The form and function processing of lexical tone and intonation in tone-language-speaking children with autism spectrum disorder

Wang, L., Xiao, S., Jiang, C., Hou, Q., Chan, A. H. D., Wong, P. C. M. and Liu, F. ORCID: https://orcid.org/0000-0002-7776-0222 (2023) The form and function processing of lexical tone and intonation in tone-language-speaking children with autism spectrum disorder. Journal of the Acoustical Society of America, 154. pp. 467-481. ISSN 0001-4966 doi: https://doi.org/10.1121/10.0020271 Available at https://centaur.reading.ac.uk/112521/

To link to this article DOI: http://dx.doi.org/10.1121/10.0020271

Publisher: Acoustical Society of America

The form and function processing of lexical tone and intonation in tone-language-speaking children with autism spectrum disorder

Li Wang[1], Sanrong Xiao[2], Cunmei Jiang[3], Qingqi Hou[4], Alice H. D. Chan[5], Patrick C. M. Wong[6,7], & Fang Liu[8]

[1] Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, 518057, China

[2] School of Public Administration, Nanchang University, Nanchang, 330031, China

[3] Music College, Shanghai Normal University, Shanghai, 200234, China

[4] Department of Music and Dance, Nanjing Normal University of Special Education, Nanjing, 210038, China

[5] Linguistics and Multilingual Studies, School of Humanities, Nanyang Technological University, 639798, Singapore

[6] Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

[7] Brain and Mind Institute, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

[8] School of Psychology and Clinical Language Sciences, University of Reading, Reading, RG6 6AL, UK

Address for correspondence and reprints:

Fang Liu, School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK. E-mail: f.liu@reading.ac.uk

Abstract:

Studies on how the form versus function aspect of tone and intonation is processed by autistic individuals have mainly focused on speakers of non-tonal languages (e.g., English), and have produced equivocal results. While the samples' heterogeneous cognitive abilities may be contributing factors, the phenotype of tone and intonation processing in autism may also vary with one's language background. Thirty-eight autistic and 32 non-autistic Mandarin speaking children completed tone and intonation perception tasks, each containing a function and form condition. Results suggested that the abilities to discriminate tone and intonation were not impaired at either the form or function level in some autistic children, and that these abilities were positively associated with one another in both autistic and non-autistic groups. Additionally, the more severe the ASD symptoms, the worse the form- and function-level of tone and intonation processing. While enhanced tone and intonation processing has been found in a subgroup of autistic children, it may not be a general characteristic of the autistic population even for those with long-term tone language experience. These findings reveal typical tone and intonation processing at both the form and function levels in cognitively competent autistic children and provide evidence for associated tone and intonation processing abilities across levels.

Keywords: Autism, Tone language, Lexical tone, Intonation, Pitch Perception

I. INTRODUCTION

Tone and intonation refer to the different types of pitch variation in spoken language. More specifically, tonal pitch encodes segments and morphemes, whilst intonational pitch gives further discourse meaning independent of the meaning of the words themselves (Gussenhoven, 2004). As such, the ability to process pitch, which corresponds to the perceptual attribute of the fundamental frequency of sound (Plack et al., 2006, 2014), is crucial for distinguishing between different tones and intonations in language (Xu, 2019). In addition, research has shown that the ability to process pitch varies across individuals, with moderating factors including intelligence (Acton & Schroeder, 2001; Deary et al., 1989; Helmbold et al., 2006; Raz et al., 1987; Spearman, 1904; Watson, 1991), age (Fancourt et al., 2013; Lamont, 1998), memory (Moore et al., 2007; Tillmann et al., 2016), music aptitude (Bidelman et al., 2013; Lynn et al., 1989; Schellenberg & Weiss, 2013; Wong et al., 2007), and tone language background (Bidelman et al., 2013; Pfordresher & Brown, 2009). Thus, comparing pitch processing, as well as tone and intonation processing across different groups and samples requires consideration of various background measures, which could confound the findings even in matched case-control studies (Pearce, 2016).

Apart from being embedded in speech, pitch is also carried in other sounds, such as musical melodies and chords (Krumhansl, 2004; Levrero et al., 2018). In contrast to function processing of tone and intonation in language (Xu, 2005, 2019), pitch processing of non-linguistic stimuli can be considered to be at a form level because of their lack of semantic meaning (Patel, 2008). For example, complex tone sequences matching the pitch and temporal patterns of speech stimuli have been used as musical analogues in comparative studies of pitch processing in music and language (Patel et al., 1998, 2005, 2008). It has been found that one of the reasons for the finer precision required to process pitch in music than in language may be related to the different roles it plays in each domain, namely the form-driven aesthetic role in music and the function-driven communicative role in language (Bidelman et al., 2013; F. Liu, Jiang, et al., 2013; Mantell & Pfordresher, 2013; Patel, 2008, 2011; Pfordresher & Brown, 2009).

A. Pitch processing at form- and function-level in typical development

In typical development (TD), form- and function-level processing of pitch are closely associated with one another (Asaridou & McQueen, 2013; Bidelman et al., 2011, 2013; Wong & Perrachione, 2007). Specifically, after years of sensory–motor training on music, musicians outperform non-musicians not only on form processing of pitch, such as detecting pitch changes in complex tones (Bidelman et al., 2013) and discriminating non-native lexical tones and their low-pass filtered and violin versions (Burnham, Brooker, et al., 2015), but also on function processing of pitch, such as discrimination and identification of native lexical tones (Ong et al., 2020), learning to identify non-native lexical tones (Wong & Perrachione, 2007), and detection of subtle pitch changes in intonational contours (Deguchi et al., 2012).

Like musicianship, tone language experience also increases pitch processing at both form and function levels (Bidelman et al., 2013; Burnham, Kasisopa, et al., 2015; Creel et al., 2018; Li et al., 2021). Compared to speakers of non-tonal languages (e.g., English), tone-language speakers (e.g., Mandarin) use pitch contours to distinguish lexical meaning at the syllable or word level on a daily basis (Yip, 2002). As a result of this lifelong tone experience, tone-language speakers show enhanced abilities to discriminate/recognize not only linguistic pitch contours and lexical tones (Stevens et al., 2013; Sun & Huang, 2012; Xu et al., 2006) but also fine-grained pitch and interval changes (Bidelman et al., 2013; Bradley, 2012; Giuliano et al., 2011; Pfordresher & Brown, 2009). Thus, the advantage in pitch processing among musicians and tone-language speakers suggests that a more fine-grained form processing of pitch (i.e., in musicians) can extrapolate to a more elaborate functional representation in language, and vice versa (Hirst, 2005; Patel, 2011; Wong et al., 2007).

B. Pitch processing at form- and function level in autism spectrum disorder

However, the findings of pitch processing in autism spectrum disorder (ASD) suggest a dissociation between form and function processing especially among earlier studies (Heaton, 2009; Heaton et al., 1998; O'Connor, 2012; Ouimet et al., 2012). ASD is a complex neurodevelopmental disorder characterized by impairments in social communication and interaction, as well as restricted and repetitive behaviours and interests (American Psychiatric Association, 2013). Previous research has shown that function-level processing of pitch seems to be selectively impaired in ASD, with the ability to process form-level pitch being intact (for reviews, see O'Connor, 2012; Ouimet et al., 2012). Specifically, in most studies investigating form processing, enhanced or at least unimpaired pitch sensitivity has been reported in autistic individuals, such as perception of pure tones (Bonnel et al., 2003, 2010; Heaton et al., 1998; O'Riordan & Passetti, 2006) and laryngographic sounds (Järvinen-Pasley et al., 2008; Peppé et al., 2007). By contrast, for function-level tasks, autistic individuals have been reported to have difficulties, such as distinguishing lexical stress contrasts (e.g., REcall versus reCALL) (Paul et al., 2005) and speakers' affective states (e.g., liking versus disliking) (Järvinen-Pasley et al., 2008; McCann et al., 2007; Peppé et al., 2007). It has been proposed that preferential processing of form-level pitch may be a consequence of reduced attention to functional information during development in ASD, resulting in atypical language development as seen in some autistic individuals (O'Connor, 2012).

Subsequent research nevertheless revealed that the ability to perceive pitch across form and function levels among autistic individuals is not as straightforward as it may seem. Instead, a growing body of evidence has shown that pitch processing abilities in autistic individuals may be modulated by cognitive factors, including IQ (Chowdhury et al., 2017; Mayer et al., 2016), verbal ability (Heaton, Hudry, et al., 2008), and short-term memory (Quintin et al., 2013; Sota et al., 2018). Thus, these factors should be considered while examining pitch processing in ASD. Indeed, Kargas et al. (2015) conducted a well-controlled study matching the ASD and TD groups on age, IQ, and musical training background, where autistic individuals showed poorer performance in discriminating pitch differences between pure tones compared to those without ASD. Yet, enhanced pitch discrimination (i.e., 100% accuracy) was observed in a subgroup of participants, including 9% in the ASD group (n = 2) and 14% in the TD group (n = 3) (Kargas et al., 2015). Hence, while enhanced form processing of pitch has been observed in ASD, it may only be evident among a subgroup of autistic individuals and may not necessarily be a general characteristic of the ASD population (Jones et al., 2009).

Similarly, the notion that autistic individuals are associated with impaired function-level processing of pitch has also been challenged by a recent study that carefully controlled for participants' characteristics between groups, including age, nonverbal IQ, musical training background, receptive verbal ability, and short-term memory (L. Wang et al., 2021). Specifically, this study utilized statement-question stimuli (e.g., "He just turned one./?") that differed primarily in the direction of the pitch contour of the final word (i.e., falling in statements and rising in questions) to examine intonation perception and production in 84 English-speaking children, adolescents and adults with and without ASD. Results indicated that intonation perception and production performance were comparable between the ASD and TD groups within each age cohort (L. Wang et al., 2021). Consistent with the finding of no response bias (e.g., judging the same items as different) from Mandarin speakers with ASD (Jiang et al., 2015), English speakers with ASD also showed no response bias (L. Wang et al., 2021). Thus, the comparable accuracy rates between the two groups were not confounded by response bias. Taken together, this study suggests that some autistic individuals may have genuinely unimpaired abilities to rely primarily on pitch cues to perceive and produce statement-question intonation at the function level (L. Wang et al., 2021).

In summary, much research has been done on the processing of form and function of pitch in ASD (Bonnel et al., 2003, 2010; Heaton et al., 1998; Heaton, 2005; Heaton, Williams, et al., 2008; Järvinen-Pasley et al., 2008; McCann et al., 2007; Paul et al., 2005; Peppé et al., 2007; L. Wang et al., 2021). However, it remains unresolved whether there is a dissociation between form- and function-level of pitch processing in autistic individuals especially after controlling for musical training experience and cognitive factors (e.g., IQ). It is also yet to be determined whether enhanced form-level processing of pitch accounts for or leads to impaired function processing of pitch in ASD (O'Connor, 2012). Answering these questions would not only help understand the pitch processing phenotypes of ASD, but also provide implications for studies of the development of language particularly language difficulties in autistic individuals (Lai et al., 2012; Sharda et al., 2015; Williams et al., 2021).

C. Pitch processing in ASD across tone and non-tonal languages

Previous research has primarily focused on speakers of non-tonal languages with ASD (O'Connor, 2012; Ouimet et al., 2012). It is known that the roles pitch play across languages are not always the same, with one difference being how pitch is used to convey meaning (Xu, 2019). Accordingly, the world's languages can be classified as tone and non-tonal languages (Yip, 2002). Specifically, in tone languages (e.g., Mandarin), pitch serves a lexically distinctive function. That is, the different pitch registers or contours recognized over strings of otherwise identical phonemes distinguish different words from one another (Klein et al., 2001). For example, with the same pitch contours (e.g., the rising tone, Tone 2, in Mandarin), words (e.g., "bai2", 白, "white") convey meaningful linguistic information to native listeners, whereas nonwords (e.g., "dai2") have no meaning in the lexicon (Zhou & Marslen-Wilson, 1994). Such tone languages are to be differentiated from non-tonal languages in which pitch variations are usually not contrastive at the syllable or word level. Namely, varying pitch contours does not change the lexical meaning of individual words, though it may alter the meaning of a sentence as a whole (Krishnan & Gandour, 2009). Across tone and non-tonal languages, the communicative function of speech intonation is signified through pitch contours in different ways (Xu, 2005). In English, the distinction between a statement and a yes/no question with neutral/final focus lies in the pitch contour of the final word: falling in statements and rising in yes/no questions (Eady & Cooper, 1986; F. Liu, Xu, et al., 2013). In Mandarin, pitch contours of lexical tones are encoded in parallel with focus and statement-question intonation. The distinction between a statement and a yes/no question with neutral/final focus lies in the pitch range of the final tones: compressed and lowered in statements and expanded and raised in yes/no questions (F. Liu & Xu, 2005). In both English and Mandarin, pitch range of the focused word is expanded in statements as well as in questions. In Mandarin, post-focus pitch range is compressed and lowered in both statements and questions, although the latter is smaller in magnitude (F. Liu & Xu, 2005). In English, post-focus pitch range is compressed and lowered in statements but compressed and raised in yes/no questions (Eady & Cooper, 1986; F. Liu, Xu, et al., 2013). Given the typological differences in pitch processing between tone and non-tonal languages, focusing mainly on non-tonal languages may lead to an incomplete understanding of pitch processing in ASD and how pitch processing ability affects language development in ASD generally.

To date, only a handful of studies have investigated pitch processing in speakers of tone languages with ASD from behavioural and electrophysiological perspectives (Chen & Peng, 2021; Cheng et al., 2017; Jiang et al., 2015; Lau et al., 2020; Rong et al., 2022; X. Wang et al., 2017; Yu et al., 2015; J. Zhang et al., 2019). Generally speaking, behavioural studies fail to observe the dissociation between form and function processing of pitch in tone-language speakers with ASD for discrimination of pitch differences between pairs of real words, nonwords, and non-speech stimuli (Cheng et al., 2017) and categorization of pairs of real words, nonwords, iterated rippled noise, and pure tones based on pitch contours (Chen & Peng, 2021). Regardless of stimulus type, Mandarin/Cantonese speakers with ASD did not outperform the TD group on pitch perception (Chen & Peng, 2021; Cheng et al., 2017).

However, evidence from electrophysiological studies suggests that Mandarin/Cantonese speakers with ASD manifest impaired ability to process functional pitch (e.g., in lexical tones), whereas the ability to process form pitch can be either enhanced or unimpaired, as indicated by the presentation of ERP responses including mismatch negativity and P3a (X. Wang et al., 2017; Yu et al., 2015; J. Zhang et al., 2019). In addition, Lau et al. (2020) examined early neural sensory encoding of pitch in lexical tones by measuring frequency-following response (FFR) and found a linguistic pitch encoding impairment in Cantonese speakers with ASD.

The inconsistencies between the above-mentioned electrophysiological and behavioural studies may be explained by the different processing mode (i.e., passive versus active) required in these different types of studies. Indeed, Whitehouse and Bishop (2008) have tested this possibility by setting up two processing modes: a passive mode where participants were told to ignore the sounds, and an active mode where participants responded by clicking on the mouse when hearing nonstandard sounds. They found that autistic individuals showed diminished ERP responses (indicated by P1-N2-P3-N4-P3a, suggesting poor sound encoding) to speech stimuli but not to nonspeech stimuli in the passive mode, whereas the speech encoding deficits disappeared in the active mode where participants were asked to pay attention to the sound stream. Thus, it appears that while autistic individuals can process linguistic information in speech when in an active mode as suggested by previous behavioural studies, they do not do so spontaneously when in a passive mode as shown in some of the electrophysiological studies. However, with only a few behavioural studies investigating pitch processing in tone-language speakers with ASD, one cannot draw well-founded conclusions about how tone-language experience affects pitch processing in ASD.

A. Present study

To obtain a more comprehensive understanding of pitch processing in ASD, the current study investigated tone and intonation processing in Mandarin-speaking autistic individuals while controlling for various possible contributing factors including nonverbal IQ, verbal ability, short-term memory, and musical training. Particularly, we examined whether tone-language experience can compensate for the possible deficit in function processing of pitch that has been suggested for some non-tonal-language speakers with ASD (Järvinen-Pasley et al., 2008; McCann et al., 2007; Paul et al., 2005; Peppé et al., 2007) using tone and intonation discrimination tasks at both function and form levels. Our main research question was (a) Would Mandarin-speaking autistic individuals differ from non-autistic individuals in terms of pitch processing across the form and function levels in tone and intonation tasks? Our secondary research questions included (b) Would the performance of the two groups be confounded by response bias? (c) Would a subgroup of autistic individuals show enhanced form and function processing of pitch in tone and intonation tasks? (d) Would the associations between musical training experience/cognitive ability and pitch processing differ between the two groups at the form and function levels? (e) Would the severity of ASD be associated with pitch processing across form and function levels? Based on existing limited findings from tone-language speakers on pitch perception between autistic and non-autistic individuals (Chen & Peng, 2021; Cheng et al., 2017; Jiang et al., 2015), we hypothesized that: (a) when various possible contributing factors (e.g., nonverbal IQ, verbal ability, short-term memory, and musical training) are controlled for between groups, Mandarin-speaking autistic individuals would not show a deficit in function-level pitch processing in tone and intonation tasks, whereas they might show an enhancement when processing form-level pitch compared with non-autistic individuals; (b) The performance between the two groups would not be confounded by response bias; (c) A subgroup of autistic individuals would show enhanced form and function processing of pitch in tone and intonation tasks; (d) Across form- and function-level pitch processing, the associations between musical training experience, cognitive ability, and pitch processing would be similar between the two groups; (e) Severity of ASD would be associated with pitch processing at both the form and function levels.

II. METHODS

A. Participants

Forty-five autistic children (aged between 7 and 16) and thirty-three age-matched non-autistic children participated in the study. All were native speakers of Mandarin and were recruited from mainstream schools and special educational facilities in Nanchang and Nanjing, China. The children with ASD all had a clinical diagnosis of ASD, which was further supported by the Autism Diagnostic Observation Schedule – Second Edition (ADOS-2) (Lord et al., 2012) conducted by the first author (with clinical and research reliability for administration and scoring). All participants with ASD were administrated using the ADOS-2 Module 3 according to their developmental and language levels. Total scores on the ADOS-2 consisting of Social Affect (SA) and Restricted and Repetitive Behavior (RRB) were converted to a comparative score (CS) of 1-10, with 10 representing the highest severity of autism-related symptoms (Duda et al., 2014; Gotham et al., 2009). All participants had normal hearing in both ears, with pure-tone air conduction thresholds of 25 dB HL or better at frequencies of 0.5, 1, 2, and 4 kHz, as assessed using an Amplivox manual audiometer (Model 116). Participants' musical training background was collected using a questionnaire, and their years of formal musical training were summed across all instruments including voice (L. Wang et al., 2021). Participants completed a nonverbal IQ test using the Raven's Standard Progressive Matrices Test (RSPM) (Raven et al., 1998) and a receptive vocabulary test using the Peabody Picture Vocabulary Test-Revised (PPVT-R) (Dunn & Dunn, 1981). The Chinese version of the forward digit span task was used to assess verbal short-term memory (Wechsler, 2003). The standardized scores for RSPM were calculated based on the means and standard deviations obtained from a Chinese normative study (H. Zhang, 1989). Given that the Chinese norms for PPVT-R only included ages from 3.5 to 9 (Sang & Miao, 1990), standardized scores were calculated based on American norms (Dunn & Dunn, 1981). Correlation analysis revealed a significant positive relationship between the standardized scores obtained based on the Chinese norms and those based on the American norms ($r = 0.95$) for participants at or below 9 years old, thus confirming the validity of this approach.

To avoid the results being confounded by impaired intelligence and receptive verbal ability, we only included participants with IQ and verbal ability in the typical range (> 70), resulting in 7 ASD and 1 TD participants excluded from the current study. Finally, a total of 38 autistic children (5 females and 33 males) and 32 non-autistic children (5 females and 27 males) were included in the data analysis. Two-sample *t*-tests were conducted to test whether the two groups were matched on the background measures. TABLE I shows the characteristics of the participants and the results of the two-sample *t*-tests. The two groups were largely matched on the background measures, with the exception that the ASD group showed lower scores of PPVT-R than the TD group. To control for the possible contribution of receptive verbal ability to the current results, scores on PPVT-R were entered as covariates in the statistical analysis.

A post hoc power analysis was conducted using G*Power (Faul et al., 2009). To detect the interaction of Group (ASD vs. TD) by Stimulus type (form vs. function) and a covariance of receptive verbal ability in the present design, a total of 70 participants reached a power of 0.91 with a large effect size ($f = 0.40$) at an alpha of 0.05. This suggests that our sample size was large enough to detect statistically significant effects examined here. The study was approved by the University of Reading Research Ethics Committee. Written informed consent/assent was obtained from the participants and their parents prior to the experiment.

TABLE I. Characteristics of the ASD (n = 38) and TD groups (n = 32).

Variables	ASD	TD	t	p	Cohen's *d*
Age					
Mean (SD)	10.37 (2.54)	11.47 (2.75)	1.73	0.09	0.25
Musical training					
Mean (SD)	0.88 (1.31)	0.47 (1.08)	1.45	0.15	0.24
RSPM					
Mean (SD)	110.4 (15.05)	112.94 (9.98)	0.85	0.40	0.24
PPVT-R					
Mean (SD)	121.90 (27.12)	141.41(12.83)	3.94	< .001	0.26
Digit span					
Mean (SD)	8.32 (1.07)	8.13 (1.10)	0.73	0.47	0.24

Note: Musical training: years of musical training; RSPM: standard score of Raven's Standard Progressive Matrices Test; PPVT-R: standard score of Peabody Picture Vocabulary Test-Revised; Digit span: raw score of verbal short-term memory.

B. Tasks

The experiments consisted of one disyllabic lexical tone discrimination task and one intonation discrimination task. Stimuli from all tasks were recorded or generated using Praat (Boersma & Weenink, 2001), with 44.1 kHz sampling rate and 16-bit amplitude resolution.

1. Disyllabic lexical tone discrimination task

The stimuli in the lexical tone discrimination task consisted of seventy-two disyllabic pairs (see Appendix TABLE III. for the whole list), among which thirty-six were real words in Mandarin (e.g., 仙人-闲人, xian1ren2-xian2ren2, 'celestial being'-'idler') and thirty-six were nonwords (e.g., 相牌-相拍, xiang1pai2-xiang1pai1). In contrast to the real words which conveyed lexically meaningful information, the nonwords were not part of the Mandarin lexicon and thus were only non-functional representations of form processing of the whole compounds (Zhou & Marslen-Wilson, 1994). The reasons that we chose disyllabic rather than monosyllabic words/nonwords as stimuli were that a) most words are disyllabic in Modern Chinese; and that b) it was difficult to find monosyllabic nonword pairs with the same segments (Duanmu, 2007). In Mandarin Chinese, disyllabic words and nonwords are processed at a whole-word level, rather than at a syllable or morpheme level (Zhou & Marslen-Wilson, 1994). Therefore, although each of the two syllables that make up a disyllabic non-word has its own meaning, it is the lack of meaning of the whole compound that is being processed and recognised as a nonword in native listeners' mental lexicon. Consequently, the distinction between the processing of words and nonwords can be seen as a function versus form contrast.

Half the pairs differed in the first syllable and the other half differed in the second syllable. The frequencies of usage of the words in each pair were closely matched (two-sample *t*-test: $t(70) = -0.34$, $p = 0.73$), using the lexicon of common words in contemporary Chinese (The National Language Working Committee, 2008). The two (non-)words in each pair were manipulated through a cross-splicing procedure using a custom-written Praat script so that they shared the same segments but differed in tonal composition (Liu et al., 2012). Each of the 72 stimulus pairs appeared in both 'same' and 'different' conditions, leading to 144 stimulus pairs (72 'same' pairs and 72 'different' pairs) in total. The stimuli were randomized and presented to each participant in a different order, with 750 ms interstimulus interval and 1500 ms intertrial interval after receiving a response. The duration of the stimuli was normalised to 450 ms, with intensity normalised at 65 dB. The stimuli were presented with PsychoPy (Peirce et al., 2019) through the built-in speakers of a Lenovo ThinkPad laptop. Eight practice trials (with different stimuli than the experimental trials) were given to familiarize the participants with the experimental procedure and materials, with the sound volume adjusted to a comfortable listening level for each participant. Participants were asked to judge whether the tones of the stimuli in each pair sounded the same or different.

2. *Intonation discrimination task*

The intonation discrimination task consisted of two subtests assessing discrimination of statements and questions at a function level using natural speech and at a form level using gliding tone analogues of natural speech (F. Liu et al., 2012). The stimuli comprised both natural speech utterances and their gliding tone analogues, ranging from 3 to 7 syllables and consisting of only High/Falling tones (see Appendix TABLE IV. for a full list). In the natural speech condition, 20 statement-question pairs shared the same word sequence but differed in intonation. They were naturally spoken with either an initial or a final focus. To create gliding tone analogues of natural speech, 20 complex tone pairs were created using Praat to match the pitch contours and durations of the natural speech utterances word by word. Specifically, the complex tones consisted of the fundamental frequency (F0) plus seven odd harmonics of the syllable(s) in the stimuli, leading to clarinet-like sound quality. To achieve roughly equal loudness, the amplitudes of all stimuli were normalized by increasing the peak value to the maximum utilizing Praat. It should be noted that these statement and question pairs differed in multiple prosodic cues including pitch, duration, and intensity, whereas their musical analogues were created using complex tones replacing the voiced speech segments (F. Liu et al., 2012). The reasons that we did not manipulate or equalise the duration and intensity cues across the statement-question pairs were that a) in Mandarin Chinese, lexical tone, focus, and statement-question intonation are realised in parallel through a single pitch channel together with variations in other acoustic cues such as duration and intensity (F. Liu & Xu, 2005; Yuan, 2006); and that b) this multiplexing of several linguistic functions makes the identification and discrimination of intonation challenging (F. Liu et al., 2012; F. Liu & Xu, 2005; M. Liu et al., 2016; Yuan, 2011). To maintain naturalness of the stimuli and to avoid potential floor performance of participants, we decided to examine form and function processing of prosody in the intonation task, rather than focusing on pitch alone (F. Liu et al., 2012).

There were 40 stimulus pairs (20 same and 20 different pairs) in each subtest, with 750 ms interstimulus interval and 1500 ms intertrial interval after receiving a response. The stimuli were presented with PsychoPy (Peirce et al., 2019) through the built-in speakers of a Lenovo ThinkPad laptop. Four practice trials (with different stimuli than the experimental trials) were given before each task to familiarize the participants with the experimental procedure and materials, and to adjust the sound volume to a comfortable listening level for each participant, based on participant feedback. Participants were required to achieve 100% correct on the practice trials (with feedback) before proceeding to the testing sessions. During testing, participants were required to judge whether the two sequences sounded the same or different across pitch, duration, and intensity cues. Given that inattention may impact the performance of the task, especially in children (L. Wang et al., 2021), participants were required to make their responses orally for the experimenters to enter into the laptop, in order to maintain their attention. The natural speech and gliding tone subtests were presented in counterbalanced order across participants.

C. Data analysis

D-prime (d') from signal detection theory was used to measure participants' performance, as the standardized difference between hits (i.e., correct responses to "different" trials) and false alarms (i.e., wrong responses to "same" trials) (Green & Swets, 1966; Macmillan & Creelman, 2005). Larger values of d' indicate higher sensitivity and hence better discrimination. We calculated d' using the psycho package in Rstudio (Makowski, 2018; RStudio Team, 2020), where extreme values (e.g., hits = 100% or 0%) were corrected following the log-linear rule by increasing the frequency of each category (e.g., hits and false alarms) by 0.5 (Hautus, 1995).

Both classical frequentist and Bayesian analyses were run using JASP software (JASP Team, 2020). Bayes Factor (BF) is the ratio of the likelihood of one hypothesis (e.g., an alternative hypothesis) over the likelihood of another (e.g., a null hypothesis). Unlike frequentist statistics, BF can quantify the strength of evidence in favor of one of the two hypotheses (Dienes, 2014). Thus, BF is particularly useful in evaluating the strength of the null hypothesis attained, compared with only reporting the probability using cut-off values (e.g., p = 0.05) of the data given the null hypothesis. We therefore conducted both analyses in the present study. Specifically, BF_{10} indicates the Bayes factor in favor of the alternative hypothesis over the null hypothesis, whereas BF_{01} indicates the Bayes factor in favor of the null over the alternative hypothesis. The relationship between the two is $BF_{10} = 1/BF_{01}$. For the interpretation of BF as support for hypotheses, Raftery (1995, p.139) suggested that a BF value above 1 and less than 3 is "weak" evidence, and a BF range between 3-20 is "positive" evidence.

III. RESULTS

A. Disyllabic lexical tone discrimination task

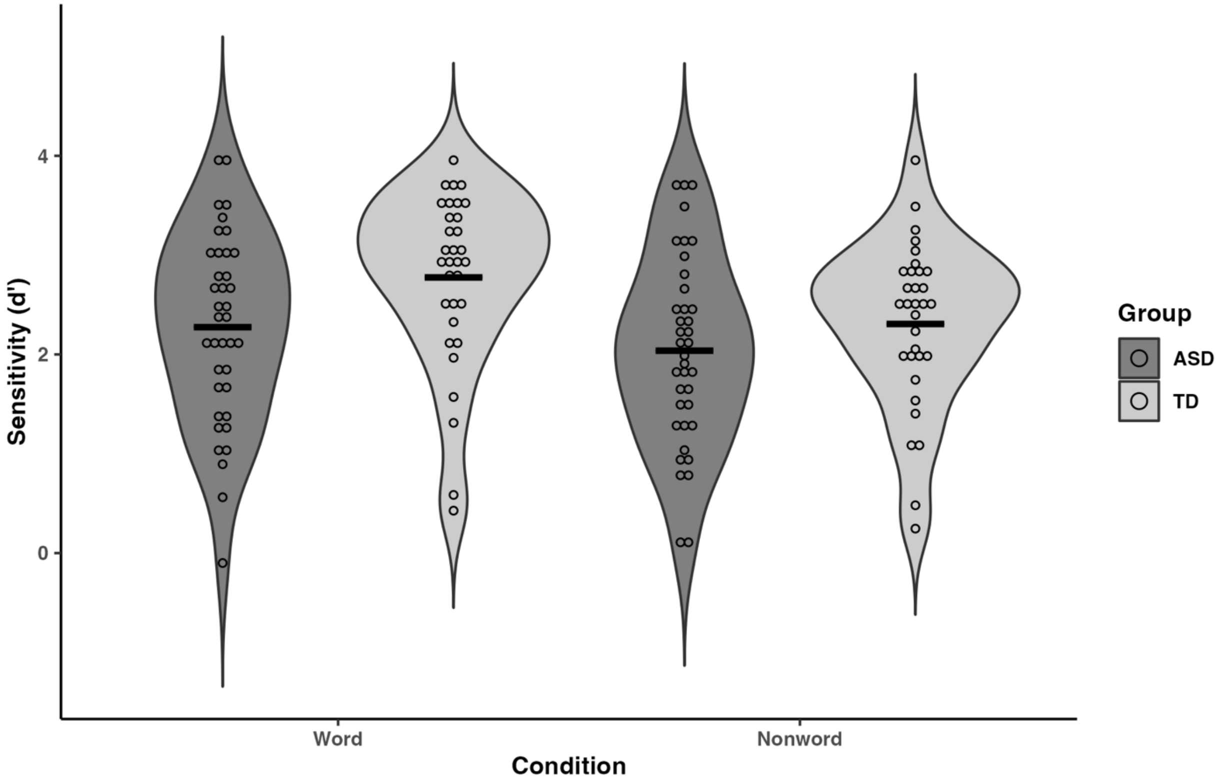

FIG. 1. Sensitivity index (d') observed for each group in discrimination of words (function) and nonwords (form) (Black lines represent mean values).

FIG. 1. shows the means and distributions of the performance of the two groups on the lexical tone discrimination task. A repeated measure ANCOVA with Bayesian and frequentist analysis was conducted to examine the effects of Condition (Words vs. Nonwords) and Group (ASD vs. TD) on lexical tone discrimination, after controlling for receptive vocabulary scores (i.e., PPVT-R). Results revealed a significant main effect of Condition, $F(1, 67) = 25.06$, $p < 0.001$, with both groups showing better performance on the word condition (ASD: M(SD) = 2.28 (0.95); TD: M(SD) = 2.78 (0.83)) than the nonword condition (ASD: M(SD) = 2.04 (0.95); TD: M(SD) = 2.31 (0.84)). The main effect of Group ($F(1,67) = 0.05$, $p = 0.83$) and the interaction between Group and Condition ($F(1, 67) = 1.09$, $p = 0.30$) did not reach significance. The main effect of PPVT-R was significant, $F(1,67) = 16.19$, $p < 0.001$.

Bayesian model comparisons led to a similar conclusion. There was positive evidence in favor of the effects of Condition ($BF_{10} = 2069.94$) and PPVT-R ($BF_{10} = 376.30$). There was week evidence against an effect of Group ($BF_{01} = 1.85$) and a Group by Condition interaction ($BF_{01} = 1.08$).

B. Intonation discrimination task

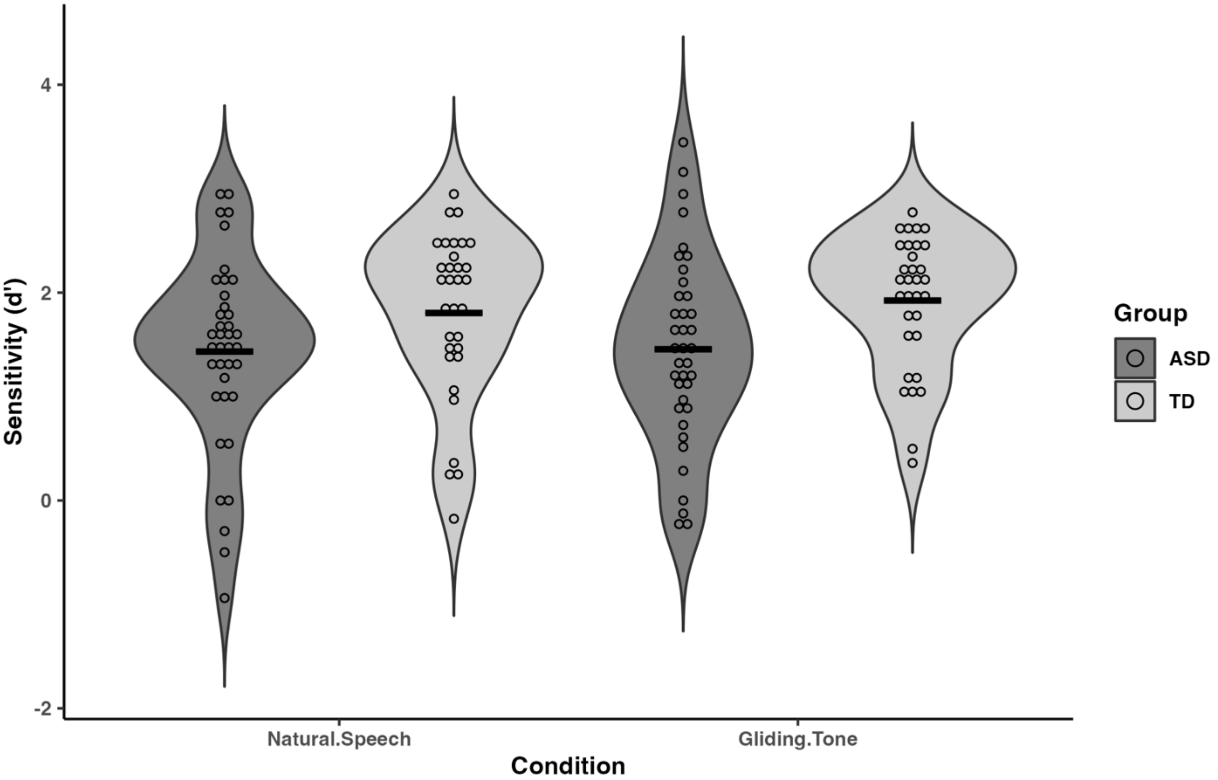

FIG. 2. Sensitivity index (d') observed for each group in discrimination of natural speech (function) and gliding tones (form) (Black lines represent mean values).

FIG. 2. shows the means and distributions of the two groups' performance on the intonation discrimination task. Similarly, a repeated measure ANCOVA with Bayesian and frequentist analysis was conducted to examine the effects of Condition (Speech vs. Gliding tone analogue) and Group (ASD vs. TD) on intonation discrimination, after controlling for PPVT-R. The analysis revealed that the main effect of Group (F(1, 67) = 0.53, $p = 0.47$) was not significant (Speech: ASD: M(SD) = 1.43 (0.92); TD: M(SD) = 1.80 (0.72); Gliding tone analogue: ASD: M(SD) = 1.46 (0.91); TD: M(SD) = 1.92 (0.67)), nor was the main effect of Condition (F(1, 67) = 0.91, $p = 0.34$) or the Group × Condition interaction (F(1, 67) = 0.32, $p = 0.57$). There was a significant effect of PPVT-R, F(1, 67) = 13.14, $p < 0.001$.

Again, Bayesian model comparisons produced similar results. A model with an effect of PPVT-R was more likely than the null model supported by positive evidence ($BF_{10} = 158.05$). There was weak evidence against an effect of Group ($BF_{01} = 2.58$) and positive evidence against an effect of Condition ($BF_{01} = 5.05$) and Group × Condition interaction ($BF_{01} = 11.36$).

C. Correlation between performance on form- and function-level pitch processing

Given that previous studies suggested a dissociation between form- and function-level auditory processing in ASD (O'Connor, 2012; Ouimet et al., 2012), we investigated the relationship between form and function processing in the present tone and intonation tasks. To calculate the scores on form and function processing respectively, we averaged the corresponding d' scores across conditions. Specifically, the score on form processing was the average d' score in the nonword condition of the tone discrimination task and the gliding tone analogue condition of the intonation discrimination task, and the score on function processing was the average d' score in the word condition of the tone discrimination task and the natural speech condition of the intonation discrimination task. Kendall's tau correlation with a positive hypothesis of the association (one-tailed) indicated that the scores on form and function processing were positively correlated in both groups (ASD: tau = 0.70, $p < .001$, $BF_{+0} = 32360000$; TD: tau = 0.57, $p < 0.001$, $BF_{+0} = 11392.96$).

D. Response bias

To investigate whether autistic individuals showed response biases between the same versus different pairs, i.e., tending to judge the same pairs as different or vice versa, we summed the number of trials of hits and false alarms across tasks (i.e., lexical tone and intonation discrimination tasks) and calculated not only d' but also beta. In particular, beta is an index of response bias and reflects participants' bias toward "same" or "different" with the unbiased response pattern having a value approaching 1 (Makowski, 2018). Values of beta less than 1 indicate a bias toward judging pairs as "different", whereas values greater than 1 indicate a bias toward responding "same". By comparing with beta = 1 (i.e., unbiased response pattern), one-sample *t* tests suggested that the values of beta from both groups were significantly higher than 1 (ASD: M(SD) = 2.23 (1.95), t(37) = 3.88, $p < 0.001$, BF_{10} = 67.17; TD: M(SD) = 4.95 (5.18), t(31) = 4.31, $p < 0.001$, BF_{10} = 174.74), indicating a bias toward judging the pairs to be the same for both groups. However, two-sample *t* tests revealed that this trend was significantly more pronounced in the TD group than the ASD group (t(68) = 3.00, $p =$ 0.004, BF_{10} = 10.09).

E. Subgroup analysis

Given previous studies suggesting that enhanced pitch perception may only exist in a small subgroup with ASD (Heaton, Williams, et al., 2008; Jones et al., 2009; Kargas et al., 2015), we further explored participants' performance scores on form-level and function-level tasks in order to determine whether there was a subgroup of exceptional pitch perception in ASD in the present experiment. Exceptional pitch perception performance was defined as 1 and 1.5 SD above the mean value of the TD group, since only one ASD participant performed above two SDs (i.e., Sub1 in TABLE II., who had 4 years of musical training and relatively high nonverbal IQ and receptive verbal ability, as well as verbal short-term memory) (Jones et al., 2009). For form processing, two participants with ASD (5% of the ASD group) performed above 1.5 SDs of the TD mean (M (SD)= 2.14 (0.68)) and five participants above 1 SD (13% of the ASD group), compared to one (3% of the TD group) and three participants in the TD group (9% of the TD group), respectively. While no TD participants showed exceptional pitch processing at the function level with 1.5 SDs above the TD mean (M (SD)= 2.29 (0.70)), two participants with ASD did (5% of the ASD group). In addition, four participants with ASD (11% of the ASD group), in comparison to six TD participants (19% of the TD group), performed 1 SD above the mean value of the TD group (see TABLE II. for details). If we only focused on the characteristics of those who performed 1.5 SDs above the mean value of the TD group across form and function levels (i.e., Sub1, 2, 3 and 8), they all had musical training for at least 2 years, had relatively high nonverbal IQ and receptive verbal scores, as well as enhanced short-term memory.

TABLE II. The characteristics of participants who showed exceptional pitch sensitivity at form and function levels.

ID	Group	Above 1.5 SDs		Above 1 SD		Musical training	RSPM	PPVT-R	Digit span
		Form	Function	Form	Function				
Sub1	ASD	■	■	■	■	4	120.0	150	9
Sub2	ASD	■		■	■	4	141.7	96	9
Sub3	ASD		■	■	■	2	131.4	159	9
Sub4	ASD			■	■	0	121.7	87	9
Sub5	ASD			■		0	114.6	142	9
Sub6	TD			■	■	0	118.9	160	8
Sub7	TD			■	■	0	118.9	148	9
Sub8	TD	■		■		4	130.9	159	9
Sub9	TD				■	0	121.0	144	8
Sub10	TD				■	0	120.2	150	9
Sub11	TD				■	0	124.0	159	9
Sub12	TD				■	1	121.6	159	9

Note: Musical training: years of musical training; RSPM: standard score of Raven's Standard Progressive Matrices Test; PPVT-R: standard score of Peabody Picture Vocabulary Test-Revised; Digit span: raw score of verbal short-term memory.

F. The relationship between ASD severity and pitch processing performance

To explore the relationship between ASD severity levels and pitch processing performance at the form and function level, we conducted correlation analysis between the sub-scores of ADOS-2, namely, Social Affect (SA) and Restricted and Repetitive Behavior (RRB), comparison scores (CS), and d' sensitivity for each level. Given that previous studies reported both positive and negative correlations between ASD severity and pitch processing abilities (Diehl et al., 2009; Mayer et al., 2016; Nadig & Shaw, 2012), Kendall's tau correlations without a specific hypothesis of the direction of the association (two-tailed) were conducted. The results showed that the performance on form and function processing was largely negatively correlated with scores on SA, RRB, and CS: the greater the pitch sensitivity, the lower the SA, RRB, and CS scores (FIG. 3.).

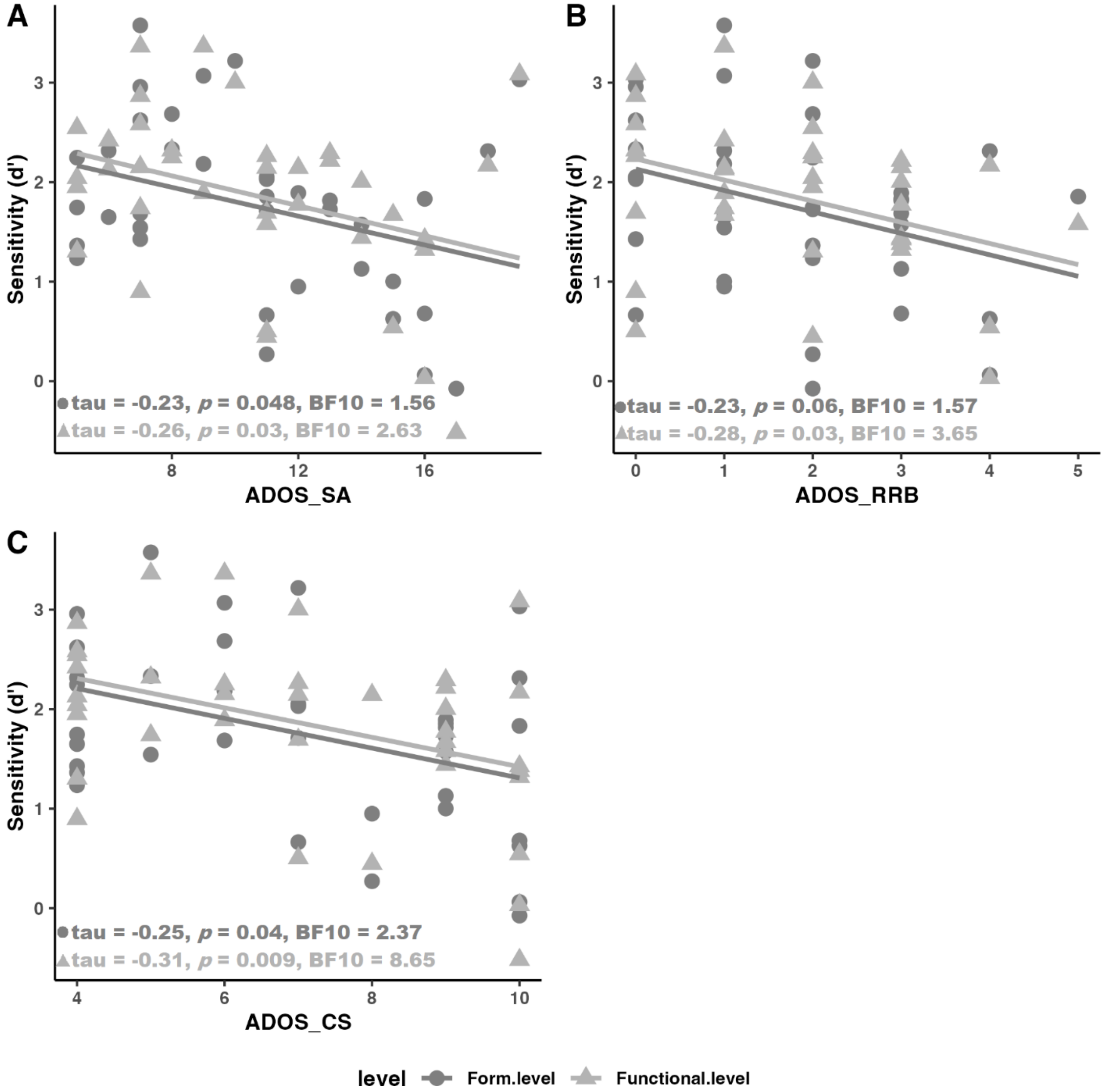

FIG. 3. Scatter plots of ASD severity levels against pitch processing performance.

G. The relationship between cognitive factors and pitch processing performance

To further explore how chronological age, musical training background, nonverbal IQ, receptive verbal ability, and short-term memory influence form and function processing for both groups, we examined the correlations between these factors and performance on form/function processing by group. Results indicate significant correlations between these cognitive factors, except for the factor of chronological age, and both form and function processing of pitch/prosody in the ASD group, whereas there was only a significant correlation between receptive verbal ability and form processing of pitch in the TD group (See FIG. 4.).

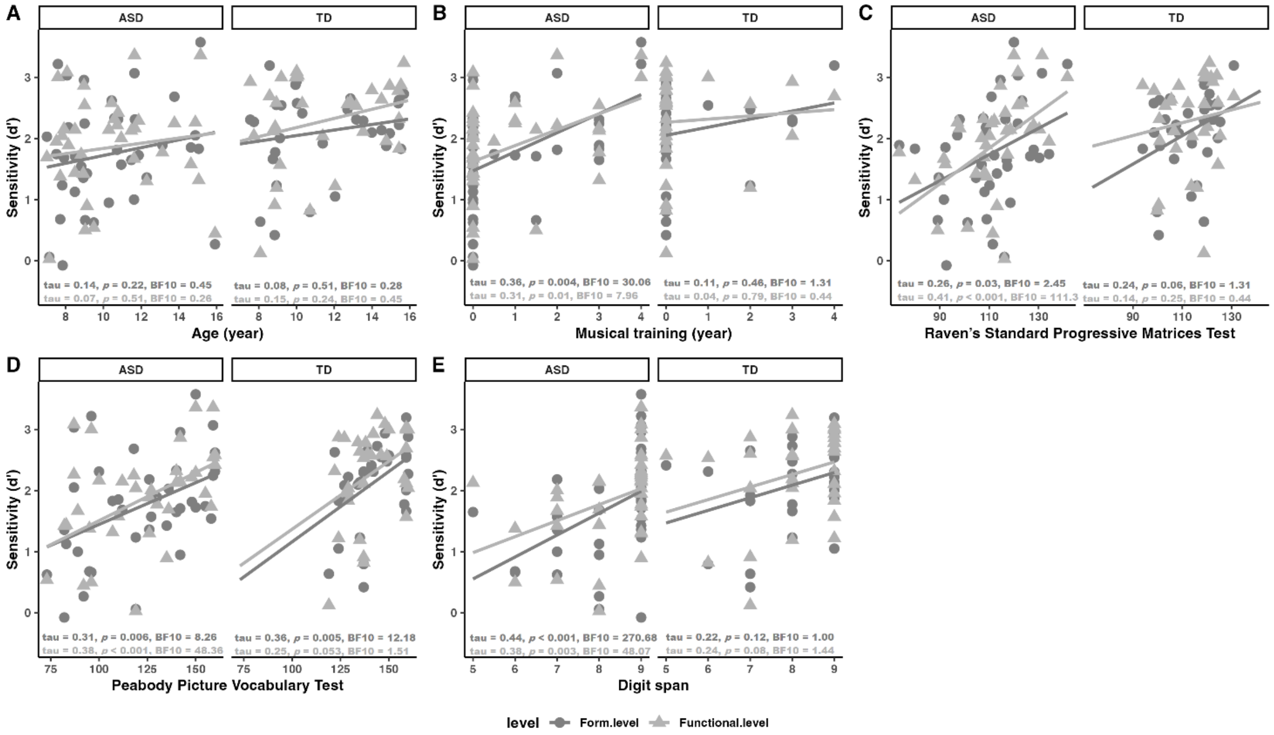

FIG. 4. Scatter plots of cognitive factors against pitch processing performance at form and function levels by group.

IV. DISCUSSION

Using lexical tone and intonation perception tasks, the present study examined whether there was a dissociation between form and function processing of pitch in Mandarin-speaking children with and without ASD, while controlling for a variety of possible contributing factors (i.e., nonverbal IQ, receptive verbal ability, short-term memory, and musical training). The main results showed that the abilities to discriminate lexical tone and statement-question intonation were typical at both form and function levels in some autistic children relative to non-autistic children, and that the abilities to process form- and function-level pitch were positively associated with one another in both groups. In addition, while enhanced pitch processing has been found in a subgroup of autistic children, it may not necessarily be a general characteristic of the ASD population, even for Mandarin-speaking autistic individuals who have lifelong tone-language experience. Furthermore, the more severe the ASD symptoms, the worse the form- and function-level of pitch processing among autistic participants. Finally, musical training experience and cognitive abilities (i.e., nonverbal IQ, receptive verbal ability, short-term memory) were significantly correlated with both form and function processing of pitch for the ASD group. Given that no group-specific response bias was observed in either task, the present findings indicate that cognitively competent Mandarin-speaking autistic children may have genuinely unimpaired pitch processing abilities in tone and statement-question intonation at both the form and function levels.

The current finding of unimpaired processing of function-level pitch (i.e., words and natural speech of statement/question intonation) in the Mandarin-speaking ASD group is consistent with our hypothesis, although we did not find an enhancement in processing form-level pitch (i.e., nonwords and gliding tone analogues of intonation), relative to the TD group. Specifically, regarding lexical tone processing, we found no group difference in discrimination accuracy across word (function) and nonword (form) conditions among Mandarin speakers with and without ASD. This finding is consistent with previous behavioural results on tone language speakers suggesting that Mandarin/Cantonese speakers with ASD did not outperform the TD group on pitch perception of lexical tones, regardless of stimulus type (Chen & Peng, 2021; Cheng et al., 2017).

In the intonation discrimination task, Mandarin-speaking participants with ASD performed comparably to those without ASD across natural speech (function) and gliding tone analogue (form) conditions. These findings are in line with previous studies suggesting typical perception of statement-question intonation in ASD (Chevallier et al., 2009; Filipe et al., 2014; Järvinen-Pasley et al., 2008; Paul et al., 2005; L. Wang et al., 2021). However, they are inconsistent with other findings indicating impaired statement-question intonation perception among English speakers with ASD (McCann et al., 2007; Peppé et al., 2007) and Mandarin speakers with ASD (Jiang et al., 2015). It is worth noting that McCann et al. (2007) and Peppé et al. (2007) reported the same set of results from the same sample of participants (31 ASD versus 72 TD) in two different publications. In addition, they assessed statement-question intonation discrimination using a short-item discrimination task that included not only the laryngographic sounds of statement-question pairs but also liking-disliking pairs from an affective subtask within the same test battery—PEPS-C (Profiling Elements of Prosodic Systems - Children) (Peppé & McCann, 2003). Thus, the impaired intonation discrimination performance reported in these two studies for English speakers with ASD is likely confounded by participants' ability to discriminate affective prosody. In addition, while the stimuli used in the present study and Jiang et al. (2015) were both in Mandarin, there were some differences in the design of the experiments and stimuli that may explain the contradictory findings between these two studies. Specifically, Jiang et al. (2015) used disyllabic statement-question pairs that were manipulated to only differ in pitch of the second syllable. In the present study, the intonation stimuli consisted of naturally spoken statements and questions of 3-7 syllables that differed not only in pitch but also in duration and intensity across the entire utterances. Thus, it is likely that participants used multiple acoustic cues in discriminating the sentences and their gliding tone analogues in the current study, as in Liu et al. (2012). To test this possibility, future studies should examine whether stimuli with different lengths and acoustic manipulations would lead to different performance on intonation discrimination in the same sample of participants.

Our results and those from some earlier study (L. Wang et al., 2021) provide converging evidence that Mandarin- and English-speaking autistic participants may have unimpaired form and function processing of pitch across languages. However, one may argue that these results may be due to sampling error or variability. Indeed, ASD is a condition with heterogeneity across different symptoms, including language ability, IQ, and memory (Chiang et al., 2014; Eigsti et al., 2011; Mottron & Bzdok, 2020; Nowell et al., 2015; Tager-Flusberg & Joseph, 2003). Our correlation results on Mandarin speakers suggested that compared to the TD group, the ASD group's performance on form and function processing of pitch/prosody was more likely to be modulated by these cognitive factors (see FIG. 4.). In the present study, participants with ASD had typical ranges of receptive verbal ability, nonverbal IQ, and short-term memory, which might be the reason why we did not observe impaired function processing of pitch or enhanced form processing of pitch (Chowdhury et al., 2017; Heaton, Hudry, et al., 2008; Jones et al., 2009). While our current results may not be representative of a less cognitively capable ASD group, some earlier studies did report cases of musical savants with exceptional pitch sensitivity (e.g., absolute pitch) in the presence of relatively low cognitive abilities (e.g., IQ) (Mottron et al., 1999; Young & Nettelbeck, 1995). Nevertheless, the current null results are unlikely due to issues related to sample size or sampling error, as our sample sizes were large enough to detect an effect based on our post hoc power analysis. In addition, no group-specific response bias was observed in either tone or intonation discrimination, and the abilities to process form and function of pitch were positively correlated with each other in both groups. Therefore, the current results may reflect a genuinely unimpaired ability of the cognitively competent autistic participants to discriminate tone and intonation at both form and function levels (L. Wang et al., 2021).

Our findings from Mandarin speakers with ASD support the notion reported by previous English studies that enhanced pitch/prosodic processing might only exist in a small subgroup of autistic individuals (Heaton, Williams, et al., 2008; Jones et al., 2009; Kargas et al., 2015). In addition, pitch processing abilities in participants with ASD across form and function levels were largely subject to ASD severity levels, i.e., the more severe the ASD symptoms, the worse the form- and function-level of pitch processing. However, these findings were weakly supported by Bayesians and should be interpreted with caution. Indeed, conflicting results have been reported in the literature, with some suggesting that pitch sensitivity and ASD severity were correlated and others suggesting otherwise (Diehl et al., 2009; Mayer et al., 2016; Nadig & Shaw, 2012). Nevertheless, Mayer et al. (2016) proposed that the scores from the ADOS might not be suitable to be used in empirical analysis. While Mayer et al. (2016) observed correlated ASD severity levels and pitch processing, the finding lacked support from the correlation between scores on pitch discrimination and on the Communication Checklist—a self-reported questionnaire used to reflect communication difficulties (Bishop et al., 2009). Therefore, with limited but mixed findings in the literature (Diehl et al., 2009; Mayer et al., 2016; Nadig & Shaw, 2012), as well as the questioning of the appropriateness of using ADOS scores in empirical analysis, the relationship between pitch processing and ASD severity level warrants further studies.

Moreover, it is worth mentioning that both groups in the present study performed better on the word condition than on the nonword condition in the lexical tone discrimination task, with receptive verbal ability contributing positively to the performance. These findings suggest that participants captured the lexical meaning of the words while discriminating their differences in lexical tone, with the presence of semantic meaning facilitating both groups' performance. However, a previous study indicated that over-focusing on semantic meaning under speech conditions hampered the performance of English-speaking TD participants on pitch discrimination, compared to their performance under music conditions, while ASD participants performed equally well across speech and music conditions (Järvinen-Pasley & Heaton, 2007). Whether capturing semantic information facilitates or hampers pitch discrimination performance may be dependent on the language involved. Indeed, as we mentioned in the Introduction, Mandarin is a tone language where pitch variation affects semantic meanings of individual words, whereas English is an intonation language where pitch carries no semantic information at the lexical level (Krishnan & Gandour, 2009). Therefore, lexical meaning may play a more essential role in facilitating discrimination of pitch differences for tone language speakers than for intonation language speakers. However, this phenomenon should be explained with caution since the facilitation of lexical meaning in pitch discrimination tasks is also dependent on task demand; that is, whether and to what extent the task requires lexical meaning to differentiate pitch variations. For example, a pitch discrimination task using Cantonese stimuli failed to observe an advantage in the processing of words compared to nonwords in both TD and ASD groups (Cheng et al., 2017). In contrast to the present study where stimulus pairs differed in lexical meaning (e.g., 'celestial being' versus 'idler'), the stimuli used in Cheng et al. (2017) were mainly differentiated in pitch intervals (e.g., no difference versus one to four semitones) while maintaining pitch contours. Another explanation may lie in the lexical status of the syllables used. While our nonwords contained real syllables in Mandarin, Cheng et al. (2017) used pseudo-syllables that were not in the Cantonese lexicon, leading to a more distinctive difference between experimental conditions. Future studies should explore how segmental and tonal features affect lexical processing independently and in combination in tone-language speakers with and without ASD.

V. CONCLUSION

In the present study, we examined whether pitch processing ability in Mandarin-speaking children with and without ASD would be affected by the nature of processing (form versus function) in tone and intonation tasks while controlling for other cognitive factors. The results suggested that the abilities to discriminate lexical tone and intonation may be unimpaired at both form and function levels in some cognitively competent autistic children. Similar to individuals with typical development, the abilities to process form- and function-level pitch were closely associated with one another in autistic individuals. Compared to the abilities to process form- and function-level pitch in the TD group, those abilities in the ASD group were more susceptible to cognitive factors, including nonverbal IQ, receptive ability, short-term memory, and musical training background. In addition, pitch processing was associated with ASD severity levels, and the more severe the ASD symptoms, the worse the form- and function-level of pitch processing. Consistent with the literature on speakers of non-tonal languages, our current findings on Mandarin speakers support the notion that enhanced pitch processing might only occur in a subgroup of ASD individuals rather than being a general characteristic of the ASD population, even for those who have lifelong tone language experience. These results are unlikely due to issues related to sample size or sampling error, as evidenced by our post hoc power analysis. Given that no group-specific response bias was observed in either tone or intonation discrimination, the current findings on Mandarin-speaking children with ASD likely reflect their genuine ability to process pitch, rather than due to random response or bias in their responses. Thus, these findings suggest unimpaired pitch processing abilities at both the form and function levels and provide evidence for associated pitch processing across levels in cognitively competent ASD.

ACKNOWLEDGMENTS

This work was supported by a European Research Council (ERC) Starting Grant, ERC-StG-2015, CAASD, 678733, to F.L. and C.J., and a National Science Foundation Grant (BCS-1125144), a Research Grants Council (RGC) of Hong Kong grant (34000118) to P.C.M.W., a Singapore Ministry of Education Academic Research Fund (AcRF) Tier 1 RG71/18 and Tier 2 MOE2019-T2-1-125 to A.H.D.C. We would like to thank all our participants who volunteered their time, and Chen Zhao, Ellie Packham, and Roshni Vagadia for assistance with data collection. Li Wang and Sanrong Xiao contributed equally to this work.

APPENDIX

TABLE III. The list of 36 pairs of nonwords and 36 pairs of words used in the tone discrimination task.

No.	Nonword1	Nonword2	Word1	Word2
1	翻通 (fan1tong1)	翻同 (fan1tong2)	翻拍 (fan1pai1)	翻牌 (fan1pai2)
2	清分 (qing1fen1)	清愤 (qing1fen4)	高低 (gao1di1)	高地 (gao1di4)
3	猜间 （cai1jian1）	财间 （cai2jian1）	挥师 （hui1shi1）	回师 （hui2shi1）
4	孤修 （gu1xiu1）	故修 （gu4xiu1）	包修 （bao1xiu1）	报修 （bao4xiu1）
5	相牌 （xiang1pai2）	相拍 （xiang1pai1）	风蚀 （feng1shi2）	风湿 （feng1shi1）
6	恭劫 （gong1jie2）	公借 （gong1jie4）	初十 （chu1shi2）	初试 （chu1shi4）
7	申条 （shen1tiao2）	神条 （shen2tiao2）	仙人 （xian1ren2）	闲人 （xian2ren2）
8	东绝 （dong1jue2）	动绝 （dong4jue2）	私刑 （si1xing2）	肆行 （si4xing2）
9	心帐 （xin1zhang4）	新章 （xin1zhang1）	心迹 （xin1ji4）	心机 （xin1ji1）
10	剥过 （bo1guo4）	菠国 （bo1guo2）	剥落 （bo1luo4）	菠萝 （bo1luo2）
11	批照 （pi1zhao4）	皮照 （pi2zhao4）	方位 （fang1wei4）	防卫 （fang2wei4）
12	昌势 （chang1shi4）	唱世 （chang4shi4）	孤寂 （gu1ji4）	顾忌 （gu4ji4）
13	糖声 （tang2sheng1）	汤声 （tang1sheng1）	晴天 （qing2tian1）	青天 （qing1tian1）
14	神超 （shen2chao1）	神潮 （shen2chao2）	陪都 （pei2du1）	陪读 （pei2du2）
15	形筐	形况	结晶	洁净

	（xing2kuang1）	（xing2kuang4）	（jie2jing1）	（jie2jing4）
16	遗机 （yi2ji1）	异基 （yi4ji1）	鱼鹰 （yu2ying1）	育婴 （yu4ying1）
17	墙文 （qiang2wen2）	枪文 （qiang1wen2）	银元 （yin2yuan2）	因缘 （yin1yuan2）
18	随情 （sui2qing2）	随青 （sui2qing1）	呈祥 （cheng2xiang2）	城厢 （cheng2xiang1）
19	游骸 （you2hai2）	游害 （you2hai4）	求实 （qiu2shi2）	求是 （qiu2shi4）
20	爬杰 （pa2jie2）	怕劫 （pa4jie2）	豪杰 （hao2jie2）	浩劫 （hao4jie2）
21	神地 （shen2di4）	申地 （shen1di4）	详尽 （xiang2jin4）	相近 （xiang1jin4）
22	雄动 （xiong2dong4）	雄东 （xiong2dong1）	传动 （chuan2dong4）	船东 （chuan2dong1）
23	残露 （can2lu4）	残炉 （can2lu2）	尝试 （chang2shi4）	常识 （chang2shi2）
24	豪践 （hao2jian4）	浩件 （hao4jian4）	实践 （shi2jian4）	事件 （shi4jian4）
25	现息 （xian4xi1）	先息 （xian1xi1）	信封 （xin4feng1）	新风 （xin1feng1）
26	会缸 （hui4gang1）	回缸 （hui2gang1）	浴缸 （yu4gang1）	鱼缸 （yu2gang1）
27	放方 （fang4fang1）	放房 （fang4fang2）	大虾 （da4xia1）	大侠 （da4xia2）
28	上星 （shang4xing1）	上性 （shang4xing4）	夜班 （ye4ban1）	夜半 （ye4ban4）
29	向楼 （xiang4lou2）	相楼 （xiang1lou2）	现行 （xian4xing2）	先行 （xianx1ing2）

30	泄人 (xie4ren2)	邪人 (xie2ren2)	菜园 (cai4yuan2)	裁员 (cai2yuan2)
31	复祥 (fu4xiang2)	付厢 (fu4xiang1)	过时 (guo4shi2)	过失 (guo4shi1)
32	复逃 (fu4tao2)	复套 (fu4tao4)	富于 (fu4yu2)	富裕 (fu4yu4)
33	向质 (xiang4zhi4)	乡制 (xiang1zhi4)	向背 (xiang4bei4)	相悖 (xiang1bei4)
34	橡饰 (xiang4shi4)	详侍 (xiang2shi4)	世纪 (shi4ji4)	实际 (shi2ji4)
35	报线 (bao4xian4)	报先 (bao4xian1)	惦记 (dian4ji4)	奠基 (dian4ji1)
36	问样 (wen4yang4)	问阳 (wen4yang2)	定计 (ding4ji4)	定级 (ding4ji2)

TABLE IV. The list of 20 statement-question pairs used in the intonation discrimination task with bolded texts representing the position of focus in the sentences.

No.	Initial focus	No.	Final focus	English meaning
1	**象**最大。/? (**Xiang4**zui4da4)	11	象最**大**。/? (Xiang4zui4**da4**)	The elephant is the biggest. /?
2	**猫**喝汤。/? (**Mao1**he1tang1)	12	猫喝**汤**。/? (Mao1he1**tang1**)	The cat drinks soup. /?
3	**顾俊**做饭。/? (**Gu4jun4**zuo4fan4)	13	顾俊**做饭**。/? (Gu4jun4**zuo4fan4**)	Gu Jun cooks. /?
4	**郭珊**出差。/? (**Guo1shan1**chu1chai1)	14	郭珊**出差**。/? (Guo1shan1**chu1chai1**)	Guo Shan is on a business trip. /?
5	**陆丽**在放假。/? (**Lu4li4**zai4fang4jia4)	15	陆丽在**放假**。/? (Lu4li4zai4**fang4jia4**)	Lu Li is on holiday. /?
6	**方晶**刚出生。/? (**Fang1jing1**gang1 chu1sheng1)	16	方晶刚**出生**。/? (Fang1jing1gang1 **chu1sheng1**)	Fang Jing was just born. /?

7	**赵亮**在看电视。/? (**Zhao4liang4**zai4kan4dian4shi4)	17	赵亮在看**电视**。/? (Zhao4liang4zai4kan4**dian4shi4**)	Zhao Liang is watching TV. /?
8	**张欣**刚刚出发。/? (**Zhang1xin1**gang1gang1chu1fa1)	18	张欣刚刚**出发**。/? (Zhang1xin1gang1gang1**chu1fa1**)	Zhang Xin just set off. /?
9	**杜秀**最爱做运动。/? (**Du4xiu4**zui4ai4zuo4yun4dong4)	19	杜秀最爱做**运动**。/? (Du4xiu4zui4ai4zuo4**yun4dong4**)	Du Xiu loves doing sports. /?
10	**张丹枫**关心周菲。/? (**Zhang1dan1feng1**guan1xin1zhou1fei1)	20	张丹枫关心**周菲**。/? (Zhang1dan1feng1guan1xin1**zhou1fei1**)	Zhang Danfeng cares for Zhou Fei. /?

REFERENCES

Acton, G. S., & Schroeder, D. H. (2001). Sensory discrimination as related to general intelligence. *Intelligence*, *29*(3), 263–271. https://doi.org/10.1016/S0160-2896(01)00066-6

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. American Psychiatric Association. http://dsm.psychiatryonline.org/doi/book/10.1176/appi.books.9780890425596

Asaridou, S. S., & McQueen, J. M. (2013). Speech and music shape the listening brain: Evidence for shared domain-general mechanisms. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00321

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Cross-domain Effects of Music and Language Experience on the Representation of Pitch in the Human Auditory Brainstem. *Journal of Cognitive Neuroscience*, *23*(2), 425–434. https://doi.org/10.1162/jocn.2009.21362

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

Bishop, D. V., Whitehouse, A. J. O., & Sharp, M. (2009). *The Communication Checklist-Self Report (CC-SR)*. Pearson.

Boersma, P., & Weenink, D. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*, 341–345.

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

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

Bradley, E. D. (2012). Tone language experience enhances sensitivity to melodic contour. *LSA Annual Meeting Extended Abstracts*, *3*(0), Article 0. https://doi.org/10.3765/exabs.v0i0.612

Burnham, D., Brooker, R., & Reid, A. (2015). The effects of absolute pitch ability and musical training on lexical tone perception. *Psychology of Music*, *43*(6), 881–897. https://doi.org/10.1177/0305735614546359

Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., Rattanasone, N. X., Schwarz, I.-C., & Webster, D. (2015). Universality and language-specific experience in the perception of lexical tone and pitch. *Applied Psycholinguistics*, *36*(6), 1459–1491. https://doi.org/10.1017/S0142716414000496

Chen, F., & Peng, G. (2021). Categorical Perception of Pitch Contours and Voice Onset Time in Mandarin-Speaking Adolescents With Autism Spectrum Disorders. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/2021_JSLHR-20-00725

Cheng, S. T. T., Lam, G. Y. H., & To, C. K. S. (2017). Pitch Perception in Tone Language-Speaking Adults With and Without Autism Spectrum Disorders. *I-Perception*, *8*(3), 2041669517711200. https://doi.org/10.1177/2041669517711200

Chevallier, C., Noveck, I., Happé, F., & Wilson, D. (2009). From acoustics to grammar: Perceiving and interpreting grammatical prosody in adolescents with Asperger Syndrome. *Research in Autism Spectrum Disorders*, *3*(2), 502–516. https://doi.org/10.1016/j.rasd.2008.10.004

Chiang, H.-M., Tsai, L. Y., Cheung, Y. K., Brown, A., & Li, H. (2014). A meta-analysis of differences in IQ profiles between individuals with Asperger's disorder and high-functioning autism. *Journal of Autism and Developmental Disorders*, *44*(7), 1577–1596. https://doi.org/10.1007/s10803-013-2025-2

Chowdhury, R., Sharda, M., Foster, N. E. V., Germain, E., Tryfon, A., Doyle-Thomas, K., Anagnostou, E., & Hyde, K. L. (2017). Auditory Pitch Perception in Autism Spectrum Disorder Is Associated With Nonverbal Abilities. *Perception*, 0301006617718715. https://doi.org/10.1177/0301006617718715

Creel, S. C., Weng, M., Fu, G., Heyman, G. D., & Lee, K. (2018). Speaking a tone language enhances musical pitch perception in 3–5-year-olds. *Developmental Science*, *21*(1), e12503. https://doi.org/10.1111/desc.12503

Deary, I. J., Head, B., & Egan, V. (1989). Auditory inspection time, intelligence and pitch discrimination. *Intelligence*, *13*(2), 135–147. https://doi.org/10.1016/0160-2896(89)90012-3

Deguchi, C., Boureux, M., Sarlo, M., Besson, M., Grassi, M., Schön, D., & Colombo, L. (2012). Sentence pitch change detection in the native and unfamiliar language in musicians and non-musicians: Behavioral, electrophysiological and psychoacoustic study. *Brain Research*, *1455*, 75–89. https://doi.org/10.1016/j.brainres.2012.03.034

Diehl, J. J., Watson, D., Bennetto, L., Mcdonough, J., & Gunlogson, C. (2009). An acoustic analysis of prosody in high-functioning autism. *Applied Psycholinguistics*, *30*(3), 385–404. https://doi.org/10.1017/S0142716409090201

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00781

Duanmu, S. (2007). *The Phonology of Standard Chinese* (Second Edition). The Phonology of the World's Languages.

Duda, M., Kosmicki, J. A., & Wall, D. P. (2014). Testing the accuracy of an observation-based classifier for rapid detection of autism risk. *Translational Psychiatry*, *4*(8), e424. https://doi.org/10.1038/tp.2014.65

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised*. American Guidance Service.

Eady, S. J., & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *The Journal of the Acoustical Society of America*, *80*(2), 402–415. https://doi.org/10.1121/1.394091

Eigsti, I.-M., de Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders*, *5*(2), 681–691. https://doi.org/10.1016/j.rasd.2010.09.001

Fancourt, A., Dick, F., & Stewart, L. (2013). Pitch-change detection and pitch-direction discrimination in children. *Psychomusicology: Music, Mind, and Brain*, *23*(2), 73–81. https://doi.org/10.1037/a0033301

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

Filipe, M. G., Frota, S., Castro, S. L., & Vicente, S. G. (2014). Atypical Prosody in Asperger Syndrome: Perceptual and Acoustic Measurements. *Journal of Autism and Developmental Disorders*, *44*(8), 1972–1981. https://doi.org/10.1007/s10803-014-2073-2

Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S., & Wicha, N. Y. Y. (2011). Native Experience with a Tone Language Enhances Pitch Discrimination and the Timing of Neural Responses to Pitch Change. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00146

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS Scores for a Measure of Severity in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *39*(5), 693–705. https://doi.org/10.1007/s10803-008-0674-3

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics (Vol. 1)*. New York.

Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511616983

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d'. *Behavior Research Methods, Instruments, & Computers*, *27*(1), 46–51. https://doi.org/10.3758/BF03203619

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

Heaton, P. (2009). Assessing musical skills in autistic children who are not savants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1522), 1443–1447. https://doi.org/10.1098/rstb.2008.0327

Heaton, P., Hermelin, B., & Pring, L. (1998). Autism and Pitch Processing: A Precursor for Savant Musical Ability? *Music Perception: An Interdisciplinary Journal*, *15*(3), 291–305. https://doi.org/10.2307/40285769

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

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

Helmbold, N., Troche, S., & Rammsayer, T. (2006). Temporal information processing and pitch discrimination as predictors of general intelligence. *Canadian Journal of Experimental Psychology = Revue Canadienne De Psychologie Experimentale*, *60*(4), 294–306. https://doi.org/10.1037/cjep2006027

Hirst, D. J. (2005). Form and function in the representation of speech prosody. *Speech Communication*, *46*(3–4), 334–347. https://doi.org/10.1016/j.specom.2005.02.020

Järvinen-Pasley, A., & Heaton, P. (2007). Evidence for reduced domain-specificity in auditory processing in autism. *Developmental Science*, *10*(6), 786–793. https://doi.org/10.1111/j.1467-7687.2007.00637.x

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*(7), 1328–1340. https://doi.org/10.1007/s10803-007-0520-z

JASP Team. (2020). *JASP (Version 0.13.1) [Computer software]*.

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

Jones, C. R. G., Happé, F., Baird, G., Simonoff, E., Marsden, A. J. S., Tregay, J., Phillips, R. J., Goswami, U., Thomson, J. M., & Charman, T. (2009). Auditory discrimination and auditory sensory behaviours in autism spectrum disorders. *Neuropsychologia*, *47*(13), 2850–2858. https://doi.org/10.1016/j.neuropsychologia.2009.06.015

Kargas, N., López, B., Reddy, V., & Morris, P. (2015). The Relationship Between Auditory Processing and Restricted, Repetitive Behaviors in Adults with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *45*(3), 658–668. https://doi.org/10.1007/s10803-014-2219-2

Klein, D., Zatorre, R. J., Milner, B., & Zhao, V. (2001). A Cross-Linguistic PET Study of Tone Perception in Mandarin Chinese and English Speakers. *NeuroImage*, *13*(4), 646–653. https://doi.org/10.1006/nimg.2000.0738

Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and Language*, *110*(3), 135–148. https://doi.org/10.1016/j.bandl.2009.03.005

Krumhansl, C. L. (2004). The Cognition of Tonality – as We Know it Today. *Journal of New Music Research*, *33*(3), 253–268. https://doi.org/10.1080/0929821042000317831

Lai, G., Pantazatos, S. P., Schneider, H., & Hirsch, J. (2012). Neural systems for speech and song in autism. *Brain*, *135*(3), 961–975. https://doi.org/10.1093/brain/awr335

Lamont, A. (1998). Music, Education, and the Development of Pitch Perception: The Role of Context, Age and Musical Experience. *Psychology of Music*, *26*(1), 7–25. https://doi.org/10.1177/0305735698261003

Lau, J. C. Y., To, C. K. S., Kwan, J. S. K., Kang, X., Losh, M., & Wong, P. C. M. (2020). Lifelong Tone Language Experience does not Eliminate Deficits in Neural Encoding of Pitch in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-020-04796-7

Levrero, F., Mathevon, N., Pisanski, K., Gustafsson, E., & Reby, D. (2018). The pitch of babies' cries predicts their voice pitch at age 5. *Biology Letters*, *14*(7), 20180065. https://doi.org/10.1098/rsbl.2018.0065

Li, Y., Tang, C., Lu, J., Wu, J., & Chang, E. F. (2021). Human cortical encoding of pitch in tonal and non-tonal languages. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-21430-x

Liu, F., Jiang, C., Pfordresher, P. Q., Mantell, J. T., Xu, Y., Yang, Y., & Stewart, L. (2013). Individuals with congenital amusia imitate pitches more accurately in singing than in speaking: Implications for music and language processing. *Attention, Perception, & Psychophysics*, *75*(8), 1783–1798. https://doi.org/10.3758/s13414-013-0506-1

Liu, F., Jiang, C., Thompson, W. F., Xu, Y., Yang, Y., & Stewart, L. (2012). The Mechanism of Speech Processing in Congenital Amusia: Evidence from Mandarin Speakers. *PLOS ONE*, *7*(2), e30374. https://doi.org/10.1371/journal.pone.0030374

Liu, F., & Xu, Y. (2005). Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Phonetica*, *62*(2–4), 70–87. https://doi.org/10.1159/000090090

Liu, F., Xu, Y., Prom-on, S., & Yu, A. C. L. (2013). Morpheme-like prosodic functions: Evidence from acoustic analysis and computational modeling. *Journal of Speech Sciences*, *3*(1), 85–140.

Liu, M., Chen, Y., & Schiller, N. O. (2016). Online processing of tone and intonation in Mandarin: Evidence from ERPs. *Neuropsychologia*, *91*, 307–317. https://doi.org/10.1016/j.neuropsychologia.2016.08.025

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule–2nd edition (ADOS-2)*. CA: Western Psychological Corporation.

Lynn, R., Graham Wilson, R., & Gault, A. (1989). Simple musical tests as measures of spearman's g. *Personality and Individual Differences*, *10*(1), 25–28. https://doi.org/10.1016/0191-8869(89)90173-6

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide*. Lawrence Erlbaum Associates.

Makowski, D. (2018). The psycho Package: An Efficient and Publishing-Oriented Workflow for Psychological Science. *Journal of Open Source Software*, *3*(22), 470. https://doi.org/10.21105/joss.00470

Mantell, J. T., & Pfordresher, P. Q. (2013). Vocal imitation of song and speech. *Cognition*, *127*(2), 177–202. https://doi.org/10.1016/j.cognition.2012.12.008

Mayer, J. L., Hannent, I., & Heaton, P. F. (2016). Mapping the Developmental Trajectory and Correlates of Enhanced Pitch Perception on Speech Processing in Adults with ASD. *Journal of Autism and Developmental Disorders*, *46*(5), 1562–1573. https://doi.org/10.1007/s10803-014-2207-6

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. https://doi.org/10.1080/13682820601170102

Moore, R. E., Keaton, C., & Watts, C. (2007). The Role of Pitch Memory in Pitch Discrimination and Pitch Matching. *Journal of Voice*, *21*(5), 560–567. https://doi.org/10.1016/j.jvoice.2006.04.004

Mottron, L., & Bzdok, D. (2020). Autism spectrum heterogeneity: Fact or artifact? *Molecular Psychiatry*, 1–8. https://doi.org/10.1038/s41380-020-0748-y

Mottron, L., Peretz, I., Belleville, S., & Rouleau, N. (1999). Absolute pitch in autism: A case study. *Neurocase*, *5*(6), 485–501. https://doi.org/10.1080/13554799908402744

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

Nowell, K. P., Schanding, G. T., Kanne, S. M., & Goin-Kochel, R. P. (2015). Cognitive Profiles in Youth with Autism Spectrum Disorder: An Investigation of Base Rate Discrepancies using the Differential Ability Scales—Second Edition. *Journal of Autism and Developmental Disorders*, *45*(7), 1978–1988. https://doi.org/10.1007/s10803-014-2356-7

O'Connor, K. (2012). Auditory processing in autism spectrum disorder: A review. *Neuroscience & Biobehavioral Reviews*, *36*(2), 836–854. https://doi.org/10.1016/j.neubiorev.2011.11.008

Ong, J. H., Wong, P. C. M., & Liu, F. (2020). Musicians show enhanced perception, but not production, of native lexical tones. *The Journal of the Acoustical Society of America*, *148*(6), 3443–3454. https://doi.org/10.1121/10.0002776

O'Riordan, M., & Passetti, F. (2006). Discrimination in Autism Within Different Sensory Modalities. *Journal of Autism and Developmental Disorders*, *36*(5), 665–675. https://doi.org/10.1007/s10803-006-0106-1

Ouimet, T., Foster, N. E. V., Tryfon, A., & Hyde, K. L. (2012). Auditory-musical processing in autism spectrum disorders: A review of behavioral and brain imaging studies. *Annals of the New York Academy of Sciences*, *1252*, 325–331. https://doi.org/10.1111/j.1749-6632.2012.06453.x

Patel, A. D. (2008). *Music, Language, and the Brain.* Oxford University Press, USA.

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., Foxton, J. M., & Griffiths, T. D. (2005). Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. *Brain and Cognition*, *59*(3), 310–313. https://doi.org/10.1016/j.bandc.2004.10.003

Patel, A. D., Peretz, I., Tramo, M., & Labreque, R. (1998). Processing Prosodic and Musical Patterns: A Neuropsychological Investigation. *Brain and Language*, *61*(1), 123–144. https://doi.org/10.1006/brln.1997.1862

Patel, A. D., Wong, M., Foxton, J. M., Lochy, A., & Peretz, I. (2008). Speech intonation perception deficits in musical tone deafness (congenital amusia). *Music Perception*, *25*(4), 357–368. https://doi.org/10.1525/mp.2008.25.4.357

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. https://doi.org/10.1007/s10803-004-1999-1

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Peppé, S., & McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: The PEPS-C test and the revised version. *Clinical Linguistics & Phonetics*, *17*(4–5), 345–354. https://doi.org/10.1080/0269920031000079994

Peppé, S., McCann, J., Gibbon, F., O'Hare, A., & Rutherford, M. (2007). Receptive and Expressive Prosodic Ability in Children With High-Functioning Autism. *Journal of Speech, Language, and Hearing Research*, *50*(4), 1015–1028. https://doi.org/10.1044/1092-4388(2007/071)

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

Plack, C. J., Barker, D., & Hall, D. A. (2014). Pitch coding and pitch processing in the human brain. *Hearing Research*, *307*, 53–64. https://doi.org/10.1016/j.heares.2013.07.020

Plack, C. J., Oxenham, A. J., & Fay, R. R. (2006). *Pitch: Neural Coding and Perception*. Springer Science & Business Media.

Quintin, E.-M., Bhatara, A., Poissant, H., Fombonne, E., & Levitin, D. J. (2013). Processing of musical structure by high-functioning adolescents with autism spectrum disorders. *Child*

Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence, *19*(3), 250–275. https://doi.org/10.1080/09297049.2011.653540

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology*, *25*, 111–163. JSTOR. https://doi.org/10.2307/271063

Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven manual: Section 3. Standard progressive matrices.* Oxford Psychologists Press.

Raz, N., Willerman, L., & Yama, M. (1987). On sense and senses: Intelligence and auditory information processing. *Personality and Individual Differences*, *8*(2), 201–210. https://doi.org/10.1016/0191-8869(87)90175-9

Rong, Y., Weng, Y., Chen, F., & Peng, G. (2022). Categorical perception of Mandarin lexical tones in language-delayed autistic children. *Autism*, 13623613221138688. https://doi.org/10.1177/13623613221138687

RStudio Team. (2020). *RStudio: Integrated Development for R*. RStudio, PBC, Boston, MA. http://www.rstudio.com/

Sang, B., & Miao, X. (1990). The revision of trail norm of Peabody picture vocabulary test revised (PPVT-R) In Shanghai proper. *Psychological Science*, *5*, 20–25.

Schellenberg, E. G., & Weiss, M. W. (2013). 12—Music and Cognitive Abilities. In D. Deutsch (Ed.), *The Psychology of Music (Third Edition)* (pp. 499–550). Academic Press. https://doi.org/10.1016/B978-0-12-381460-9.00012-2

Sharda, M., Midha, R., Malik, S., Mukerji, S., & Singh, N. C. (2015). Fronto-Temporal Connectivity is Preserved During Sung but Not Spoken Word Listening, Across the Autism Spectrum. *Autism Research*, *8*(2), 174–186. https://doi.org/10.1002/aur.1437

Sota, S., Hatada, S., Honjyo, K., Takatsuka, T., Honer, W. G., Morinobu, S., & Sawada, K. (2018). Musical disability in children with autism spectrum disorder. *Psychiatry Research*, *267*, 354–359. https://doi.org/10.1016/j.psychres.2018.05.078

Spearman, C. (1904). 'General Intelligence,' Objectively Determined and Measured. *The American Journal of Psychology*, *15*(2), 201–292. JSTOR. https://doi.org/10.2307/1412107

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

Sun, K.-C., & Huang, T. (2012). A cross-linguistic study of Taiwanese tone perception by Taiwanese and English listeners. *Journal of East Asian Linguistics*, *21*(3), 305–327. https://doi.org/10.1007/s10831-012-9092-9

Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1430), 303–314. https://doi.org/10.1098/rstb.2002.1198

The National Language Working Committee. (2008). *Lexicon of Common Words in Contemporary Chinese* (Draft Edition). The Commercial Press.

Tillmann, B., Lévêque, Y., Fornoni, L., Albouy, P., & Caclin, A. (2016). Impaired short-term memory for pitch in congenital amusia. *Brain Research*, *1640*, 251–263. https://doi.org/10.1016/j.brainres.2015.10.035

Wang, L., Beaman, C. P., Jiang, C., & Liu, F. (2021). Perception and Production of Statement-Question Intonation in Autism Spectrum Disorder: A Developmental Investigation. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-021-05220-4

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*(1), 43254. https://doi.org/10.1038/srep43254

Watson, B. U. (1991). Some relationships between intelligence and auditory discrimination. *Journal of Speech and Hearing Research*, *34*(3), 621–627. https://doi.org/10.1044/jshr.3403.621

Wechsler, D. (2003). *Wechsler intelligence scale for children–Fourth Edition (WISC-IV)*. TX: The Psychological Corporation.

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

Williams, T. I., Loucas, T., Sin, J., Jeremic, M., Aslett, G., Knight, M., Fincham-Majumdar, S., & Liu, F. (2021). A randomised controlled feasibility trial of music-assisted language telehealth intervention for minimally verbal autistic children—The MAP study protocol. *Pilot and Feasibility Studies*, *7*(1), 182. https://doi.org/10.1186/s40814-021-00918-9

Wong, P. C. M., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, *28*(04), 565–585.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420–422. https://doi.org/10.1038/nn1872

Xu, Y. (2005). Speech melody as articulatorily implemented communicative functions. *Speech Communication*, *46*(3–4), 220–251. https://doi.org/10.1016/j.specom.2005.02.014

Xu, Y. (2019). Prosody, tone, and intonation. In *The Routledge Handbook of Phonetics* (1st Edition, pp. 314–356). Routledge.

Xu, Y., Gandour, J. T., & Francis, A. L. (2006). 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. https://doi.org/10.1121/1.2213572

Yip, M. (2002). *Tone*. Cambridge University Press.

Young, R. L., & Nettelbeck, T. (1995). The abilities of a musical savant and his family. *Journal of Autism and Developmental Disorders*, *25*(3), 231–248. https://doi.org/10.1007/BF02179286

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

Yuan, J. (2006). Mechanisms of Question Intonation in Mandarin. In Q. Huo, B. Ma, E.-S. Chng, & H. Li (Eds.), *Chinese Spoken Language Processing* (Vol. 4274, pp. 19–30). Springer Berlin Heidelberg. https://doi.org/10.1007/11939993_7

Yuan, J. (2011). Perception of intonation in Mandarin Chinese. *The Journal of the Acoustical Society of America*, *130*(6), 4063–4069. https://doi.org/10.1121/1.3651818

Zhang, H. (1989). Standardization research on Raven's standard progressive matrices in China. *Acta Psychologica Sinica*, *21*(02), 3–11.

Zhang, J., Meng, Y., Wu, C., Xiang, Y.-T., & Yuan, Z. (2019). Non-speech and speech pitch perception among Cantonese-speaking children with autism spectrum disorder: An ERP study. *Neuroscience Letters*, *703*, 205–212. https://doi.org/10.1016/j.neulet.2019.03.021

Zhou, X., & Marslen-Wilson, W. (1994). Words, morphemes and syllables in the Chinese mental lexicon. *Language and Cognitive Processes*, *9*, 393–422. https://doi.org/10.1080/01690969408402125