Successful lexical tone production of Mandarin Chinese autistic children with intellectual impairment

Ao Chen[1,2] | **Ru Zhao**[1,3] | **Gan Huang**[4] | **Aijun Li**[4] | **Hintat Cheung**[5]

[1]School of Psychology, Beijing Language and Culture University, Beijing, China
[2]School of Communication Science, Beijing Language and Culture University, Beijing, China
[3]The Children's Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China
[4]Institute of Linguistics, Chinese Academy of Social Sciences, Beijing, China
[5]Department of Audiology and Speech-Language Pathology, Asia University, Taichung, Taiwan, China

Correspondence
Ao Chen, School of Psychology, Beijing Language and Culture University, Room 314, Office Building, Beijing Language and Culture University, Beijing, 100083, China.
Email: irischen71@hotmail.com

Funding information
Beijing Social Science Fund, Grant/Award Number: 18YYC013; National Social Science Fund of China, Grant/Award Number: 17CYY041;
Fundamental Research Funds for the Central Universities, Grant/Award Number: 21PT01

Abstract

Background: Atypical speech prosody has been commonly found among autistic children. Yet it remains unknown whether prosody impairment originates from poor pitch ability in general or whether it is the result of the difficulty in understanding and using prosody for communicative purposes.

Aims: To investigate whether native Mandarin Chinese-speaking autistic children with intellectual impairment were able to accurately produce native lexical tones, which are pitch patterns that distinguish word meaning lexically and serve little social purpose.

Methods & Procedures: Using a picture-naming task, thirteen 8–13-year-old Mandarin Chinese-speaking autistic children with intellectual impairment were tested on their production of Chinese lexical tones. Chronical age-matched typically developing (TD) children were included as the control group. Perceptual assessment and phonetic analyses were conducted with the produced lexical tones.

Outcomes & Results: The majority of the lexical tones produced by the autistic children were perceived as accurate by adult judges. Phonetic analysis of the pitch contours found no significant difference between the two groups, and the autistic children and TD children used the phonetic features in comparable ways when differentiating the lexical tones. However, the lexical tone accuracy rate was lower among the autistic children than among the TDs, and the larger individual difference was observed among the autistic children than the TD children.

Conclusions & Implications: These results indicate that autistic children are able to produce the global contours of the lexical tones, and pitch deficits do not seem to qualify as a core feature of autism.

KEYWORDS
autism spectrum disorders (ASD), lexical tone, prosody, speech

What this paper adds

What is already known on the subject

- Atypical prosody has been considered a maker of the speech of autistic children, and meta-analysis found a significant difference in mean pitch and pitch range between TD children and autistic children. Yet it remains unknown whether the pitch deficits are the result of impaired perceptual–motoric ability or if they reflect failure in learning sentential prosody, which requires an understanding of the interlocutors' mind. In addition, research on pitch ability of autistic children with intellectual disabilities has been scarce, and whether these children are able to produce pitch variation is largely unknown.

What this paper adds to existing knowledge

- We tested native Mandarin Chinese autistic children with intellectual impairment on their production of native lexical tones. The lexical tones in Chinese are pitch variations realized on individual syllables that distinguish lexical meaning, but they do not serve social pragmatic purposes. We found that although these autistic children had only developed limited spoken language, the majority of their lexical tones were perceived as accurate. They were able to use the phonetic features in comparable ways with the TD children when distinguishing the lexical tones.

What are the potential or actual clinical implications of this work?

- It seems unlikely that pitch processing at the lexical level is fundamentally impaired in autistic children, and pitch deficits do not seem to qualify for a core feature of their speech. Practitioners should be cautious when using pitch production as a clinical marker for autistic children.

INTRODUCTION

Individuals with autism spectrum disorder (ASD) have persistent deficits in social communication and interpersonal interaction and show restricted repetitive patterns of behaviour, interests and activities (Lord et al., 2020). Such symptoms start early in life. As speech and language serve as the most important tool for communication, it is unsurprising that impairment in language development is common among autistic children. Atypical language development and language delay are among the most evident red flags for young children who are later diagnosed with ASD (Robins et al., 2014).

Disordered prosody is one of the most frequently reported atypical features of speech produced by autistic children (Baltaxe & Simmons, 1985; Baltaxe & Guthrie, 1987; Paul et al., 2005), and anomalous prosody of autistic children occurs early. Infants who were later diagnosed with ASD showed less complex prosodic modulation than typically developing (TD) children in the first a few months of life (Brisson et al., 2014). Speech of autistic children has been described as monotonic or machine-like, showing inappropriate volume and stress (Arciuli & Benjamin, 2019; DeMyer et al., 1973; Kanner, 1971; Rutter & Lockyer, 1967; Shriberg Lawrence et al., 2001; Simmons & Baltaxe, 1975). The most reliable deficits among autistic children are perhaps atypical pitch mean and pitch range (Fusaroli et al., 2017). In particular, lifelong sensory experience does not seem to eliminate linguistic pitch processing deficits in perception (Lau et al., 2021). Nevertheless, since pitch deficits are not unequivocally found in all studies (Fusaroli et al., 2017), it remains inconclusive whether pitch deficits can be considered as a core feature of the atypical prosody of autistic children.

Perhaps one reason for the inconsistent findings on pitch production among autistic children is the different functions of pitch. Cross-linguistically, pitch variation serves social–pragmatic purposes at sentential level. For example, the distinction between questions, which are often realized with rise of fundamental frequency (f0, perceived as voice pitch), and statements, which are often realized with pitch fall (Gussenhoven, 2004), reflects different intentions of speakers. Prosody also cues emotions such as joy, sadness, anger and fear, with anger and fear being associated with increased f0 level and f0 range (Banse & Scherer, 1996). Speakers tend to accent new and/or important information with increased pitch and loudness (Halliday, 1970; Xu, 2011). For the sentence, *I bought the red table*, an accent will be placed on 'red' if the speaker aims to emphasize the colour, probably contrasting with other tables in different colours, while if the important information is the object, an accent will be placed on 'table' rather than 'red'. As can be seen, efficient processing of sentential level pitch variation entails understanding other people's communicative purposes. By definition, autistic children are impaired in social communication across multiple contexts (*The Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. (DSM−5). Thus, it is unsurprising that they have difficulties mastering the social function of pitch. It is widely reported that even those children with high-functioning autism (HFA), are unable to produce sentential prosody accurately. They use pitch accents inappropriately, produce atypical boundary tones, and demonstrate greater pitch range and variance (Diehl et al., 2009; Fosnot & Jun, 1999; Green & Tobin, 2009; Hubbard & Trauner, 2007; Nadig & Shaw, 2012).

At the lexical level, on the contrary, pitch can contrast word meaning without serving any social–communicative purpose. For example, in stress languages such as English, location of the stressed syllable, which is often realized with higher pitch and longer duration, differentiates word meaning, such as REcord and reCORD (Cutler & Van Donselaar, 2001). In tone languages such as Mandarin Chinese, lexical tones are pitch variations realized on single syllables that distinguish lexical meaning. A common example is that the same syllable /ma/ means *mom*, *hemp*, *horse* and *scold*, respectively when carrying high-level (T1), mid-rising (T2), low-dipping (T3) and high-falling (T4). The phonetic implementation of lexical tones is barely influenced by intonation (Chao, 1933). Therefore, lexical tones distinguish meaning *lexically*, and do not encode speakers' intention or emotion. It has been hypothesized that the grammatical aspects of speech prosody are relatively intact among people with ASD while they have persistent difficulties with the pragmatic and affective aspects of prosody (Järvinen-Pasley et al., 2008; Shriberg et al., 2001, 2011). Indeed, some previous studies seem to suggest that HFA children were able to contrast lexical stress (i.e., strong–weak and weak–strong words), and word level pitch variation did not seem to differ significantly between TD and HFA children, although atypical long duration and more variable f0 have been observed (Grossman et al., 2010; Nakai et al., 2014; Van Santen et al., 2010; Xu, 2013). Other studies, however, found that stress, both pragmatic and lexical, formed an area of difficulty for HFA children (Arciuli et al., 2020; Paul et al., 2005; Shriberg Lawrence et al., 2001). When participants were not restricted to children with HFA or Asperger Syndrome (AS), autistic children as a whole misplaced and reduced lexical stress more often than their TD peers (McAlpine et al., 2014). It should be noted that among these studies that investigated word-level prosody, different methods were employed, such as non-word repetition (Grossman et al., 2010; Van Santen et al., 2010) and picture naming (Naikai et al., 2014). In some studies, (non-)words were embedded in multi-syllable sequences with intonational patterns (Grossman et al., 2010) while in others these were elicited in isolation (van Santen et al., 2010). The different paradigms made it difficult to ascertain whether atypical lexical level prosody was driven by pitch deficits at sentential level as the result of communicative impairment, or it reflects broad impairment in pitch production regardless whether or not pitch serves social–pragmatic purposes.

Anomalous sensory perceptual processing might underlie the atypical production of pitch patterns. Autistic children have been found to pay excessive attention to details (Happé & Frith, 2006) and displayed enhanced perceptual processing and low-level discrimination (Mottron et al., 2006; Plaisted, 2001). Such detail-focused processing styles may undermine their ability to categorize phonetic and phonological categories, which may consequently lead to atypical expressive phonology (Oller et al., 2010). With regard to pitch, although aberrant cortical discrimination of frequency and pitch change has been widely documented in autistic children (Kujala et al., 2013), it remains debated whether autistic children were superior (Bonnel et al., 2003; Gomot et al., 2002; Heaton, 2003, 2005; Heaton et al., 2008) or disordered (Jansson-Verkasalo et al., 2005; Kujala et al., 2010; Russo et al., 2008) in pitch processing compared with TD children. If perceptual pitch encoding is deficient among autistic children, we would expect less accurate pitch production in general, no matter whether pitch serves lexical or social–communicative purposes; whereas if autistic children are enhanced in pitch perception, we would expect accurate production of *lexical* level pitch patterns, such as lexical tones, since these do not require understanding of other people's intention.

Another reason that may explain the inconsistent findings on pitch processing among autistic children is heterogeneity of the population, in terms of both severity and language profile. Kjelgaard and Tager-Flusberg (2001) showed that autistic children between 4 and 14 years of age were heterogenous in terms of language as well as articulation ability. (Severe) intellectual impairment often co-occurs with autism, in particular among those children at the lower end of the autism spectrum (Goldin et al., 2014). These children may never develop verbal language while children with HFA may exhibit unimpaired structural language (Boucher, 2012; McCann & Peppé, 2003). It is still debated whether articulatory impairment can differentiate subtypes of autism. When the children were divided into language impaired, typical language and language-borderline subgroups, only those language impaired children showed mild articulatory impairment, while the other two groups showed intact articulation. Among school aged autistic children, cluster analysis found two major subtypes of language impairment: either with or without impaired expressive phonology (Rapin et al., 2009). For children with HFA, some studies found articulatory distortions or errors (Cleland et al., 2010; Shriberg Lawrence et al., 2001) while others found that articulation was spared among various language disorders (Kjelgaard & Tager-Flusberg, 2001). It should be noted that these studies often tested children with HFA or AS, while research on less able autistic children is scarce. Since previous studies have found different prosodic profiles of children with HFA and children/adolescents with ASD (Peppé et al., 2011), it is plausible that compared with the children with HFA and AS, autistic children with concomitant intellectual impairment may produce pitch patterns with more notable or different prosodic errors. Therefore, to understand whether pitch production impairment can be considered a core feature of autistic children, it is crucial to test autistic children with intellectual impairment.

To shed light on whether atypical pitch production could be considered a labelling feature of autistic children, the current study tested Mandarin Chinese-speaking autistic children with intellectual impairment on lexical tone production. Pitch contour alone, with f0 being its physical attribute, is sufficient for distinguishing the Mandarin Chinese lexical tones. When f0 information is present, other acoustic cues such as duration or amplitude are negligible for lexical tone identification (Howie, 1976; Whalen & Xu, 1992). Importantly, lexical tones maintain their pitch contours in the face of intonational pitch variation, and they faithfully approach their tonal target despite pitch modification at the sentential level (e.g., focus realization and affect expression, Chao, 1933; Xu, 2011; Xu, 1997). Therefore, lexical tones are a suitable vehicle to delineate productive pitch ability among autistic children, in that accurate production of the lexical tones is not dependent on learning sentential level prosody, thus it does not require understanding the intention of the interlocutor. If the atypical pitch production of autistic children results from impaired communicative ability, lexical tone production should be intact as it serves no social–pragmatic purposes. On the other hand, if pitch processing is fundamentally impaired among autistic children as the result of deficient sensory and/or motoric ability, then their production of lexical tones should be disordered.

For native Mandarin Chinese speakers, lexical tones have been found to be acquired earlier than consonants and vowels (Zhu & Dodd, 2000). Studies with TD children showed that, by 3 years, children's production of the T1, T2 and T4 reached an accuracy above 70%, and T3 had a production accuracy of about 40% (Wong et al., 2005). Xu (2013) tested lexical tone production of TD teenagers and teenagers with HFA or AS without language impairment, and found no significant difference between the two groups on phonetic measurements regardless of whether the lexical tones were produced on isolated syllables or in sentences. Wu et al. (2020) showed that the lexical tones produced by children with autism, both those between 3 and 6 years and those over 6 years of age, were assessed as accurate by adult judges. This study, however, did not specify whether the children were high- or low-functioning, and the authors tested the production of only four words (each carried a different lexical tone), hence whether their findings are representative of lexical tone production among autistic children remains unknown. Another recent study showed that native Mandarin speaking children with HFA were able to contrast the Cantonese lexical tones in a repetition task, yet phonetic analyses found larger individual variability compared with TD children (Chen et al., 2022). So far, there is little evidence on whether low functioning autistic children can produce lexical tones accurately.

The current study tested lexical tone production of autistic children with intellectual impairment. Both perceptual assessment and phonetic analysis were conducted with their lexical tone productions. If these autistic children were able to produce the lexical tones accurately, it is unlikely that pitch processing impairment qualifies for a core feature of autism, and their difficulties with sentential prosody cannot be attributed to inability of pitch perception or production per se.

METHODS AND MATERIALS

Participants

A total of 13 autistic children (11 boys, mean age (SD) = 11.51 (2.02) years) and 13 TD children (eight boys, mean age (SD) = 11.35 (0.77) years) participated in the current study.

Mandarin Chinese was the home and school language for all the children. All children were reported by their parents to have normal (corrected to normal) vision and normal hearing. The autistic children were recruited from special education schools in Beijing, except one boy who was attending a mainstream school at the time of the experiment but moved to a special education school shortly after. All the children were urban residents in Beijing. They had a formal clinical diagnosis of autism given by qualified clinicians in local hospitals in Beijing. In China, clinical diagnosis is based on interview of the children and parents as well as results of commonly used assessment batteries that have been adapted for use in China, such as Autism Behaviour Checklist (Krug et al., 1980) and Childhood Autism Rating Scale (Schopler et al., 1980). Children's general development such as height, weight, cognitive ability, etc. was also assessed. The guidelines for clinical assessment of autism can be found at the website of The Central People's Government of the People's Republic of China (http://www.gov.cn/zwgk/2010-08/16/content_1680727.htm). Wechsler Intelligence Scale for Children (WISC) scores were available for eight autistic children, with the mean full-scale score being 57 (SD = 10). The remaining six children had a formal statement of intellectual disability from China Disabled Persons' Federation, and the severity of the disability was classified as level 2 (equals to 20–34 WISC full-scale score) or level 3 (equals to 35–49 WISC full-scale score). All the autistic children had oral language, and were able to understand their teacher's instruction during the experiment, yet none of them was able to converse fluently with spoken language. The TD children were recruited from mainstream schools, and were reported by their parents to have no intellectual or neurological impairments. It should be acknowledged that the age and gender of the autistic children and the TD children were not completely matched. However, Mandarin Chinese does not contrast between level tones and f0 contour alone is sufficient to define the lexical tones, therefore individual speaker's pitch level is unlikely to distort the lexical tones. All the parents gave oral consent for their children's participation in the study.

Materials

A picture-naming task was adopted. Clipart pictures of common objects (e.g., a backpack) and movements (e.g., running) were used as visual stimuli to elicit lexical tones. Only one object or one movement was presented on each picture, except for 'mouth' and 'foot', which were presented on one single picture of a boy. The 25 target words and one example picture are listed in Appendix A1 and Figure A1. These words were tested among TD 3–4-year-old kindergarteners in Beijing before the experiment, and all these kindergarteners were able to name the pictures correctly. For the bisyllabic words, tone sandhi was avoided (Duanmu, 2006), and the lexical tones of the two syllables were compatible (i.e., the tonal offset of the first syllable and the tonal onset of the second syllable were comparable in terms of pitch level, such as T1T4, where T1 has a high offset and T4 has a high onset). Compatible bisyllabic words ensured that the lexical tones of the bisyllabic words maintained its citation form (Xu, 1997), which allowed us to pool the monosyllabic and bisyllabic words together for phonetic analysis.

Procedure

Given the autistic children's social difficulty, each child was tested by a teacher familiar to them from the special education school they were attending at the time of the experiment. Before the experiment, all the teachers were trained with the testing materials and procedure so that a standard protocol could be followed. The pictures were presented on A4-sized papers or on a screen of comparable size. All the teachers were native Mandarin Chinese speakers. The instructions were as follows: 'I have many pictures here, let's go through them together one by one, could you please tell me what you see on each picture?' The teacher only started presenting the visual stimuli if the participant was attending to them. Two practice trials preceded the experiment, where pictures of a baby and a girl were presented in the first and second trial, respectively. If a participant gave an oral response in one of the practice trials, regardless whether the picture was correctly named, the experiment proceeded, and if a participant failed to give any oral response in the practice trial, the experiment terminated. All the participants passed practice trials and proceeded to the experiment. The teachers were allowed to give positive feedbacks orally after each trial, such as 很好, meaning 'good job' to keep the child interested in the task, and they were asked to redirect the participant to the stimuli if he or she got distracted in the experiment. If a participant failed to give a response for a picture, the teacher would say the target word and then ask the participant to repeat. If the participant failed to respond after two attempts to elicit repetition, the experiment proceeded to the next trial. If the participant failed to name the picture with the target word but a synonym (e.g., /men/ meaning 'door' instead of the target /mumen/ meaning 'wooden door'), the teacher would say the target word and ask the participant to repeat. A smartphone was placed within 30 cm of the participant and video recorded the experiment. The children were tested in a quiet room, and the sound quality of the smartphone was sufficient for pitch analysis.

The TD children were tested by a familiar teacher, who was trained and followed the same testing procedure as for the autistic children. It took the autistic children about 15 min to finish the procedure, and the TD children were able to finish the test within 5 min.

Data analysis

The audio was split from the video using a PC app (http://www.pcgeshi.com/). Each syllable was first manually transcribed by a trained phonetician. All the syllables were analysed (i.e., monosyllabic words, the first and second syllables in the bisyllabic words). Target words produced in the sentence were excluded (e.g., when the target is *pingguo* /p^{h}iŋ1kuo3/, meaning 'apple', and participant responded with the sentence *zhe shi pingguo* meaning 'this is an apple', then this trial was excluded from analysis). For the autistic children, 10 words were excluded for whispering or overlapping with the teacher's speech. They produced 17 words on average (SD = 4.5). A total of 222 words (377 lexical tones) were included in the final analysis, among which 53 words were produced as repetitions. The number of T1, T2, T3 and T4 produced was as follows: 133, 126, 52 and 66, respectively. The TD children produced 325 words (532 lexical tones), with the number of T1, T2, T3 and T4 being 182, 181, 78 and 91, respectively.

The lexical tones produced by the participants were assessed both perceptually and phonetically. For the perceptual assessment, one native Mandarin Chinese adult judge (age 35, female) listened to the words produced by the children through earphones. For each word, the target word was presented on a computer screen in Chinese characters, and she was asked to indicate whether the lexical tones were produced accurately or not on an excel spreadsheet. For both the TD and the autistic children, a randomly selected 80 words (160 words in sum) were selected and presented to another native Mandarin Chinese judge (age 36, male) for consistency assessment, and the second judge followed the same protocol as the first one.

For the phonetic analysis, for each syllable, the lexical tone was extracted from the vocalic portion of the target syllable (and nasal if present). For each speaker, a proper setting for pitch floor and ceiling were selected, where the range of pitch floor was 75–100 Hz, and the range of pitch ceiling is 500–650 Hz. Following Wong (2012), for T1 and T4, within the lexical tone, two landmarks were annotated, which were pitch valley (L, located at F0 minimum) and pitch peak (H, located at F0 maximum). Since T2 and T3 are similar in contour (i.e., f0 first decreases and then increases), and the timing of the 'turning point' (i.e., the point of F0 direction change) and the decrease in F0 from the onset of the tone to the turning point are relevant phonetic features for identification (Moore & Jongman, 1997), three landmarks were annotated on the f0 contour of T2 and T3, namely H1 (i.e., maximum f0 before minimum f0), L (i.e., minimum f0) and H2 (i.e., maximum f0 after minimum f0). For T2 and T3, the max f0 was defined as highest f0 value of the vocalic part, thus f0 value at H1 if this was larger than the f0 value at H2, and f0 value at H2 if the opposite. Table 1 lists the acoustical parameters analysed in the current study and their operational definitions (Wong, 2012). Using these parameters, the contours of the lexical tones could be characterized despite potentially variable pitch levels of individual speakers. To compare whether the autistic children differed from TD children in lexical tone production, for each phonetic parameter listed in Table 1, except slope 1 and slope 2, a mixed effect ANOVA was conducted with lexical tones (T1, T2, T3, T4) as the within-subject variable and group (autistic children or TD children) as between-subject variable. To compare the contours of T2 and T3, a mixed effect ANOVA was conducted with slopes 1 and 2 with lexical tones (T2 or T3) and slope (slope 1 or 2) as within-subject variables, and group (autistic children or TD children) as between-subject variable. The statistical analyses were conducted in SPSS 22.

To understand how individual participants may differ in their lexical tone production, we plotted the time normalized f0 contour of each lexical tone of each individual participant. For each individual production, 10 equally distanced f0 values (i.e., the f0 value at 10%, 20%, 30%, etc. of the duration of the tonal part of one syllable) were extracted from the f0 contour. Then for each individual participant, for each tone, for each of the 10 points, a mean f0 value was calculated by averaging the f0 values of all the productions. Finally, for each individual lexical tone, a tonal contour was generated by connecting the mean f0 values of the 10 points.

Results and discussion

The two judges gave highly consistent assessment for the lexical tones produced by the children, with an overall Cronbach's alpha of 0.82 when the two groups were pooled together, and a Cronbach's alpha of 0.88 for the autistic children. The assessment of Judge 1 was used for further analysis. For each lexical tone and for each participant, the accuracy was calculated as the percentage of productions rated as 'accurate'. Table 2 lists the mean accuracy of the lexical tones based on the perception of Judge 1. A mixed-effect analysis of variance (ANOVA) with lexical tones as the within-subject variable and group as the between subject variable found significant main effect of lexical tones, $F(3, 72) = 3.87$, $p = 0.02$, partial $\eta^2 = 0.14$ as well as group,

TABLE 1 Definition of the acoustical parameters analysed in the current study

Parameters	Definition	Predictions
Tone duration (s)	Duration of the lexical tones	T3 was expected to be longer than the other tones, and T4 shorter than the other tones (Xu, 1997)
f0 range (Hz)	f0 max–f0 min	Level tone (T1) was expected to have smaller f0 range than the contour tones (T2–T4)
Height of min f0 (Hz)	Mean min f0 of each of the four lexical tones – mean f0 across all the lexical tones by the speaker	High tone (T1) was expected to have higher height of min f0 than low tone (T3)
Height of max f0 (Hz)	Mean max f0 of each of the four lexical tones – mean f0 across all the lexical tones by the speaker	High tone (T1), rising tone (T2) and falling tone (T4) were expected to have higher height of max f0 than low tone (T3)
Directional excursion (Hz)	f0 max–f0 min, positive if L precedes H, negative if the opposite	Rising tone (T2) was expected to have positive directional excursion, and falling tone (T4) was expected to have negative directional excursion
Slope 1 (Hz/s) (for T2 and T3 only)	Absolute value of directional excursion/duration between H1 and L	T3 was expected to have a steeper slope than T2
Slope 2 (Hz/s) (for T2 and T3 only)	Absolute value of directional excursion/duration between L and H2	T2 was expected to have a steeper slope than T3
Min f0 alignment	Duration between lexical tone onset and L/duration of the lexical tone	Rising tone (T2) was expected to have earlier min f0 alignment than falling tone (F4), and the low tone (T3) was expected to have later min f0 alignment than the rising tone (T2) but earlier min f0 alignment than the falling tone (T4)
Max f0 alignment	Duration between lexical tone onset and H/duration of lexical tone	Rising tone (T2) was expected to have later max f0 alignment than falling tone (F4)

$F(1, 24) = 7.36$, $p = 0.01$, partial $\eta^2 = 0.24$, but the interaction between the two was not significant, $F(3, 72) = 1.92$, $p = 0.13$. Bonferroni-corrected post-hoc analysis found no significant pair wise differences between the lexical tones among the TD children, while for the autistic children, the accuracy of T3 was significantly lower than T1 ($p = 0.03$). As can be seen from Table 2, the lexical tones produced by the TD children were perceived as highly accurate. The least accurate lexical tone was T3, yet it still had an accuracy of 96%. For the autistic children, although most of their lexical tones were perceived as accurate, they still had a lower accuracy than the TD children. Similar to the TD children, they produced T3 with an accuracy much lower than the other three lexical tones, which was possibly due to the complex tonal contour of T3. In addition, the autistic children showed larger individual variation than the TD children.

With all the lexical tones pooled together, a repeated-measures ANOVA was conducted with the accuracy of individual participants, with number of syllables (monosyllabic, bisyllabic) being the within-group variable and group (TD, autism) being the between-group variable. We found that number of syllables showed a significant main effect, $F(1, 24) = 7.30$, $p = 0.012$, partial $\eta^2 = 0.23$, and lexical tones produced in bisyllabic words were perceived to be more accurate than those produced in monosyllabic words. The main effect of group was significant as well, $F(1, 24) = 6.86$, $p = 0.015$, partial $\eta^2 = 0.22$, with the TD children having a higher accuracy. The interaction between the two factors was not significant, $F(1, 24) = 2.63$, $p = 0.12$. Therefore, the children's tone production seemed to be facilitated when the lexical tones occurred in bisyllabic words with the two lexical tones being compatible. It is possible that the judges found lexical tones easier to perceive in bisyllabic words due to the additional context.

For the acoustic analyses, Table 3 lists the mean tone duration, f0 range, height of min f0, height of max f0, directional excursion, slope 1, slope 2, min f0 alignment and max f0 alignment of the autistic children and the TD children.

For duration, lexical tone showed a significant main effect, $F(3, 72) = 85.59$, $p < 0.001$, partial $\eta^2 = 0.78$, group showed no significant main effect, $F(1, 24) = 1.78$, $p = 0.20$, and the interaction between the two was significant, $F(3, 72) = 5.34$, $p = 0.002$, partial $\eta^2 = 0.18$. Bonferroni corrected post-hoc analysis showed that T1 was significantly shorter than T2 and T3, but significant longer than T4; T2 was significantly shorter than T3, but longer than T4; T3 was significantly longer than all the other lexical tones; T4 was significantly shorter than all the other tones. Such durational differences between the lexical tones were consistent with previous studies, namely that T3 was the longest while T4 the shortest (Xu, 1997). The

TABLE 2 Mean accuracy (SD) of the lexical tones produced by the typically developing (TD) children and the autistic children

	TD children	Autistic children
T1	100% (0)	94% (17%)
T2	100% (0)	86% (26%)
T3	96% (12%)	67% (42%)
T4	99% (4%)	81% (28%)

TABLE 3 Mean (SD) of the acoustic measurements of the typically developing (TD) children and autistic children

		TD children	Autistic children
Lexical tone duration	T1	0.29 (0.08)	0.25 (0.09)
	T2	0.33 (0.08)	0.28 (0.08)
	T3	0.42 (0.10)	0.34 (0.12)
	T4	0.23 (0.05)	0.23 (0.08)
Pitch range (Hz)	T1	11.03 (3.35)	23.53 (17.05)
	T2	35.28 (11.68)	47.49 (25.10)
	T3	50.93 (19.16)	69.55 (55.96)
	T4	63.15 (32.58)	64.97 (44.71)
Height of min f0 (Hz)	T1	18.31 (8.55)	0.39 (15.55)
	T2	−30.93 (9.37)	−36.15 (18.89)
	T3	−63.66 (25.99)	−64.66 (28.60)
	T4	−11.18 (26.62)	−7.09 (29.67)
Height of max f0 (Hz)	T1	29.23 (11.13)	23.73 (15.23)
	T2	4.33 (6.05)	11.33 (10.04)
	T3	−13.17 (13.03)	4.89 (47.40)
	T4	50.99 (16.97)	56.92 (29.08)
Directional excursion	T1	−1.87 (5.02)	−5.70 (16.98)
	T2	29.99 (12.69)	31.82 (25.81)
	T3	−23.83 (28.57)	−14.68 (32.89)
	T4	−51.49 (32.30)	−51.11 (52.34)
Slope 1	T2	155.44 (62.37)	276.71 (304.14)
	T3	292.53 (109.49)	640.87 (714.59)
Slope 2	T2	205.49 (72.81)	299.78 (129.26)
	T3	382.60 (193.14)	407.33 (481.45)
Min f0 alignment	T1	0.52 (0.07)	0.56 (0.16)
	T2	0.36 (0.06)	0.37 (0.11)
	T3	0.56 (0.08)	0.53 (0.16)
	T4	0.73 (0.06)	0.65 (0.15)
Max f0 alignment	T1	0.43 (0.12)	0.36 (0.17)
	T2	0.76 (0.11)	0.66 (0.17)
	T3	0.33 (0.23)	0.39 (0.15)
	T4	0.28 (0.06)	0.32 (0.13)

autistic children produced shorter lexical tones than the TD children for T1–T3, but their T4s had a comparable duration to those produced by TD children.

For pitch range, lexical tone showed a significant main effect, $F(3, 72) = 16.78$, $p < 0.001$, partial $\eta^2 = 0.41$, group showed no significant main effect, $F(1, 24) = 1.97$, $p = 0.17$, and the interaction between the two was not significant, $F(3, 72) = 0.44$, $p = 0.72$. Bonferroni corrected post-hoc analysis showed that T1 had a smaller pitch range than T2, T3 and T4; T2 had a smaller pitch range than T4, and none of the other pair-wise comparison turned to be significant. These results indicated that both the autistic children and the TD children produced T1 with a flatter contour than the other lexical tones, and both groups produced T4 with a steep falling contour. Therefore, both groups of children captured the different pitch contours of the lexical tones.

For pitch height of min f0, lexical tone showed a significant main effect, $F(3, 72) = 59.88$, $p < 0.001$, partial $\eta^2 = 0.72$, group showed no significant main effect, $F(1, 24) = 1.09$, $p = 0.31$, and the interaction between the two was not significant, $F(1, 24) = 1.31$, $p = 0.28$. Bonferroni corrected post-hoc analysis showed that T1 had a significantly higher min f0 height than T2–T4; T2 had a significantly higher min f0 height than T3, but significantly lower min f0 height than T4; T3 had significantly lower min f0 height than all the other three lexical tones. These findings indicate that for both groups of children, T1 had a higher min f0 level than the other three lexical tones, which was in accordance with its defining 'high' feature, while T3 had a lower min f0 level than the other lexical tones, consistent with its 'low' feature (Duanmu, 2006). In other words, both groups of children were able to distinguish T1 and T3 by using the 'high' and 'low' acoustic features correctly.

For height of max f0, lexical tone showed a significant main effect, $F(3, 72) = 31.00$, $p < 0.001$, partial $\eta^2 = 0.56$, group showed no significant main effect, $F(1, 24) = 2.57$, $p = 0.12$, the interaction between the two was not significant, $F(3, 72) = 1.12$, $p = 0.35$. Bonferroni post-hoc analysis showed that the height of max f0 of T1 was significantly higher than that of T2 and T3, but significantly lower than T4; the height of max f0 of T2 was significantly lower than T4, the height of max f0 of T3 was significantly lower than T1 and T4. The height of max f0 of T2 was not significantly different from that of T3. These findings indicate that for both the autistic children and the TD children, T2 and T3 had lower max f0s than T1 and T4, and T4 has the highest max f0. Therefore, both groups of children faithfully produced the 'high' feature of T1 and T4.

For directional excursion, lexical tone showed a significant main effect, $F(3, 72) = 24.07$, $p < 0.001$, partial $\eta^2 = 0.50$, group showed no significant main effect, $F(1, 24) = 0.07$, $p = 0.79$, and the interaction between the two was not significant, $F(3, 72) = 0.09$, $p = 0.96$. As can be seen from Table 3, T1 had the smallest directional excursion among the four lexical tones. For both groups, T1 directional excursion was not significantly different from zero, $t_{TD}(12) = -1.34$, $p = 0.20$ (two tailed), $t_{ASD}(12) = -1.21$, $p = 0.25$ (two tailed). These findings indicate that, both TD children and autistic children produced T1 with a flat contour. For both groups, T2 had a positive directional excursion hence a rising contour, while T4 had a negative directional excursion, hence a falling contour. For T3, the results showed that the onset of T3 had a higher f0 than the offset, hence the max f0 occurred earlier than min f0, leading to negative directional excursion for T3. Therefore, both groups of children successfully produced the 'level', 'rising' and 'falling' feature of T1, T2 and T4, respectively.

For min f0 alignment, lexical tone showed a significant main effect, $F(3, 72) = 40.23$, $p < 0.001$, partial $\eta^2 = 0.63$, and group showed no significant main effect, $F(1, 24) = 0.31$, $p = 0.58$, and the interaction between the two was not significant, $F(3, 72) = 1.67$, $p = 0.18$. Bonferroni corrected post hoc analysis showed that T1 had a later min f0 alignment than T2, and an earlier min f0 alignment than T4; min f0 alignment between T1 and T3 was not significant. T2 had an earlier min f0 than T3 and T4, and T3 had earlier min f0 alignment than T4. These findings indicate that for both groups of children, T4 had the latest min f0, which was consistent with its falling contour. In addition, both groups of children produced T2 with an earlier min f0 (i.e., an earlier turning point) than T3, and successfully distinguished between T2 and T3 in terms of turning point, consistent with findings among the adult native speakers (Moore & Jongman, 1997; Xu, 1997).

For max f0 alignment, lexical tone showed a significant main effect, $F(3, 72) = 44.53$, $p < 0.001$, partial $\eta^2 = 0.65$, group showed no significant main effect, $F(1, 24) = 0.36$, $p = 0.56$, and interaction between the two was not significant, $F(3, 72) = 2.03$, $p = 0.12$. Bonferroni corrected post hoc analysis showed that T1 had an earlier max f0 than T2, and T2 had a later max f0 than T3 and T4. None of the other pair-wise comparisons turned out to be significant. These findings indicate that, for both the TD children and the autistic children, T2 had a rising contour, leading to its late max f0 than the other lexical tones.

For slopes 1 and 2, repeated-measures ANOVA with lexical tone (T2 or T3) and slope (slope 1 or 2) as within subject variable and group as between subject variable showed that lexical tone had a significant main effect, $F(1, 24) = 10.67$, $p = 0.003$, partial $\eta^2 = 0.31$. The main effect of slope was not significant, $F(1, 24) = 0.14$, $p = 0.71$, and neither was the main effect of group, $F(1, 24) = 2.16$, $p = 0.16$. The interaction between slope and group was marginally significant, $F(1, 24) = 3.51$, $p = 0.073$. For both groups, T3 had a larger slopes 1 and 2 than T2. The larger slopes of T3 were consistent with its dipping contour. Bonferroni corrected post hoc analysis indicated that the TD children had a larger slope 2 than slope 1 while the opposite was true for autistic children, yet such differences were not statistically significant.

The pitch contours of each lexical tone produced by each participant are plotted in Figure 1. As can be seen, for all four lexical tones, autistic children showed larger individual variation than the TD children. For the TD children, the tonal contours of different children clustered tightly, indicating consistent tonal production across the

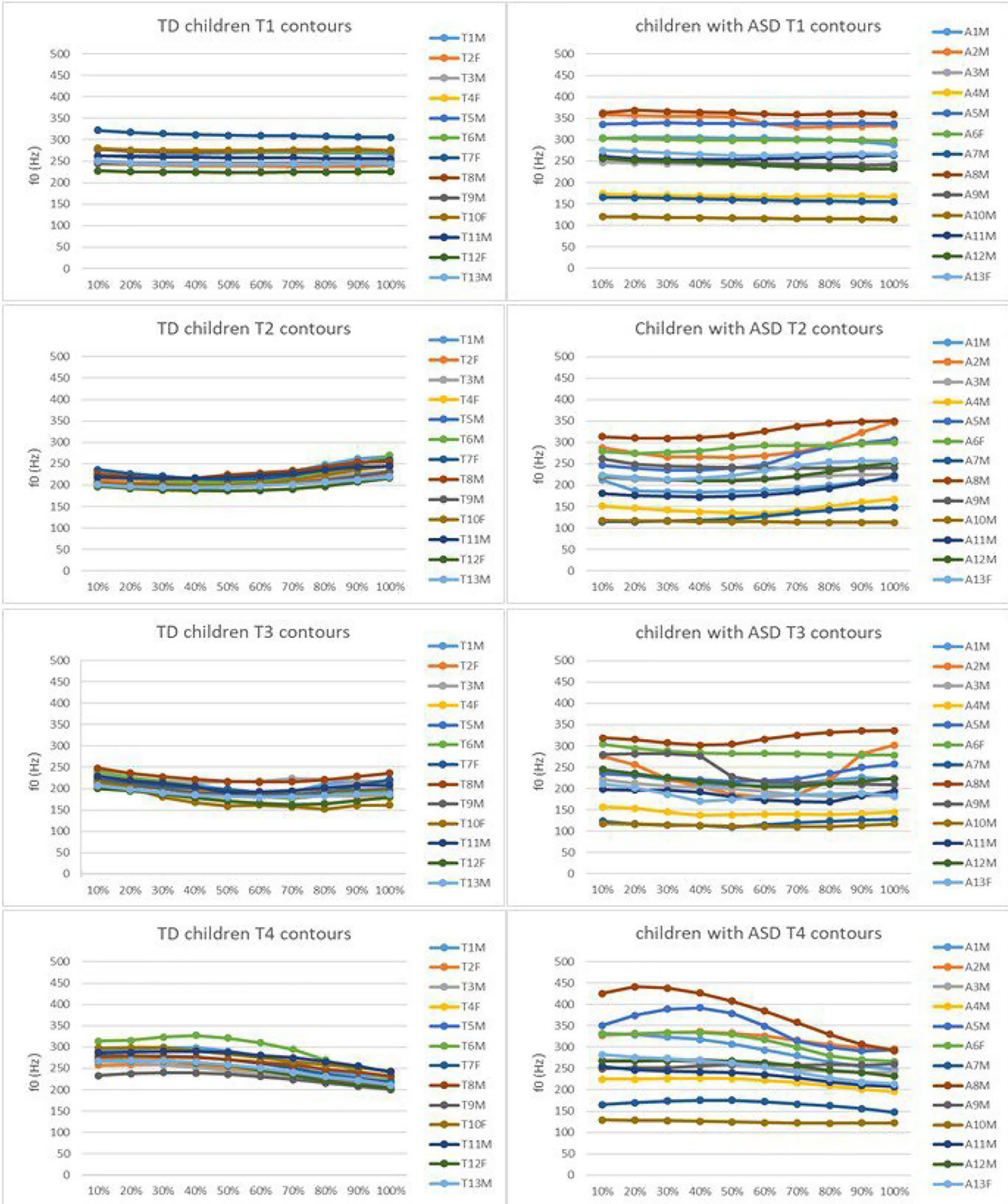

FIGURE 1 The f0 contours of T1–T4 of individual children. Each colour represents an individual child.
Note: The *x*-axis represents the total normalized duration of the lexical tones. T, typically developing children; A, autistic children; M, male; and F, female.
[Colour figure can be viewed at wileyonlinelibrary.com]

participants. In other words, by about 10 years of age, among the TD children, regardless of who the speaker is, for each tone, the f0 register and contour are highly stable. Although the autistic children produced the lexical tones with similar global contours compared with the TD children, their productions were far less convergent than the TD children. Therefore, it seems that the autistic children were able manipulate their articulators in a way that the global f0 contours of the lexical tones can be produced, yet their f0 register and range were heterogeneous.

To summarize, the majority of the lexical tones produced by the autistic children were perceived to be accurate by adult judges, yet they still had lower accuracies compared with their TD peers. The phonetic analysis of the lexical tones failed to find any significant difference between the two groups for all the pitch measures. Duration wise, however, the autistic children produced T1–T3 with a shorter duration than their TD peers. In addition, the two groups of children produced the lexical tones more accurately in the bisyllabic words.

GENERAL DISCUSSION

In the current study, lexical tones were elicited with a picture-naming task from TD children and autistic children with intellectual impairment, and these were both perceptually and phonetically analysed. We found that for the perceptual judgement, the majority of the lexical tones produced by the autistic children were perceived to be accurate, yet they still had lower accuracies than their TD counterparts. Phonetic analyses with f0 showed that, for all the measurements, namely f0 range, height of max f0, height of min f0, directional excursion, alignment of min f0 and alignment of max f0, no significant difference was found between the TD children and autistic children, and both groups were able to phonetically differentiate the lexical tones. Importantly, there was no significant interaction between group and the lexical tones. In other words, there was no evidence that the autistic children used the phonetic features differently than the TD children when differentiating the lexical tones in production. Our findings are consistent with Xu (2013), who found no difference between Chinese children with HFA/AS and TD children in the phonetic analysis of lexical tone production. Since in Xu (2013) the participants were older than the children tested in the current study, and the lexical tones were elicited in carrier sentences, it would be informative for future studies to compare verbal autistic children to autistic children with intellectual impairment, and TD children of the same age on tasks targeting different functions of pitch, and investigate how these groups differ in terms of pitch manipulation.

Lexical tones are different from sentential prosody in that its production and perception do not require an understanding of the interlocutor's intention. The Mandarin Chinese lexical tones distinguish word meaning, and they maintain their pitch contours regardless of intonation variation (Chao, 1933; Xu, 1997). Our results showed that when prosody (f0 in the current study) served no communicative purpose, the autistic children were able to produce lexical prosody in a way comparable with TD children, even when these lexical tones had complex contours and required precise manipulation of the vocal tract in the duration of one syllable (e.g., change in pitch direction in T2, T3 and T4). In particular, the autistic children tested in the current study were intellectually impaired and minimally verbal. Nevertheless, the severely impaired language ability did not impede them from learning the lexical tones. Therefore, the prosody deficits commonly observed among autistic children cannot be solely attributed to general impairment sensory–motoric processing (e.g., Kujala et al., 2010, 2013), at least not for pitch processing. Instead, it seems likely that their impairment in social communication prevents autistic children from fully acquiring social–pragmatically oriented sentential prosody. It would be interesting for future studies to investigate whether disordered perceptual response to pitch still exhibits among those autistic children who are able to produce the lexical tones correctly.

Nevertheless, the autistic children were still different from the TD children in lexical tone production. First, the autistic children produced the lexical tones less accurately than the TD children. As can be seen, by late childhood, the TD children barely made any errors in their native lexical tone production. In contrast, the autistic children were commensurate with the 3-year-olds tested in Wong (2012). Second, much larger individual variation was observed among the autistic children. For typical language development, larger individual variation is commonly found among younger rather than older children (Wong, 2012; Xu et al., 2018). It seems that the autistic children were at the early stage of lexical tone production. Given their age, it is unlikely that they will reach the consistency of TD children as they grew older. Although most of the autistic children were able to produce the global contours of the lexical tones, some showed idiosyncratic patterns. For example, as can be seen in Figure 1, participant A10M produced all the lexical tones with a flat contour, and the perceptual assessment showed that all his T1s were accurate, while all the other tones were inaccurate except one T2 and one T4. Therefore, although as a group, the autistic children were able to make use of the same phonetic features as the TD children when differentiating the lexical tones, some individuals may be less capable of producing f0 variations. It would be interesting for future studies to investigate whether the autistic children who are unable to produce pitch variations form a separate subgroup compared with those who can.

Compared with the findings in the systematic review of Fusaroli et al. (2017), the current study confirmed the larger pitch variability of the autistic children. The participants in the current study were complementary to those reported in Fusaroli et al. (2017). The children in the current study were low functioning, while the vast majority in Fusaroli et al. (2017) were children with HFA or AS. Therefore, it seems that larger pitch variability was consistent across different subtypes of autistic children. This is not surprising given the heterogeneity of autism. However, we failed to find higher pitch mean of these children as shown in Fusaroli et al. (2017). The discrepancy might be due to participant characteristics, testing procedure, the stimuli or combination of all. Nevertheless, it does not seem that phonetic realization of pitch can be a reliable marker of autism. Therefore, to understand how speech prosody characterizes the autistic children, it is crucial for future studies to compare autistic children that differ in severity on tasks targeting different functions of pitch. Interestingly, we found that the autistic children exhibited different durational patterns of the lexical tones compared with the TD children. For the TD children, T4 was the shortest and T3 the longest, which was consistent with previous findings (Xu, 1997). Autistic children, on the other hand, seemed to produce the lexical tones with reduced durational differences. Some previous studies found disordered duration in autistic children's production (Diehl et al., 2013; Grossman et al., 2010) while others did not (Diehl et al. 2012; Nadig & Shaw, 2012). Our findings seem to suggest that the autistic children might be less able to integrate durational and pitch cues than the TD children.

It should be acknowledged that the autistic children tested in the current study had both autism and intellectual impairment. To better disentangle the influence of the two, it would be useful to compare intellectually impaired children with and without autism in the future. Furthermore, it would be informative if phonetic analysis could be conducted with intonational and lexical pitch patterns produced by the same autistic children, and to investigate whether the deviation, if any, at the lexical and sentential level is of the same kind. As the children tested in the currently study were minimally verbal, it was not possible to test their sentential prosody.

In sum, although the autistic children produced the lexical tones less accurately than the TD children, the majority of the lexical tones produced by the autistic children were accurate. They were able to produce the global contours of the lexical tones and they made use of the phonetic features in comparable ways with the TD children. Therefore, disordered pitch production does not seem to qualify for a core deficit among autistic children, at least not when it serves no social pragmatic purpose.

ACKNOWLEDGEMENTS

We are deeply grateful for all the children and teachers who participated in this study. I thank the reviewers for their very helpful suggestions.

The research for this article was supported by Beijing Social Science Fund with the grant number 18YYC013, the National Social Science Fund of China with the grant number 17CYY041, the Chinese Academy of Social Sciences Innovation Project, and the Fundamental Research Funds for the Central Universities, with the approval number 21PT01. I am deeply grateful for all the children and teachers that participated in this study. I thank all the reviewers for their very helpful suggestions.

CONFLICT OF INTEREST STATEMENT

The authors claim no conflict of interest. The research was approved by Ethic Committee for Research in Beijing Language and Culture University.

DATA AVAILABILITY STATEMENT

The data are available from the authors on request.

ORCID

Ao Chen https://orcid.org/0000-0002-9741-9179

REFERENCES

Arciuli, J. & Benjamin, B. (2019) An acoustic study of lexical stress contrastivity in children with and without autism spectrum disorders. *Journal of Child Language*, 46(1), 142–152. https://doi.org/10.1017/S0305000918000272

Arciuli, J., Colombo, L. & Surian, L. (2020) Lexical stress contrastivity in Italian children with autism spectrum disorders: an exploratory acoustic study. *Journal of Child Language*, 47(4), 870–880. https://doi.org/10.1017/S0305000919000795

Baltaxe, C.A.M. & Simmons, J.Q. (1985) Prosodic development in normal and autistic children. In: Schopler, E., Mesibov, G.B. (Eds.). *Communication problems in Autism*. Current Issues in Autism. Boston, MA: Springer. https://doi.org/10.1007/978-1-4757-4806-2_7

Baltaxe, C.A.M. & Guthrie, D. (1987) The use of primary sentence stress by normal, aphasic, and autistic children. *Journal of Autism and Developmental Disorders*, 17(2), 255–271. https://doi.org/10.1007/BF01495060

Banse, R. & Scherer, K.R. (1996) Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3), 614–636. https://doi.org/10.1037/0022-3514.70.3.614

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

Boucher, J. (2012) Research review: structural language in autistic spectrum disorder – characteristics and causes. *Journal of Child Psychology and Psychiatry*, 53(3), 219–233. https://doi.org/10.1111/j.1469-7610.2011.02508.x

Brisson, J., Martel, K., Serres, J., Sirois, S. & Adrien, J.-L. (2014) acoustic analysis of oral productions of infants later diagnosed with autism and their mother. *Infant Mental Health Journal*, 35(3), 285–295. https://doi.org/10.1002/imhj.21442

Chao, Y.R. (1933) Tone and intonation in Chinese. *Bulletin of the Institute of History and Philology, Academia Sinica*, 4, 121–134.

Chen, F., Cheung, C.C. & Peng, G. (2022) Linguistic tone and non-linguistic pitch imitation in children with autism spectrum disorders: a cross-linguistic investigation. *Journal of Autism and Developmental Disorders*, 52(5), 2325–2343. https://doi.org/10.1007/s10803-021-05123-4

Cleland, J., Gibbon, F.E., Peppé, S.J.E., O'Hare, A. & Rutherford, M. (2010) Phonetic and phonological errors in children with high functioning autism and Asperger syndrome. *International Journal of Speech–Language Pathology*, 12(1), 69–76. https://doi.org/10.3109/17549500903469980

Cutler, A. & Van Donselaar, W. (2001) Voornaam is not (really) a homophone: lexical prosody and lexical access in Dutch.

Language and Speech, 44(2), 171–195. https://doi.org/10.1177/00238309010440020301

DeMyer, M.K., Barton, S., DeMyer, W.E., Norton, J.A., Allen, J. & Steele, R. (1973) Prognosis in autism: a follow-up study. *Journal of Autism & Childhood Schizophrenia*, 3(3), 199–246. https://doi.org/10.1007/BF01538281

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

Diehl, J. J., & Paul, R. (2012). Acoustic differences in the imitation of prosodic patterns in children with autism spectrum disorders. In Research in Autism Spectrum Disorders (Vol. 6, Issue 1, pp. 123–134). https://doi.org/10.1016/j.rasd.2011.03.012

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(1), 135–161. https://doi.org/10.1017/S0142716411000646

Duanmu, S. (2006) Chinese (Mandarin): phonology. *Encyclopedia of language and linguistics*, 2nd edn. ed. by Brown, K., 351–355. Oxford, UK: Elsevier Publishing House.

Fosnot, S.M. & Jun, S.A. (1999) Prosodic characteristics in children with stuttering or autism during reading and imitation. *International Congress of Phonetic Sciences*, 1925–1928.

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*, 10(3), 384–407. https://doi.org/10.1002/aur.1678

Goldin, R.L., Matson, J.L. & Cervantes, P.E. (2014) The effect of intellectual disability on the presence of comorbid symptoms in children and adolescents with autism spectrum disorder. *Research in Autism Spectrum Disorders*, 8(11), 1552–1556. https://doi.org/10.1016/j.rasd.2014.08.006

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

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. https://doi.org/10.1080/17549500903003060

Grossman Ruth, B., Bemis Rhyannon, H., Plesa, S.D. & Helen, T.-F. (2010) Lexical and affective prosody in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research*, 53(3), 778–793. https://doi.org/10.1044/1092-4388(2009/08-0127)

Gussenhoven, C. (2004) *The phonology of tone and intonation*. New York: Cambridge University Press.

Halliday, M.A.K. (1970) *A course in spoken English: intonation*. London: Oxford University Press.

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

Heaton, P. (2003) Pitch memory, labelling and disembedding in autism. *Journal of Child Psychology and Psychiatry*, 44(4), 543–551. https://doi.org/10.1111/1469-7610.00143

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

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

Howie, J.M. (1976) *Acoustical studies of Mandarin vowels and tones*. New York: Cambridge University Press.

Hubbard, K. & Trauner, D.A. (2007) Intonation and emotion in autistic spectrum disorders. *Journal of Psycholinguistic Research*, 36(2), 159–173. https://doi.org/10.1007/s10936-006-9037-4

Jansson-Verkasalo, E., Kujala, T., Jussila, K., Mattila, M.L., Moilanen, I., Näätänen, R., et al. (2005) Similarities in the phenotype of the auditory neural substrate in children with Asperger syndrome and their parents. *European Journal of Neuroscience*, 22(4), 986–990. https://doi.org/10.1111/j.1460-9568.2005.04216.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

Kanner, L. (1971) Follow-up study of eleven autistic children originally reported in 1943. *Journal of Autism and Childhood Schizophrenia*, 1(2), 119–145. https://doi.org/10.1007/BF01537953

Kjelgaard, M.M. & Tager-Flusberg, H. (2001) An investigation of language impairment in autism: implications for genetic subgroups. *Language and Cognitive Processes*, 16(2–3), 287–308. https://doi.org/10.1080/01690960042000058

Krug, D.A., Arick, J. & Almond, P. (1980) Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. In: *Child psychology & psychiatry & allied disciplines*, Vols. 21, Issue 3. pp. 221–229. https://doi.org/10.1111/j.1469-7610.1980.tb01797.x

Kujala, T., Kuuluvainen, S., Saalasti, S., Jansson-Verkasalo, E., von Wendt, L. & Lepistö, T. (2010) Speech-feature discrimination in children with Asperger syndrome as determined with the multi-feature mismatch negativity paradigm. In *Clinical neurophysiology*, Vols. 121, Issue 9. pp. 1410–1419. https://doi.org/10.1016/j.clinph.2010.03.017

Kujala, T., Lepistö, T. & Näätänen, R. (2013) The neural basis of aberrant speech and audition in autism spectrum disorders. *Neuroscience & Biobehavioral Reviews*, 37(4), 697–704. https://doi.org/10.1016/j.neubiorev.2013.01.006

Lau, J.C.Y., To, C.K.S., Kwan, J.S.K., Kang, X., Losh, M. & Wong, P.C.M. (2021) Lifelong tone language experience does not eliminate deficits in neural encoding of pitch in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 51(9), 3291–3310. https://doi.org/10.1007/s10803-020-04796-7

Lord, C., Brugha, T.S., Charman, T., Cusack, J., Dumas, G., Frazier, T., et al. (2020) Autism spectrum disorder. *Nature Reviews Disease Primers*, 6(1), 5. https://doi.org/10.1038/s41572-019-0138-4

McAlpine, A., Plexico, L.W., Plumb, A.M. & Cleary, J. (2014) Prosody in young verbal children with autism spectrum disorder. *Contemporary Issues in Communication Science and Disorders*, 41, 120–132.

McCann, J. & Peppé, S. (2003) Prosody in autism spectrum disorders: a critical review. *International Journal of Language & Communication Disorders*, 38(4), 325–350. https://doi.org/10.1080/1368282031000154204

Moore, C.B. & Jongman, A. (1997) Speaker normalization in the perception of Mandarin Chinese tones. *The Journal of the Acoustical Society of America*, 102(3), 1864–1877. http://link.aip.org/link/?JAS/102/1864/1

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

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 and Development*, 36(6), 516–522. https://doi.org/10.1016/j.braindev.2013.07.006

Oller, D.K., Niyogi, P., Gray, S., Richards, J.A., Gilkerson, J., Xu, D., et al. (2010) Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences*, 107(30), 13354–13359. https://doi.org/10.1073/pnas.1003882107

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

Peppé, S., Cleland, J., Gibbon, F., O'Hare, A. & Castilla, P.M. (2011) Expressive prosody in children with autism spectrum conditions. *Journal of Neurolinguistics*, 24(1), 41–53. https://doi.org/10.1016/j.jneuroling.2010.07.005

Plaisted, K.C. (2001) Reduced generalisation in autism: An alternative to weak central coherence. In: Burack, J.A., & Charman, T. (Eds.), *The development of autism: Perspectives from theory and research* (pp. 139–169). Mahwah, NJ: Lawrence Erlbaum Associates.

Rapin, I., Dunn, M.A., Allen, D.A., Stevens, M.C. & Fein, D. (2009) Subtypes of language disorders in school-age children with autism. *Developmental Neuropsychology*, 34(1), 66–84. https://doi.org/10.1080/87565640802564648

Robins, D. L., Casagrande, K., Barton, M., Chen, C.-M. A., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the Modified Checklist for Autism in Toddlers, Revised With Follow-up (M-CHAT-R/F). Pediatrics, 133(1), 37–45. https://doi.org/10.1542/peds.2013-1813

Russo, N.M., Skoe, E., Trommer, B., Nicol, T., Zecker, S., Bradlow, A., et al. (2008) Deficient brainstem encoding of pitch in children with autism spectrum disorders. *Clinical Neurophysiology*, 119(8), 1720–1731. https://doi.org/10.1016/j.clinph.2008.01.108

Rutter, M. & Lockyer, L. (1967) A five to fifteen year follow-up study of infantile psychosis: I. Description of sample. *British Journal of Psychiatry*, 113(504), 1169–1182. https://doi.org/10.1192/bjp.113.504.1169

Schopler, E., Reichler, R.J., DeVellis, R.F., et al. (1980) Toward objective classification of childhood autism: childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders*, 10, 91–103. https://doi.org/10.1007/BF02408436

Shriberg, L.D., Paul, R., Black, L.M. & van Santen, P., J. (2011) The hypothesis of apraxia of speech in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 41(4), 405–426. https://doi.org/10.1007/s10803-010-1117-5

Shriberg Lawrence, D., Rhea, P., McSweeny Jane, L., Ami, K., Cohen Donald, J. & Volkmar Fred, R. (2001) Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research*, 44(5), 1097–1115. https://doi.org/10.1044/1092-4388(2001/087)

Simmons, J.Q. & Baltaxe, C. (1975) Language patterns of adolescent autistics. *Journal of Autism and Childhood Schizophrenia*, 5(4), 333–351. https://doi.org/10.1007/BF01540680

Van Santen, J.P., Prud'hommeaux, E.T., Black, L.M. & Mitchell, M. (2010) Computational prosodic markers for autism. *Autism: The International Journal of Research and Practice*, 14(3), 215–236. https://doi.org/10.1177/1362361309363281

Whalen, D.H. & Xu, Y. (1992) Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, 49(1), 25–47. http://www.scopus.com/inward/record.url?eid=2-s2.0-0026470422&partnerID=40&md5=973bbc0e8cdda88479fd859ea69ca78a

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*, 74, 1065–1079.

Wong, P. (2012) Acoustic characteristics of three-year-olds. correct and incorrect monosyllabic Mandarin lexical tone productions. In Journal of Phonetics (Vol. 40, Issue 1, pp. 141–151). https://doi.org/10.1016/j.wocn.2011.10.005

Wu, H., Lu, F., Yu, B. & Liu, Q. (2020) Phonological acquisition and development in Putonghua-speaking children with autism spectrum disorders. *Clinical Linguistics & Phonetics*, 34(9), 844–860. https://doi.org/10.1080/02699206.2019.1702720

Xu, J. (2013) Acoustic features of lexical tones and intonation of Taiwan Mandarin-speaking adolescents with autism. Master thesis, National Taipei University of Nursing and Health sciences.

Xu, Y. (2011) Post-focus compression: cross-linguistic distribution and historical origin. *Proceedings of ICPhS XVII*, 152–155.

Xu, Y. (1997) Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25(1), 61–83.

Xu Rattanasone, N., Tang, P., Yuen, I., Gao, L., & Demuth, K. (2018). Five-Year-olds' Acoustic Realization ofMandarin Tone Sandhi and Lexical Tones in Context Are Not Yet Fully Adult-Like. *Frontiers in Psychology*, 9, 817. https://doi.org/10.3389/fpsyg.2018.00817

Zhu, H. & Dodd, B. (2000) The phonological acquisition of Putonghua (modern standard Chinese). *Journal of Child Language*, 27(01), 3. https://doi.org/10.1017/S0305000099900402X

How to cite this article: Chen, A., Zhao, R., Huang, G., Li, A. & Cheung, H. (2023) Successful lexical tone production of Mandarin Chinese autistic children with intellectual impairment. *International Journal of Language & Communication Disorders*, 1–15. https://doi.org/10.1111/1460-6984.12881

APPENDIX

TABLE A1 Target words used in the experiment

Target words	Pinyin	Lexical tones	IPA
刀	dao	T1	tɑu
花	hua	T1	xuɑ
球	qiu	T2	tɕʰiou
门	men	T2	mən
羊	yang	T2	iɑŋ
船	chuan	T2	tʂʰuan
鞋	xie	T2	ɕiɛ
嘴	zui	T3	tsuei
鸟	niao	T3	niɑu
马	ma	T3	mĄ
脚	jiao	T3	tɕiɑu
飞机	feiji	T1T1	feitɕi
西瓜	xigua	T1T1	ɕiguɑ
书包	shubao	T1T1	ʂubɑu
香蕉	xiangjiao	T1T1	ɕiɑŋtɕiɑu
熊猫	xiongmao	T2T1	ɕyŋmɑu
楼梯	louti	T2T1	loutʰi
圆圈	yuanquan	T2T1	yæntɕʰyæn
木门	mumen	T4T2	mumən
热水	reshui	T4T2	ʐɣʂuei
赛跑	saipao	T4T3	saipɑu
蔬菜	shucai	T1T4	ʂutsʰai
头发	toufa	T2T4	toufĄ
螃蟹	pangxie	T2T4	pɑŋɕiɛ
肥肉	feirou	T2T4	feiʐou

[This paper was International Journal of Language and Communication Disorders,00,1–15]