienė, R., Balan, P., Fellman, V., & Näätänen, R. (2002). Maturation of the auditory change detection response in infants: A longitudinal ERP study. *NeuroReport*, *13*(15), 1843–1848.
- Kushnerenko, E., Winkler, I., Horváth, J., Näätänen, R., Pavlov, I., Fellman, V., & Huotilainen, M. (2007). Processing acoustic change and novelty in newborn infants. *The European Journal of Neuroscience*, *26*(1), 265–274. <https://doi.org/10.1111/j.1460-9568.2007.05628.x>.
- Kwok, H. (1984). *Sentence particles in Cantonese*. Hong Kong: Centre of Asian Studies, The University of Hong Kong.
- Lambrechts, A., Falter-Wagner, C. M., & van Wassenhove, V. (2017). Diminished neural resources allocation to time processing in Autism Spectrum Disorders. *NeuroImage: Clinical*. <https://doi.org/10.1016/j.nicl.2017.09.023>.
- Leather, J. (1983). Speaker normalization in perception of lexical tones. *Journal of Phonetics*, *11*, 373–382.
- Lee, C.-Y., & Cheng, Y.-Y. (2020). Neurophysiological studies of Mandarin lexical tone acquisition in the early childhood. In H. M. Liu, F. M. Tsao, & P. Li (Eds.), *Speech perception, production and acquisition—Multidisciplinary approaches in Chinese language*. Berlin: Springer.
- Lee, Y. S., Vokoch, D., & Wurm, L. H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research*, *25*(5), 527–542.
- 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 Research*, *1066*(1–2), 147–157.

- 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. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 117(10), 2161–2171. <https://doi.org/10.1016/j.clinph.2006.06.709>.
- Leppänen, P. H. T., Richardson, U., Pihko, E., Eklund, K. M., Guttorm, T. K., Aro, M., & Lyytinen, H. (2002). Brain responses to changes in speech sound durations differ between infants with and without familial risk for dyslexia. *Developmental Neuropsychology*, 22(1), 407–422. https://doi.org/10.1207/S15326942dn2201_4.
- Li, J. P. W., Law, T., Lam, G. Y. H., & To, C. K. S. (2013). Role of sentence-final particles and prosody in irony comprehension in Cantonese-speaking children with and without autism spectrum disorders. *Clinical Linguistics & Phonetics*, 27(1), 18–32. <https://doi.org/10.3109/02699206.2012.734893>.
- Lin, M. C. (1965). The pitch indicator and the pitch characteristics of tones in Standard Chinese. *Acta Acustica (China)*, 2, 8–15.
- Liu, D., Wellman, H. M., Tardif, T., & Sabbagh, M. A. (2008). Theory of mind development in Chinese children: A meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology*, 44(2), 523–531. <https://doi.org/10.1037/0012-1649.44.2.523>.
- Liu, H.-M., Chen, Y., & Tsao, F.-M. (2014). Developmental changes in mismatch responses to mandarin consonants and lexical tones from early to middle childhood. *PLoS ONE*, 9(4), e95587. <https://doi.org/10.1371/journal.pone.0095587>.
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism diagnostic observation schedule* (2nd ed.). Torrance, CA: Western Psychological Services.
- Määttä, S., Saavalainen, P., Könönen, M., Pääkkönen, A., Muraja-Murro, A., & Partanen, J. (2005). Processing of highly novel auditory events in children and adults: An event-related potential study. *NeuroReport*, 16(13), 1443–1446.
- Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience*, 4(5), 540–545.
- Maister, L., & Plaisted-Grant, K. C. (2011). Time perception and its relationship to memory in autism spectrum conditions. *Developmental Science*, 14(6), 1311–1322.
- McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: A critical review. *International Journal of Language and Communication Disorders*, 38, 325–350.
- McCann, J., Peppé, S., Gibbon, F. E., O'Hare, A., & Rutherford, M. (2007). Prosody and its relationship to language in school-aged children with high-functioning autism. *International Journal of Language & Communication Disorders*, 42(6), 682–702.
- 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.
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, 19(3), 712–723. <https://doi.org/10.1093/cercor/bhn120>.
- Morr, M. L., Shafer, V. L., Kreuzer, J. A., & Kurtzberg, D. (2002). Maturation of mismatch negativity in typically developing infants and preschool children. *Ear and Hearing*, 23(2), 118–136.
- 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. *Neuroscience and Biobehavioral Reviews*, 37(2), 209–228. <https://doi.org/10.1016/j.neubiorev.2012.11.016>.
- Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, 36(1), 27–43. <https://doi.org/10.1007/s10803-005-0040-7>.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 118(12), 2544–2590.
- 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.

- 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. *Journal of Autism and Developmental Disorders*, 42(4), 499–511. <https://doi.org/10.1007/s10803-011-1264-3>.
- Nakai, Y., Takashima, R., Takiguchi, T., & Takada, S. (2014). Speech intonation in children with autism spectrum disorder. *Brain & Development*, 36(6), 516–522. <https://doi.org/10.1016/j.bra.indev.2013.07.006>.
- Norise, C., & Hamilton, R. H. (2016). Non-invasive brain stimulation in the treatment of post-stroke and neurodegenerative Aphasia: Parallels, differences, and lessons learned. *Frontiers in Human Neuroscience*, 10, 675. <https://doi.org/10.3389/fnhum.2016.00675>.
- Oades, R. D., Walker, M. K., Geffen, L. B., & Stern, L. M. (1988). Event-related potentials in autistic and healthy children on an auditory choice reaction time task. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 6(1), 25–37.
- Ong, J. H., Tan, S. H., Chan, A. H. D., & Wong, F. C. K. (2020). The effect of musical experience and disorder on lexical tone perception, production, and learning: A review. In H. M. Liu, F. M. Tsao, & P. Li (Eds.), *Speech perception, production and acquisition—Multidisciplinary approaches in Chinese language*. Berlin: Springer.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6(7), 674–681. <https://doi.org/10.1038/nn1082>.
- Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*, 2, 142. <https://doi.org/10.3389/fpsyg.2011.00142>.
- Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*, 308, 98–108. <https://doi.org/10.1016/j.heares.2013.08.011>.
- Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, 10(6), 717–733.
- Paul, R., Augustyn, A., Klin, A., & Volkmar, F. R. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 35(2), 205–220.
- Paul, R., Chawarska, K., Fowler, C., Cicchetti, D., & Volkmar, F. (2007). Listen to my children and you shall hear: Auditory preferences in toddlers with autism spectrum disorders. *Journal of Speech, Language and Hearing Research*, 50, 1350–1364.
- Pfordresher, P. Q., & Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception & Psychophysics*, 71(6), 1385–1398. <https://doi.org/10.3758/APP.71.6.1385>.
- Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation (MIT dissertation)*.
- 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: Official Journal of the International Federation of 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. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 111(2), 220–236.
- Purpura, D. P., & Mcmurtry, J. G. (1965). Intracellular activities and evoked potential changes during polarization of motor cortex. *Journal of Neurophysiology*, 28, 166–185.
- Quigley, J., McNally, S., & Lawson, S. (2016). Prosodic patterns in interaction of low-risk and at-risk-of-autism spectrum disorders infants and their mothers at 12 and 18 months. *Language Learning and Development*, 12(3), 295–310.
- Rahman, A., Toshev, P. K., & Bikson, M. (2014). Polarizing cerebellar neurons with transcranial direct current stimulation. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 125(3), 435–438. <https://doi.org/10.1016/j.clinph.2013.10.003>.

- Reato, D., Rahman, A., Bikson, M., & Parra, L. C. (2013). Effects of weak transcranial alternating current stimulation on brain activity—A review of known mechanisms from animal studies. *Frontiers in Human Neuroscience*, 7, 687.
- Roesler, C. P., Flax, J., MacRoy-Higgins, M., Fermano, Z., Morgan-Byrne, J., & Benasich, A. A. (2013). Sensory desensitization training for successful net application and EEG/ERP acquisition in difficult to test children. *Communication Disorders Quarterly*, 35(1), 14–20.
- Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of Autism & Developmental Disorders*, 32, 189–194.
- Rutter, M., Le Couteur, A., & Lord, C. (2003). *ADI-R. Autism diagnostic interview revised. Manual*. Los Angeles: Western Psychological Services.
- Sambeth, A., Ruohio, K., Alku, P., Fellman, V., & Huotilainen, M. (2008). Sleeping newborns extract prosody from continuous speech. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 119(2), 332–341. <https://doi.org/10.1016/j.clinph.2007.09.144>.
- Scherg, M., & Von Cramon, D. (1986). Evoked dipole source potentials of the human auditory cortex. *Electroencephalography and Clinical Neurophysiology*, 65(5), 344–360.
- Schneider, H. D., & Hopp, J. P. (2011). The use of the Bilingual Aphasia Test for assessment and transcranial direct current stimulation to modulate language acquisition in minimally verbal children with autism. *Clinical Linguistics & Phonetics*, 25(6–7), 640–654. <https://doi.org/10.3109/02699206.2011.570852>.
- Shafer, V. L., Morr, M. L., Kreuzer, J. A., & Kurtzberg, D. (2000). Maturation of mismatch negativity in school-age children. *Ear and Hearing*, 21(3), 242–251.
- Shafer, V., Morr, M., Datta, H., Kurtzberg, D., Schwartz, R. G. (2005). Neurophysiological indexes of speech processing deficits in children with specific language impairment. *Journal of Cognitive Neuroscience*, 17, 1168–1180. <https://doi.org/10.1162/0898929054475217>.
- Shafer, V. L., Shucard, D. W., & Jaeger, J. J. (1999). Electrophysiological indices of cerebral specialization and the role of prosody in language acquisition in 3-month-old infants. *Developmental Neuropsychology*, 15(1), 73–109.
- Shafer, V. L., Yu, Y. H., & Datta, H. (2010). Maturation of speech discrimination in 4- to 7-year-old children as indexed by event-related potential mismatch responses. *Ear and Hearing*, 31(6), 735–745. <https://doi.org/10.1097/AUD.0b013e3181e5d1a7>.
- Shafer, V. L., Yu, Y. H., & Datta, H. (2011). The development of English vowel perception in monolingual and bilingual infants: Neurophysiological correlates. *Journal of Phonetics*, 39(4), 527–545. <https://doi.org/10.1016/j.wocn.2010.11.010>.
- Shafer, V. L., Yu, Y. H., & Garrido-Nag, K. (2012). Neural mismatch indices of vowel discrimination in monolingually and bilingually exposed infants: Does attention matter? *Neuroscience Letters*, 526(1), 10–14. <https://doi.org/10.1016/j.neulet.2012.07.064>.
- Sharda, M., Subhadra, T. P., Sahay, S., Nagaraja, C., Singh, L., Mishra, R., ... Singh, N. C. (2010). Sounds of melody—Pitch patterns of speech in autism. *Neuroscience Letters*, 478(1), 42–45. <https://doi.org/10.1016/j.neulet.2010.04.066>.
- Sharma, A., Glick, H., Deeves, E., & Duncan, E. (2015). The P1 biomarker for assessing cortical maturation in pediatric hearing loss: A review. *Otorinolaringologia*, 65(4), 103–114.
- Shen, X. S. (1990). *The prosody of Mandarin Chinese* (Vol. 118). Berkeley: University of California Press.
- Simmons, J. Q., & Baltaxe, C. (1975). Language patterns of adolescent autistics. *Journal of Autism and Childhood Schizophrenia*, 5(4), 333–351.
- Singh, L. (2020). Early word recognition and word learning in Mandarin learning children. In H. M. Liu, F. M. Tsao, & P. Li (Eds.), *Speech perception, production and acquisition—Multidisciplinary approaches in Chinese language*. Berlin: Springer.
- Soltaninejad, Z., Nejati, V., & Ekhtiari, H. (2015). Effect of anodal and cathodal transcranial direct current stimulation on DLPFC on modulation of inhibitory control in ADHD. *Journal of Attention Disorders*. <https://doi.org/10.1177/1087054715618792>.

- Stanutz, S., Wapnick, J., & Burack, J. A. (2014). Pitch discrimination and melodic memory in children with autism spectrum disorders. *Autism: The International Journal of Research and Practice*, 18(2), 137–147. <https://doi.org/10.1177/1362361312462905>.
- Stefanics, G., Háden, G. P., Sziller, I., Balázs, L., Beke, A., & Winkler, I. (2009). Newborn infants process pitch intervals. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 120(2), 304–308. <https://doi.org/10.1016/j.clinph.2008.11.020>.
- Stevens, C. J., Keller, P. E., & Tyler, M. D. (2013). Tonal language background and detecting pitch contour in spoken and musical items. *Psychology of Music*, 41(1), 59–74. <https://doi.org/10.1177/0305735611415749>.
- Su, Y. (Esther), Jin, Y., Wan, G.-B., Zhang, J.-S., & Su, L.-Y. (2014). Interpretation of wh-words in Mandarin-speaking high-functioning children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 8(10), 1364–1372.
- Sun, X., Allison, C., Matthews, F. E., Sharp, S. J., Auyeung, B., Baron-Cohen, S., & Brayne, C. (2013). Prevalence of autism in mainland China, Hong Kong and Taiwan: A systematic review and meta-analysis. *Molecular Autism*, 4(1), 7.
- Szelag, E., Kowalska, J., Galkowski, T., & Pöppel, E. (2004). Temporal processing deficits in high-functioning children with autism. *British Journal of Psychology (London, England: 1953)*, 95(Pt 3), 269–282. <https://doi.org/10.1348/0007126041528167>.
- Tao, K. T. (1987). Brief report: Infantile autism in China. *Journal of Autism and Developmental Disorders*, 17(2), 289–296.
- Thiessen, E. D., Hill, E. A., & Saffran, J. F. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, 7, 53–71.
- Thorsen, N. G. (1980). A study of perception of sentence intonation—Evidence from Danish. *The Journal of the Acoustical Society of America*, 67(3), 1014–1030.
- Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs, N.J.: Prentice Hall.
- Tsao, F.-M., & Liu, H.-M. (2020). Lexical tone perception development in infancy. In H. M. Liu, F. M. Tsao, & P. Li (Eds.), *Speech perception, production and acquisition—Multidisciplinary approaches in Chinese language*. Berlin: Springer.
- Vaissiere, J. (1983). Language-independent prosodic features. In *Prosody: Models and measurements* (pp. 53–66). https://doi.org/10.1007/978-3-642-69103-4_5.
- Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, 10(2), 159–164.
- Wang, A. T., Dapretto, M., Hariri, A. R., Sigman, M., & Bookheimer, S. Y. (2004). Neural correlates of facial affect processing in children and adolescents with autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43(4), 481–490. <https://doi.org/10.1097/00004583-200404000-00015>.
- Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2006). Neural basis of irony comprehension in children with autism: The role of prosody and context. *Brain*, 129(4), 932–943.
- Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2007). Reading affect in the face and voice: Neural correlates of interpreting communicative intent in children and adolescents with autism spectrum disorders. *Archives of General Psychiatry*, 64(6), 698–708. <https://doi.org/10.1001/archpsyc.64.6.698>.
- Wang, J. E., & Tsao, F. M. (2015). Emotional prosody perception and its association with pragmatic language in school-aged children with high-function autism. *Research in Developmental Disabilities*, 37, 162–170. <https://doi.org/10.1016/j.ridd.2014.11.013>.
- 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. *Scientific Reports*, 7, 43254. <https://doi.org/10.1038/srep43254>.
- Wayland, R. P., & Guion, S. G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning*, 54(4), 681–712.
- Whitehouse, A. J. O., & Bishop, D. V. M. (2008). Do children with autism “switch off” to speech sounds? An investigation using event-related potentials. *Developmental Science*, 11(4), 516–524. <https://doi.org/10.1111/j.1467-7687.2008.00697.x>.

- Wong, P. C. M., Perrachione, T. K., & Parrish, T. B. (2007). Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping*, 28(10), 995–1006. <https://doi.org/10.1002/hbm.20330>.
- Wong, P. (2013). Perceptual evidence for protracted development in monosyllabic Mandarin lexical tone production in preschool children in Taiwan. *The Journal of the Acoustical Society of America*, 133(1), 434–443.
- Wong, P., Schwartz, R. G., & Jenkins, J. J. (2005). Perception and production of lexical tones by 3-year-old, Mandarin-speaking children. *Journal of Speech, Language, and Hearing Research: JSLHR*, 48(5), 1065–1079. [https://doi.org/10.1044/1092-4388\(2005\)074](https://doi.org/10.1044/1092-4388(2005)074).
- World Health Organization. (2004). ICD-10: International statistical classification of diseases and related health problems: Tenth revision, 2nd ed. *World Health Organization*. <https://apps.who.int/iris/handle/10665/42980>.
- Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, 170(1), 223–231. <https://doi.org/10.1016/j.neuroscience.2010.06.077>.
- Xu, Y., Gandour, J. T., & Francis, A. L. (2006a). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *The Journal of the Acoustical Society of America*, 120(2), 1063–1074.
- Xu, Y., Krishnan, A., & Gandour, J. T. (2006b). Specificity of experience-dependent pitch representation in the brainstem. *NeuroReport*, 17(15), 1601–1605. <https://doi.org/10.1097/01.wnr.0000236865.31705.3a>.
- Yau, S. C. (1980). Sentential connotations in Cantonese. *Fangyan*, 1, 35–52.
- You, R. S., Serniclaes, W., Rider, D., & Chabane, N. (2017). On the nature of the speech perception deficits in children with autism spectrum disorders. *Research in Developmental Disabilities*, 61, 158–171. <https://doi.org/10.1016/j.ridd.2016.12.009>.
- Yu, K., Wang, R., & Li, P. (2020). Native and nonnative processing of acoustic and phonological information of lexical tones in Chinese: Behavioral and neural correlates. In H. M. Liu, F. M. Tsao, & P. Li (Eds.), *Speech perception, production and acquisition—Multidisciplinary approaches in Chinese language*. Berlin: Springer.
- 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. *Journal of Autism and Developmental Disorders*, 45(11), 3656–3667. <https://doi.org/10.1007/s10803-015-2510-x>.
- Yu, Y. H., Shafer, V. L., & Sussman, E. S. (2017). Neurophysiological and behavioral responses of mandarin lexical tone processing. *Frontiers in Neuroscience*, 11, 95. <https://doi.org/10.3389/fnins.2017.00095>.
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science (New York, N.Y.)*, 256(5058), 846–849.
- Zatorre, R. J., & Gandour, J. T. (2008). Neural specializations for speech and pitch: Moving beyond the dichotomies. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1493), 1087–1104. <https://doi.org/10.1098/rstb.2007.2161>.
- Zhang, X., & Ji, C. (2005). Autism mental retardation of young children in China. *Biomedical and Environmental Sciences*, 18(5), 334.