



Perception and Production of Statement-Question Intonation in Autism Spectrum Disorder: A Developmental Investigation

Li Wang¹ · C. Philip Beaman¹ · Cunmei Jiang² · Fang Liu¹

Accepted: 25 July 2021 / Published online: 5 August 2021
© The Author(s) 2021

Abstract

Prosody or “melody in speech” in autism spectrum disorder (ASD) is often perceived as atypical. This study examined perception and production of statements and questions in 84 children, adolescents and adults with and without ASD, as well as participants’ pitch direction discrimination thresholds. The results suggested that the abilities to discriminate (in both speech and music conditions), identify, and imitate statement-question intonation were intact in individuals with ASD across age cohorts. Sensitivity to pitch direction predicted performance on intonation processing in both groups, who also exhibited similar developmental changes. These findings provide evidence for shared mechanisms in pitch processing between speech and music, as well as associations between low- and high-level pitch processing and between perception and production of pitch.

Keywords Autism spectrum disorder · Speech · Music · Intonation · Pitch

Introduction

Prosody is a suprasegmental feature of speech that adds additional pragmatic, affective, or grammatical information via changes in frequency, intensity, and duration of spoken utterances (McAlpine et al., 2014; McCann & Peppé, 2003; Paul et al., 2005). It plays an important role in speech communication and social interaction (Xu, 2005). While the acquisition of prosody starts from infancy (Levitt, 1993) and lays the foundations for children’s sociopragmatic development (Hübscher & Prieto, 2019), atypical prosody can become a barrier to everyday linguistic and social functioning, as seen in autism spectrum disorder (ASD) (Lloyd-Fox et al., 2013; McCann & Peppé, 2003; Paul et al., 2005).

ASD is a neurodevelopmental disorder associated with deficits in social communication and interaction as well as restricted and repetitive behaviours and interests (American Psychiatric Association, 2013). Prosodic deficits have been frequently observed in ASD across a variety of perception and production tasks (Diehl & Paul, 2012; Nakai et al.,

2014; Peppé et al., 2007; Shriberg et al., 2011; Tager-Flusberg et al., 2005). They can occur even in highly verbal individuals with ASD and tend to be lifelong even when other areas of language, such as semantics and syntax, improve (McCann & Peppé, 2003). Among the different areas of prosody, recognising and differentiating the rising from falling intonation in questions and statements represents an important aspect of conversational and linguistic competence (Dahan, 2015; Xie et al., 2021), and the literature in ASD has produced mixed findings (Chevallier et al., 2009; Filipe et al., 2014; Järvinen-Pasley et al., 2008a; Jiang et al., 2015; McCann & Peppé, 2003; McCann et al., 2007; Paul et al., 2005; Peppé et al., 2007). The current study investigated this issue by examining the roles of response bias, stimulus type, age, and pitch discrimination thresholds in the perception and production of statement-question intonation in ASD.

Perception of Statement-Question Intonation and Response Bias in ASD

Several studies used the same test battery, PEPS-C (Profiling Elements of Prosodic Systems-Children) (Peppé & McCann, 2003), to examine discrimination (e.g. same vs. different) and identification (e.g. question vs. statement) of statements and questions in ASD (Filipe et al., 2014; Järvinen-Pasley et al., 2008a; McCann et al., 2007;

✉ Fang Liu
f.liu@reading.ac.uk

¹ School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK

² Music College, Shanghai Normal University, Shanghai, China

Peppé et al., 2007). Within this battery, statement-question identification is assessed using a turn-end task with single words, e.g. “Carrots.” vs. “Carrots?”. Statement-question discrimination is assessed within a short-item discrimination task, which contains the laryngographic sounds (devoid of meaning) of the statement-question pairs, as well as those of the liking-disliking pairs, e.g. “tea” pronounced as though the speakers like it or dislike it, from the affect subtask in PEPS-C. Thus, the identification and discrimination tasks are unmatched in stimulus type (speech vs. laryngographic sounds) and in the number of relevant stimuli (only statements or questions are used in the identification task, whereas both statement-question and liking-disliking pairs are included in the discrimination task) in these studies (Filipe et al., 2014; Järvinen-Pasley et al., 2008a; McCann et al., 2007; Peppé et al., 2007). Results from these studies suggest that individuals with ASD are unimpaired in statement-question identification compared to typically developing (TD) peers (Filipe et al., 2014; Järvinen-Pasley et al., 2008a; McCann et al., 2007; Peppé et al., 2007). However, impaired discrimination between statements and questions was observed in one sample of participants (31 ASD vs. 72 TD participants) (McCann et al., 2007; Peppé et al., 2007), while a different sample showed intact discrimination (21 ASD vs. 21 TD participants) (Järvinen-Pasley et al., 2008a). In summary, studies using PEPS-C suggest intact statement-question identification but the results on statement-question discrimination in ASD are unclear, in part due to limitations of the design, as well as mixed results from different studies (Filipe et al., 2014; Järvinen-Pasley et al., 2008a; McCann et al., 2007; Peppé et al., 2007).

Using prosodic tasks other than PEPS-C (e.g. sentence stimuli from Patel et al., 1998), several studies also reported intact statement-question identification beyond single-word stimuli in ASD (Chevallier et al., 2009; Järvinen-Pasley et al., 2008a; Paul et al., 2005). However, using disyllabic phrases from Jiang et al. (2010), Jiang et al. (2015) revealed impaired identification and discrimination of statement-question intonation in Mandarin speakers with ASD. While the different results between Jiang et al. (2015) and other studies (Chevallier et al., 2009; Järvinen-Pasley et al., 2008a; Paul et al., 2005) may be attributed to the different language or cultural background of the participants: Mandarin Chinese versus English, it may also be the case that the discrepancy was due to differences in task difficulty across these studies. Indeed, participants from both the ASD and TD groups performed at ceiling in Paul et al. (2005), which used the stimuli from Patel et al. (1998). In that stimulus set, large pitch contrasts exist between the statements and questions (Patel et al., 1998), and research has shown that even individuals with congenital amusia, a neurodevelopmental disorder of pitch processing, can perform as well as TD controls on both identification and discrimination of

these statements/questions (Ayotte et al., 2002; Patel et al., 2008). Addressing the issue with ceiling performance in the literature, Liu et al. (2010) designed and created a new set of ecologically valid stimuli with relatively subtle pitch contrasts between statements and questions, and revealed prosodic deficits in congenital amusia. Thus, using stimuli from Liu et al. (2010), the current study aimed to examine whether English-speaking individuals with ASD would also show impaired statement-question identification and discrimination when task difficulty is increased.

In addition to identification/discrimination accuracy rates, it has been suggested that participants’ response patterns should also be scrutinised in order to detect possible response biases in ASD (Järvinen-Pasley et al., 2008a). Specifically, Peppé et al. (2007) observed that while children with ASD performed as well as controls in terms of judgement accuracy in statement-question identification, they were biased towards judging questions as statements. In this study, 12.9% of the ASD participants and 2.7% of the control participants judged all the questions as statements, showing a declarative bias, although this percentage difference did not reach statistical significance (Peppé et al., 2007). For discrimination, impaired performance in ASD was mainly driven by false alarms, i.e. judging the same items as different (Peppé et al., 2007). To investigate the declarative bias in ASD further, Järvinen-Pasley et al. (2008a) examined another sample of participants and included the identification task in Patel et al. (1998) in addition to the turn-end task in PEPS-C. While no significant group difference in response patterns was observed for the turn-end task in PEPS-C, a significant declarative bias was observed among 50% of participants with ASD (in comparison to 10% of controls) for the identification task from Patel et al. (1998) (Järvinen-Pasley et al., 2008a). However, no response bias emerged among Mandarin speakers with ASD for either identification or discrimination in Jiang et al. (2015), although significantly lower accuracy rates were observed in ASD compared to TD. Thus, among the studies that examined response biases in ASD, mixed findings have been presented, with some studies indicating a response bias based on either statistics or simply percentage comparison (Järvinen-Pasley et al., 2008a; McCann et al., 2007; Peppé et al., 2007), whereas others reporting no response bias (Jiang et al., 2015), depending on the tasks and samples.

In summary, despite much research (Chevallier et al., 2009; Filipe et al., 2014; Järvinen-Pasley et al., 2008a; Jiang et al., 2015; McCann & Peppé, 2003; McCann et al., 2007; Paul et al., 2005; Peppé et al., 2007), it remains unclear whether individuals with ASD are associated with deficits in identification and/or discrimination of statements and questions, and whether there are response biases driving the observed accuracy rates. These questions need to be addressed, as the answers have implications for the prosody

phenotypes of ASD. As mentioned earlier, due to the limitations of the design in PEPS-C (Peppé & McCann, 2003), the short-item discrimination task contains not only statement-question pairs but also liking-disliking pairs, and in laryngographic sounds rather than in natural speech. Thus, one cannot make inferences about the ability to discriminate statements from questions in everyday language from this task. However, if it is indeed the case that ASD is associated with intact identification but impaired discrimination as reported in Peppé et al. (2007), this dissociation between identification and discrimination may be interpreted as a special feature related to ASD phenotypes (Peppé et al., 2007). An association between identification and discrimination has been observed in other studies: both are intact (Järvinen-Pasley et al., 2008a); or both impaired (Jiang et al., 2015). To further clarify this issue and to help understand the phenotypes of ASD, the current study employed both identification and discrimination tasks from Liu et al. (2010) to investigate response patterns and the relationship between statement-question identification and discrimination in ASD.

Production of Statement-Question Intonation in ASD

In contrast to the mixed findings reported in perception studies, evidence from production studies has consistently suggested atypical intonation production in ASD (Filipe et al., 2014; Fusaroli et al., 2017; McCann & Peppé, 2003; McCann et al., 2007). Specifically, statement responses of individuals with ASD were more likely to be judged as questions or ambiguous than those of controls (McCann et al., 2007; Peppé et al., 2007). In addition, utterances by individuals with ASD were much less likely to be judged as normal or natural than those of controls (Filipe et al., 2014). These ratings were either given by the experimenter (“tester”) (McCann et al., 2007; Peppé et al., 2007) or by independent adult participants (Filipe et al., 2014). Although informative, subjective ratings do not reveal what aspects of intonation production were atypical in ASD (e.g. pitch, duration, and intensity). In studies using objective acoustic measures, individuals with ASD showed significantly greater pitch range, mean pitch, and maximum pitch than controls for both statements and questions (Filipe et al., 2014), and increased and inappropriate use of pitch accents as well as difficulty in producing high frequency boundary tones (Fosnot & Jun, 1999). These findings were supported by Fusaroli et al. (2017), who systematically reviewed the literature quantifying acoustic patterns in ASD and identified significant differences in pitch production (e.g. pitch range and mean pitch) between individuals with ASD and controls, while finding no significant differences in other acoustic features (e.g. intensity, duration). In sum, the atypical production of intonation in ASD seems to be related to

a salient acoustic parameter—pitch (DePape et al., 2012; Fusaroli et al., 2017).

However, previous studies on intonation production in ASD (Filipe et al., 2014; Fosnot & Jun, 1999; McCann et al., 2007; Peppé et al., 2007) have not conducted acoustic analysis to verify the acoustic realisation of pitch direction in statements and questions in ASD. Acoustic measures are important because question and statement intonation are heavily dependent upon pitch direction, with rising tones representing questions and falling tones representing statements (Cruttenden, 1997; Lieberman, 1960). Misuse of pitch itself can cause not only atypical intonation production but also misperception of statements and questions. Indeed, studies on congenital amusia have demonstrated the importance of acoustic analysis in quantifying pitch realisation in production, when examining the relationship between production and perception (Hutchins & Peretz, 2012; Liu et al., 2010; Loui et al., 2008). However, it remains to be determined whether intonation production and perception abilities are related or dissociated among individuals with ASD. The current study addressed this issue by including an intonation imitation task and using acoustic measures to assess pitch direction of the final words in the produced statements and questions (Liu et al., 2010).

Perception of Pitch in Speech Versus Music in ASD

As in speech, pitch is also used extensively in music to convey meaning and emotion (Patel, 2008). It has been intensely debated whether pitch processing is domain-specific or domain-general between speech and music domains (Mantell & Pfordresher, 2013; Patel, 2008; Peretz & Coltheart, 2003). In particular, Peretz and Coltheart (2003) proposed that pitch information within a musical context is processed by a tonal encoding module which is absent in spoken pitch processing. Other researchers, however, argued for shared systems underlying the processing of information across both domains (Koelsch, 2011; Koelsch & Siebel, 2005; Patel, 2008; Sammler et al., 2009). Comparing intonation perception with melodic contour perception, Jiang et al. (2015) observed enhanced/intact melodic contour identification/discrimination but impaired statement-question identification and discrimination in Mandarin speakers with ASD. This finding suggested pitch processing deficits specific to the speech domain in ASD (Jiang et al., 2015). However, other studies indicated enhanced identification of pitch contours (e.g. rising, falling, falling-rising, rising-falling) across speech and musical stimuli (Järvinen-Pasley et al., 2008b), as well as superior discrimination of pitch patterns across speech-speech and speech-music stimulus pairs in ASD versus TD (Järvinen-Pasley & Heaton, 2007). Therefore, further research is warranted to clarify the domain specificity or generality of pitch processing in ASD. To our knowledge,

no studies have yet compared pitch perception in ASD using speech and musical stimuli that are matched in global pitch contours derived from statement-question intonation. The present study aimed to fill this gap by investigating whether individuals with ASD would process intonation embedded in speech and musical stimuli differently, using the musical analogues of the statement-question discrimination task in Liu et al. (2010).

The Development of Prosodic Abilities and its Relationship With Pitch Sensitivity

Studies of prosodic development in TD children suggest that there are significant improvements in the perception and production of statement-question intonation between ages 5 and 11 (Wells et al., 2004). As children grow older, pitch becomes the primary cue for the statement-question contrast compared to intensity and duration in production (Patel & Grigos, 2006). While 4-year-olds used lengthened duration of the final syllable rather than a rising pitch contour to signify questions, 7-year-olds used multiple acoustic cues (including pitch, intensity and duration) and 11-year-olds used pitch cues predominantly to differentiate statements from questions (Patel & Grigos, 2006). Given that language delay and impairment are prevalent among children and youth with ASD (Kwok et al., 2015), it may be the case that the development of prosodic skills is also delayed in ASD. Lyons et al., (2014) investigated the developmental changes of four prosodic functions, including the perception and production of statement-question intonation, stress, phrasing, and affect, in “language-normal” and “language-impaired” preadolescents (9–12 years old) and adolescents (13–17 years old) with and without ASD. The results suggest that TD preadolescents performed as well as TD adolescents on statement-question identification and production, and thus no developmental improvement was observed among TD participants due to ceiling performance. The same pattern of results was also seen in “language-normal” ASD preadolescents and adolescents, who performed similarly to the TD groups on both identification and production of statements and questions. For the “language-impaired” ASD groups, however, significant age-related improvement was observed for identification, but not for production, of statements and questions. That is, while impaired statement-question identification was only observed among “language-impaired” ASD preadolescents, but not among adolescents, impaired statement-question production persisted among “language-impaired” ASD preadolescents and adolescents. Thus, there are developmental delays in the perception and production of statements and questions among “language-impaired” individuals with ASD (Lyons et al., 2014).

In addition to the close relationship with language abilities (Lyons et al., 2014), prosodic skills also correlate

significantly with pitch processing abilities (Liu et al., 2010, 2012; Vuvan et al., 2015). In typical development, there are age-related improvements in the ability to discriminate the direction of pitch changes between ages 6–11 (Fancourt et al., 2013). However, it has been reported that individuals with ASD show enhanced pitch discrimination early in development, and this ability maintains across children, adolescents and adults and does not correlate with receptive vocabulary (Mayer et al., 2016). By contrast, controls show significant gains in pitch discrimination performance across development, which also correlates significantly with receptive vocabulary scores (Mayer et al., 2016). This raises the questions as to whether and how pitch processing abilities influence intonation perception and production in individuals with ASD, and whether age plays a role in these abilities across the lifespan. The current study addressed these questions by examining the development of statement-question perception and production across children, adolescents and adults with and without ASD, as well as its relationship with pitch direction discrimination thresholds.

Present Study

In the current study, we matched ASD and TD children, adolescents and adults for age, sex, nonverbal IQ, receptive vocabulary, as well as verbal and nonverbal short-term memory. Focusing on the prosodic feature of statement-question intonation and the acoustic parameter of pitch, we examined intonation processing in ASD and TD from the perspectives of task condition (discrimination, identification, imitation), response bias, stimulus type (speech, music), developmental changes, and its association with pitch thresholds. We asked whether individuals with ASD differed from controls in their ability to discriminate, identify, and imitate statement-question intonation, whether individuals with ASD showed response bias in discrimination and identification tasks, and whether performance on intonation perception and production related to pitch direction discrimination thresholds. We also examined whether individuals with ASD would perform better on musical pitch processing than on linguistic pitch processing, comparing discrimination of natural speech and their musical analogues. Finally, we examined the effect of age on pitch and intonation perception and production for both ASD and control groups. Based on previous findings, we predicted that: (a) participants with ASD would show impaired performance compared to controls in intonation discrimination and identification tasks, and they would show response biases towards judging the same pairs as different and identifying questions as statements; (b) participants with ASD would show poorer performance on the imitation task compared with controls; (c) participants with ASD would perform better on the musical condition than the speech condition in the discrimination task; (d) across both groups,

performance on intonation processing would be associated with pitch direction discrimination thresholds; and (e) participants with ASD would show different developmental trajectories for pitch and intonation processing compared with controls.

Methods

Participants

A priori power analysis was conducted using G*Power (Faul et al., 2009). To detect the interaction of Group (ASD vs. control) by Condition (speech vs. music or identification vs. imitation) by Age (child, adolescent vs. adult) in the present design, 64 participants (with 32 in each group) were required to reach a power of 0.80, with a large effect size ($f=0.40$) and an alpha of 0.05. Given the mixed findings in the ASD literature and to further increase the power of our study, we recruited a total of 84 participants, 42 with ASD (12 female, 30 male) and 42 controls (12 female, 30 male), resulting in a power of 0.91.

All participants were native speakers of British English, recruited through email lists, word of mouth, online and social media advertisements, local schools, charities and organisations, as well as departmental participant databases. Participants in the ASD group all received a formal diagnosis of ASD by professional clinicians (verified by official clinical reports), and their high autistic traits were also confirmed using the cut-off scores of 32 (adults), 30 (adolescents) and 76 (children) on the Autism-Spectrum Quotient (AQ) (Auyeung et al., 2008; Baron-Cohen et al., 2001, 2006). All control participants scored below these cut-offs. In addition to the AQ, Empathy Quotient (EQ) and Systemizing Quotient (SQ) were also collected through questionnaires. 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). The study was approved by the University of Reading Research Ethics Committee. Written informed consent/assent was obtained from the participants and/or their parents prior to the experiment.

Given the significant effects of IQ, receptive vocabulary, short-term memory, and musical training on pitch and prosodic processing (Acton & Schroeder, 2001; Bidelman et al., 2013; Chowdhury et al., 2017; Heaton et al., 2008; Mayer et al., 2016; McCann et al., 2007; Peppé et al., 2007; Tillmann et al., 2016), we gathered related background measures from all participants (Table 1). Specifically, participants completed a nonverbal IQ test using the Raven's Standard Progressive Matrices Test (Raven et al., 1998) and a receptive vocabulary test using the Receptive One Word Picture

Vocabulary Test IV (ROWPVT-IV) (Martin & Brownell, 2011). The Corsi block-tapping task was used to assess participants' nonverbal short-term memory span (Kessels et al., 2000), and the digit span task was used to assess verbal short-term memory (Wechsler, 2003). Participants' musical training background was collected using a questionnaire, and their years of formal musical training were summed across all instruments including voice (Pfordresher & Halpern, 2013).

Following the age cut-offs for the Autism-Spectrum Quotient (Auyeung et al., 2008; Baron-Cohen et al., 2001, 2006), participants were divided into three age cohorts: children (7–11 years), adolescents (12–15 years), and adults (≥ 16 years). The age range of the child cohort was between 7.39 and 11.92 years, that of the adolescent cohort was between 12.08 and 15.75, and that of the adult cohort was between 18 to 55.72 years. Table 1 shows the characteristics of the participants. The groups in each of the three age cohorts were largely matched on the background measures, with the exception that the ASD adolescents showed lower receptive vocabulary and nonverbal IQ scores than the control adolescents. To control for the possible contribution of receptive verbal ability and nonverbal IQ to the current results, these scores were entered as covariates in the analysis of Bayesian ANCOVA in the Results section.

Tasks

The present study consisted of one pitch direction discrimination thresholds task and four intonation perception/production tasks from Liu et al. (2010). All tasks were conducted in a sound-proof booth at a university laboratory.

Pitch Direction Discrimination Thresholds Task

Pitch thresholds were measured using an adaptive-tracking procedure with a “two down, one up” staircase method in Matlab (MATLAB, 2010). In each trial, participants were presented with three gliding tones of 600 ms each, centring on 500 Hz, with two moving in the same direction and the other moving in the opposite direction. Participants were required to identify the “odd-one-out” among the three tones. Starting with six semitones, the initial step size was one semitone which reduced to 0.1 semitones after four reversals and 0.02 semitones after eight reversals, with a total of 14 reversals. The threshold was calculated as the mean excursion size of the target glide of the last six reversals.

A practice session of four trials was provided to ensure that all participants understood the task. Participants were required to achieve 100% correct on the practice trials (with feedback) before proceeding to the testing session. Given that inattention may impact performance on

Table 1 Characteristics of the ASD (N=42) and control groups (N=42)

Age group	Diagnostic group	Age	Musical training	NVIQ	ROWPVT-IV	Corsi	Digit span	AQ	EQ	SQ	
Children (N = 28)	ASD	9.42(1.19)	1.43(2.26)	73.93(26.11)	125.57(11.47)	4.79(0.80)	5.33(1.26)	95.86(23.18)	19.14(5.87)	29.71(10.88)	
	control	9.35(1.44)	1.29(0.99)	83.57(20.80)	123.36(11.13)	5.29(1.27)	5.64(0.93)	47.77(19.70)	39.00(9.26)	26.77(4.40)	
	Comparison statistics: Bayesian										
	W	94.5	115	113	67	112.5	104	13	171.5	73	
	BF ₀₁	2.84	2.48	2.39	2.02	2.10	2.44	0.05	0.04	2.18	
Adolescents (N = 20)	Median	0.003	-0.14	-0.17	0.28	-0.24	-0.15	1.20	-1.20	0.21	
	95%CI	[-0.64,0.66]	[-0.83,0.52]	[-0.90,0.51]	[-0.36,1.00]	[-0.99,0.42]	[-0.90,0.52]	[0.34,2.10]	[-2.05, -0.33]	[-0.45,0.93]	
	ASD	13.95(1.38)	3.40(3.06)	55(31.00)	116.2(18.71)	5.7(1.57)	5.7(1.06)	37.10(5.69)	13.00(6.67)	47.60(11.72)	
	control	13.77(1.09)	3.10(3.00)	77(20.17)	134.3(12.46)	6.2(1.40)	6.2(0.79)	15.40(7.31)	45.63(12.46)	34.88(13.94)	
	Comparison statistics: Bayesian										
Adults (N = 36)	W	47.5	45	74	78	60	68	0	40	9	
	BF ₀₁	2.43	2.21	0.78	0.64	2.10	1.60	0.06	0.27	1.07	
	Median	0.03	0.15	-0.61	-0.66	-0.19	-0.34	1.36	-1.13	0.53	
	95%CI	[-0.73,0.74]	[-0.57,0.95]	[-1.55,0.17]	[-1.61,0.14]	[-0.98,0.53]	[-1.23,0.42]	[0.32,2.49]	[-2.57,0.00]	[-0.39,1.71]	
	ASD	35.47(13.07)	4.50(6.50)	50.00(28.54)	111.47(13.74)	5.72(1.49)	7.00(1.72)	35.83(9.21)	22.89(9.85)	79.17(28.55)	
Adults (N = 36)	control	35.34(12.88)	5.14(7.03)	41.94(29.06)	108.78(13.61)	6.06(1.00)	7.06(1.11)	15.06(6.53)	48.22(13.74)	49.89(16.31)	
	Comparison statistics: Bayesian										
	W	162.5	158.5	135.5	143.5	190.5	171	15	307.5	65.5	
	BF ₀₁	3.01	2.93	2.53	2.81	2.30	2.84	0.01	0.01	0.11	
	Median	0.05	0.04	0.16	0.13	-0.23	-0.08	1.35	-1.31	0.84	
95%CI	[-0.55,0.62]	[-0.54,0.65]	[-0.44,0.81]	[-0.44,0.76]	[-0.86,0.38]	[-0.71,0.53]	[0.58,2.15]	[-2.06, -0.55]	[0.17,1.55]		

Age and Musical training are in years; NVIQ and ROWPVT-IV are percentile points of nonverbal IQ and standard scores of receptive verbal ability respectively; Corsi and Digit span are the raw scores of nonverbal and verbal short-term memory respectively; AQ, EQ and SQ are the scores of Autism Spectrum, Empathy and Systemizing Quotient respectively. Bayes factors from a default prior 2-tailed Bayesian Mann-Whitney-Wilcoxon Test are expressed in terms of the Bayes factor in favour of the null hypothesis of no difference (BF₀₁). The delta effect size in these Bayesian comparisons is given by the median of a posterior distribution and 95% credible intervals

adaptive-tracking pitch thresholds tasks, especially in children (Fancourt et al., 2013; Horváth et al., 2009; McDermott & Oxenham, 2008), participants were required to make their responses orally for the experimenters to input into the computer, in order to maintain their attention.

Intonation Tasks

The intonation tasks consisted of four subtests assessing discrimination of statements and questions in natural speech and in their musical analogues (composed of gliding tones), and identification and imitation of these statements and questions. The tasks were presented in counterbalanced order across participants using Praat. Taken from Liu et al. (2010), the speech stimuli were 18 statement-question pairs, cross-spliced so that each pair began with the same stem and differed only in the final word (See Supplementary Figure S1 for examples, and Supplementary Table S1 for full details of the speech stimuli).

Musical analogues of the sentences were created in Praat (Boersma & Weenink, 2001), matching the original sentences in pitch and temporal patterns, following the procedure in Patel et al. (1998). These musical tones were made of the fundamental frequency and its seven odd harmonics of the individual syllables in the original sentences, with peak amplitudes normalized to match those of the sentences (see Liu et al. (2010) for full details).

The speech and musical discrimination tasks were conducted in two separate blocks (order counterbalanced), where participants were presented with 36 pairs of stimuli (either speech or musical analogues) in either the same or different condition. Participants were asked whether the pairs in each trial were the same or different, with their answers recorded by an experimenter by clicking a button on the computer. The interstimulus interval was 750 ms and the intertrial interval was 2 s. Two additional pairs were included as practice trials to familiarise participants with the procedure.

The identification and imitation tasks were conducted in the same block, where participants were presented with the 36 speech sentences one at a time. They were instructed to first imitate the sentence just played as exactly as possible (while their voices were recorded), and then to indicate whether the original sentence (not their imitation) was a statement or a question. The experimenter recorded the identification responses in Praat. Prior to the experiment, participants were familiarised with the procedure using two additional sentences (one statement and one question) in a practice session.

Data Analysis

In the pitch direction discrimination task, thresholds were transformed using log transformation for parametric

statistical analysis (Howell, 2009). Nine ASD participants and two controls did not complete this task. Additionally, to screen for inattentive performers, following Moore et al. (2008), the visual tracks were closely inspected on an individual basis. Two children were identified as “non-compliant” (a term used in previous literature; Fancourt et al. (2013)) performers due to fluctuations of attention, and their data were excluded from further analysis (see Supplementary Figure S2 for their visual tracks and Supplementary Table S2 for the remaining participants’ demographic characteristics in the pitch direction discrimination task).

Signal detection analysis was carried out for the intonation data. Specifically, in the discrimination tasks, correct responses to “different” trials were coded as hits; and in the identification and imitation tasks, correct responses to “question” trials were coded as hits, and d' was corrected using the log-linear rule (Hautus, 1995). For the recordings from the imitation task, we employed acoustic analysis as a quantitative measure of performance accuracy using an earlier version of ProsodyPro (Xu, 2013) in Praat. Signed glide sizes (in Hz) were extracted from the final words of the sentences, with negative values indicating downward glides and recorded as statements, and positive values indicating upward glides and recorded as questions (Liu et al., 2010).

Bayesian analyses were run using JASP software (JASP Team, 2020). Bayes Factors indicate the strength of the evidence obtained and are particularly helpful in determining when the evidence supports the null hypothesis over an alternative. Unlike frequentist statistics, Bayes Factors test the relative probability of the two hypotheses given the data, rather than the probability of the data given the null hypothesis and so can be used to support both alternative and null hypotheses (Dienes, 2014). The Bayes Factor (BF) in favour of the null is the reciprocal of the BF in favour of the alternative. Also unlike frequentist significance testing, BFs give continuous measures of the likelihood of one hypothesis over another, which means cut-off values (e.g. $p=0.05$) are inappropriate. For the interpretation of BFs as evidence for hypotheses, Raftery (1995, p.139) suggested ranges of values equivalent to different “strengths” of evidence, where a BF value above 1 and less than 3 is “weak” evidence and a BF between 3 and 20 represents “positive” evidence for a hypothesis.

Results

Pitch Direction Discrimination Task

Figure 1 shows boxplots of the pitch thresholds for the ASD and control groups. Bayesian ANCOVA was fit to the data with group (ASD vs. control) and age cohorts (child, adolescent and adult) as the between-subjects variables

Fig. 1 Pitch threshold in semi-tone (st) of each age cohort by group from the pitch direction discrimination task

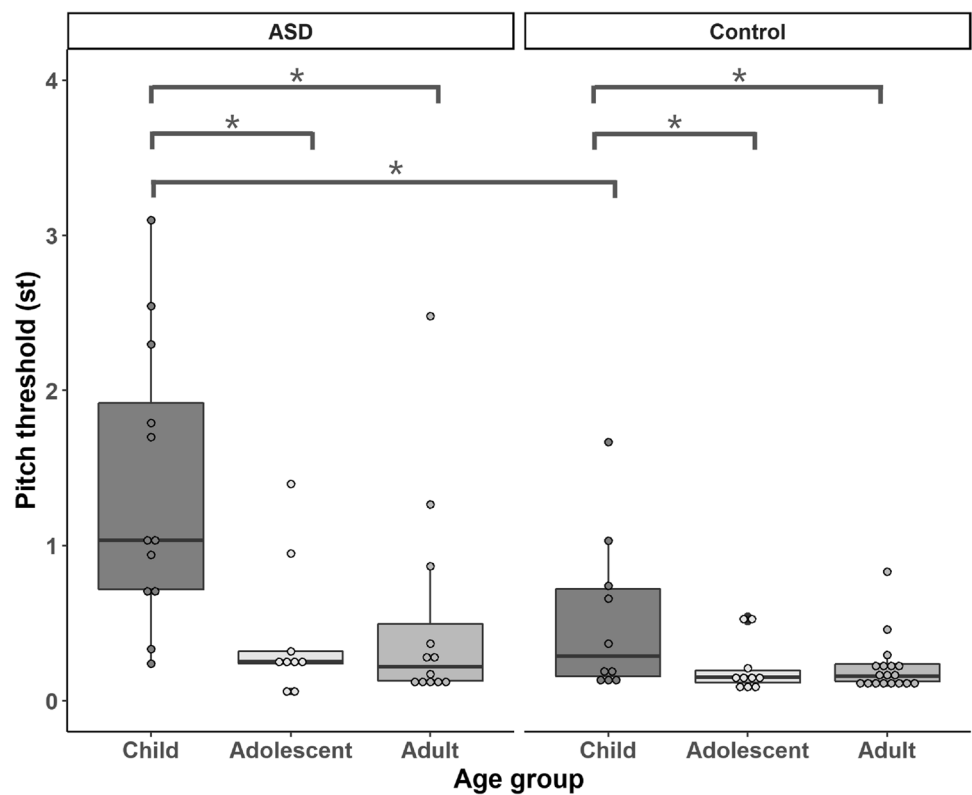


Table 2 Differences of pitch thresholds between age cohorts within each group

		ASD	Control
Adolescent	Child	BF ₁₀ = 15.00	BF ₁₀ = 1.52
Adult	Child	BF ₁₀ = 14.97	BF ₁₀ = 2.54
Adult	Adolescent	BF ₁₀ = 0.40	BF ₁₀ = 0.38

and receptive vocabulary and nonverbal IQ scores as the covariates. The model revealed positive main effects of age (BF₁₀ = 400,381.43) and group (BF₁₀ = 23.85) on pitch thresholds. The interaction between age and group also received positive support from Bayesian factor (BF₁₀ = 4.70). The Bayesian post-hoc analysis suggested that both groups showed similar developmental trajectories, with the adult and adolescent cohorts performing better than the child cohort, though this trend was more pronounced in the ASD group than in the control group (see Table 2). The main effect of group was mainly driven by the difference across the child cohorts, as ASD children showed worse pitch thresholds than control children (BF₁₀ = 7.07). The difference between the adult cohorts was weak (BF₁₀ = 1.79) and no difference was found between the adolescent cohorts (BF₁₀ = 0.46, *weak* evidence in favour of H₀).

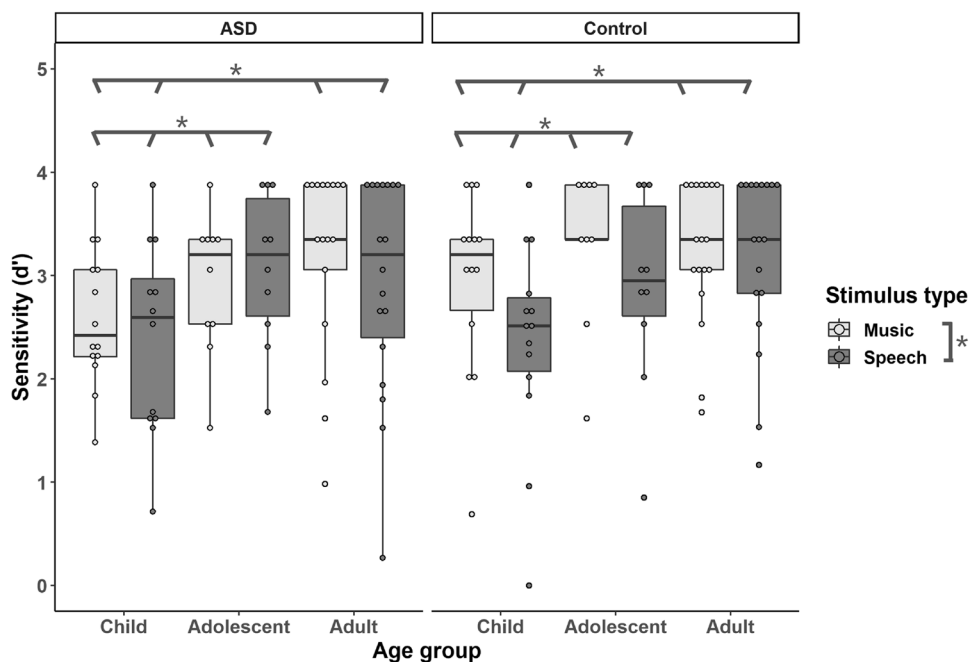
There was a main effect of nonverbal IQ (BF₁₀ = 251), whereas the main effect of receptive verbal ability was only

weakly supported by Bayes factor (BF₁₀ = 1.82). A Bayesian Kendall correlation analysis (1-tailed) showed that nonverbal IQ and receptive verbal ability were weakly associated with the performance of the control group (NVIQ: tau = -0.25, BF₀ = 1.14; ROWPVT-IV: tau = -0.18, BF₀ = 1.30) but not those of the ASD group (NVIQ: tau = -0.14, BF₀ = 0.74; ROWPVT-IV: tau = -0.04, BF₀ = 0.30).

Intonation Discrimination Tasks

Figure 2 shows boxplots of the sensitivity d' on the discrimination tasks for the ASD and control groups. Bayesian repeated measures ANCOVA was fit to the data with group (ASD vs. control) and age cohorts (child, adolescent and adult) as the between-subjects variables, stimulus type (speech vs. music) as the within subject-variable, and receptive vocabulary and nonverbal IQ scores as the covariates. The model revealed a positive main effect of stimulus type (BF₁₀ = 7.09), with all groups performing better on the music condition than on the speech condition, and a main effect of age (BF₁₀ = 3.57). The Bayesian post-hoc analysis showed that the adult cohort performed better than the child cohort (BF₁₀ = 47.19), and the adolescent cohort also performed better than the child cohort (BF₁₀ = 2.95), whereas the adult and adolescent cohorts performed comparably (BF₁₀ = 0.28, *positive* evidence in favour of H₀). No evidence for other main effects or interactions was observed, though

Fig. 2 d' of each age cohort by stimulus type and by group from the discrimination task



the evidence for the null hypotheses was also weak: group ($BF_{10}=0.35$), stimulus type by age cohort ($BF_{10}=0.59$), stimulus type by group ($BF_{10}=0.58$), age cohort by group ($BF_{10}=0.35$) and stimulus type by age cohort by group ($BF_{10}=0.34$). The evidence for main effects of receptive verbal ability ($BF_{10}=0.995$), and nonverbal IQ ($BF_{10}=0.91$), was equivocal.

To assess the relationship between pitch thresholds and the performance on intonation discrimination, 1-tailed Bayesian Kendall’s correlation analysis was carried out separately for the groups and tasks, results are reported in Table 3. Pitch thresholds were negatively correlated with performance on both tasks for both groups: the lower (better) the pitch thresholds, the better performance on the speech and musical tasks.

Intonation Identification and Imitation Tasks

Figure 3 shows boxplots of the sensitivity d' on the identification and imitation tasks for the ASD and control groups. A Bayesian repeated measures ANCOVA was conducted with group (ASD vs. control) and age (child, adolescent and adult) as the between-subjects variables, task type (identification vs. imitation) as the within-subject

variable, and receptive vocabulary and nonverbal IQ scores as the covariates. The model revealed a main effect of age ($BF_{10}=4.06$). The post-hoc analysis showed that the adult cohort performed better than the child cohort ($BF_{10}=18.01$) but similarly to the adolescent cohort ($BF_{10}=0.57$), and the adolescent and the child cohorts performed comparably ($BF_{10}=0.47$). For all other main effects and interactions, there was positive evidence in favour of the null hypotheses: task type ($BF_{10}=0.23$); group ($BF_{10}=0.14$); task type by age cohort ($BF_{10}=0.33$); task type by group ($BF_{10}=0.13$); age cohort by group ($BF_{10}=0.11$) and task type by age cohort by group ($BF_{10}=0.01$).

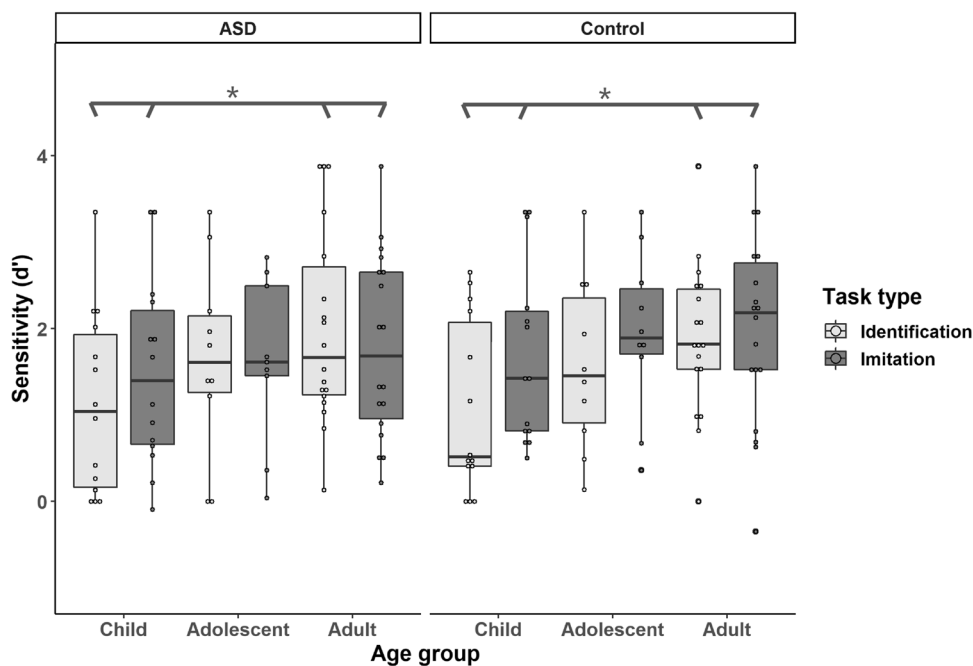
In addition, there was a weak main effect of receptive verbal ability, $BF_{10}=2.93$, whereas no main effect of nonverbal IQ was observed, $BF_{10}=0.73$. However, a Bayesian Kendall correlation analysis (1-tailed) showed that receptive verbal ability was not associated with performance on the imitation task (ASD: $\tau=0.08$, $BF_{+0}=0.40$; control: $\tau=0.10$, $BF_{+0}=0.50$) or the identification task (ASD: $\tau=0.02$, $BF_{+0}=0.23$; control: $\tau=0.07$, $BF_{+0}=0.38$) in either group, with Bayes factors supporting the null hypotheses positively.

Performance on identification and imitation was positively correlated for both groups (ASD: $\tau=0.40$,

Table 3 Kendall’s correlations between performance on pitch thresholds and intonation discrimination tasks by group

	ASD group	Natural speech	Musical analogues	Control group	Natural speech	Musical analogues
Pitch threshold	tau	-0.33	-0.34	tau	-0.38	-0.45
	BF_{-0}	11.88	15.60	BF_{-0}	91.05	765.98

Fig. 3 d' of each age cohort by group from the identification and imitation task



$BF_{+0} = 329.97$; control: $\tau = 0.28$, $BF_{+0} = 11.50$). Using 1-tailed Bayesian Kendall’s correlation analysis, we analysed the relationship between pitch thresholds and identification as well as imitation tasks (see Table 4). The results indicated that pitch thresholds were negatively correlated with the performance on both tasks in the control group, but were only weakly associated with the identification performance in the ASD group.

Response Bias

To measure whether individuals with ASD showed response biases between the same versus different pairs (discrimination task) or questions versus statements (identification task), i.e. judging the same pairs as different or questions as statements, we calculated the percentage of correct responses to same/different pairs in the discrimination task and questions/statements in the identification task for the two groups.

A 2×2 mixed ANOVA with Bayesian analysis was conducted. Group (ASD vs. control) was the between-subjects factor, and response type (same vs. different in the discrimination task and question vs. statement in the identification task) was the within-subjects factor. In the

discrimination task, there was a main effect of response type ($BF_{10} = 137.42$) with participants showing poorer performance on different pairs ($M = 0.90$, $SD = 0.12$) than on same pairs ($M = 0.94$, $SD = 0.07$). No main effect of group, or group by response type interaction was found, with Bayes factors tending to support the null hypotheses in both cases ($BF_{10} = 0.32$ and $BF_{10} = 0.52$, respectively). In the identification task, there was a main effect of response type, $BF_{10} = \infty$, with participants showing poorer performance on questions ($M = 0.62$, $SD = 0.22$) than statements ($M = 0.90$, $SD = 0.15$). No main effect of group or group by stimulus type interaction was observed, with $BF_{10} = 0.20$ and $BF_{10} = 0.24$, respectively.

To inspect individual response patterns, following Steffens et al. (2020), we calculated the probability that each individual accuracy rate was due to random guessing based on the binomial distribution. Accuracy rates with probabilities > 0.05 were interpreted as being likely due to random chance alone, whereas accuracy rates with probabilities ≤ 0.05 were interpreted as being unlikely due to chance alone (Steffens et al., 2020). We found that all participants performed above chance level in the discrimination task, while 24 participants showed chance level performance

Table 4 Kendall’s correlations between performance on pitch thresholds and intonation identification/imitation tasks by group

	ASD group	Intonation identification	Intonation imitation	Control group	Intonation identification	Intonation imitation
Pitch threshold	τ	-0.22	-0.15	τ	-0.43	-0.26
	BF_{-0}	2.06	0.80	BF_{-0}	509.47	5.09

in the identification task (12 ASD vs. 12 control). A 2×2 mixed ANOVA with Bayesian analysis with Group as the between-subjects factor and response type as the within-subjects factor on the responses of the 24 participants revealed no main effect of group, $BF_{10} = 0.28$, or group by stimulus type interaction, $BF_{10} = 0.37$. There was a main effect of stimulus type, as participants were less able to identify questions ($M = 0.27$, $SD = 0.24$) than statements ($M = 0.82$, $SD = 0.20$).

Discussion

Using pitch thresholds and intonation perception and production tasks, the present study examined the abilities of individuals with and without ASD to use pitch to differentiate, identify, and imitate intonation (statements vs. questions) and whether these abilities would be affected by response bias, age (child, adolescent vs. adult), stimulus type (speech vs. music), and pitch direction discrimination thresholds. The main results showed that the performance of intonation discrimination (in both speech and music conditions), identification, and imitation was comparable between the ASD and TD groups within each age cohort, and that performances across tasks were largely independent of participants' receptive verbal ability and nonverbal IQ, especially for participants with ASD. In addition, no response bias was observed in the discrimination and identification of statements and questions among participants with ASD. Participants' abilities to discriminate, identify and imitation intonation were associated with their pitch direction discrimination thresholds for both groups. There were also age-related improvements across all tasks for both groups. These findings suggest that some individuals with ASD may have genuinely intact abilities to differentiate, identify, and imitate statement-question intonation, and they may also show similar developmental trajectories as typically developing individuals, with performance on both intonation and pitch thresholds increasing with age.

Perception of Statement-Question Intonation and Response Bias in ASD

Regarding discrimination and identification of statements and questions, we found no group differences in response accuracy across all three age cohorts. Bayesian analyses supported our null results weakly for the discrimination task but positively for the identification task. Thus, no strong conclusions can be drawn about intonation discrimination abilities between the ASD and control groups. These findings are consistent with the majority of the literature that suggests intact statement-question identification in ASD (Chevallier et al., 2009; Filipe et al., 2014; Järvinen-Pasley et al., 2008a;

Paul et al., 2005). However, they contradict the findings indicating impaired discrimination (McCann et al., 2007; Peppé et al., 2007), impaired identification among “language-impaired” preadolescents (Lyons et al., 2014), and impaired discrimination and identification among Mandarin speakers (Jiang et al., 2015). Notably, as previously mentioned, the impaired discrimination suggested by McCann et al. (2007) and Peppé et al. (2007) was evaluated using the short-item discrimination task within PEPS-C, which contains the laryngographic sounds of the statement-question pairs, as well as those of the liking-disliking pairs from the affect subtask. Thus, the discrimination performance in ASD reported by these two studies may be confounded by the unnaturalness of the stimuli as well as by participants' ability to discriminate affective pairs.

In addition, the inconsistency between the present study and Jiang et al. (2015) may be explained by language differences. Jiang et al. (2015) used stimuli in Mandarin which is a tone language, while the present study used stimuli in English which is a non-tone language. It has been suggested that the perception of statement-question intonation in tone languages is complicated by the changes in tones, which convey lexical meaning (Jiang et al., 2015; Liu & Xu, 2005; Xu, 2005), resulting in the tasks in Jiang et al. (2015) being more difficult than the present study. Finally, impaired statement-question identification was only observed among “language-impaired” preadolescents in Lyons et al. (2014). Indeed, prosodic skills correlate significantly with language ability in ASD (McCann et al., 2007; Peppé et al., 2007). In our current study, there was a weak main effect of receptive verbal ability on intonation identification/imitation (although correlations were nonsignificant), but not on intonation discrimination, which may be because our ASD and TD participants were largely matched on a range of cognitive abilities (Table 1).

We also examined the response biases that were reported in some studies (Järvinen-Pasley et al., 2008a; Peppé et al., 2007). Inconsistent with those results, but consistent with the findings of Jiang et al. (2015), individuals with ASD in our study did not show a tendency to judge the same pairs as different or identify questions as statements and this null result receives substantial support from the Bayes factors. Similar to controls, our participants with ASD displayed poorer performance when discriminating different pairs than same pairs, and when identifying questions than statements. While the response bias in the discrimination task reported by Peppé et al. (2007) reached a significant level, the declarative bias in identification from the PEPS-C turn-end task lacked statistical support (Järvinen-Pasley et al., 2008a; Peppé et al., 2007). Using the speech stimuli from Patel et al. (1998), Järvinen-Pasley et al. (2008a) observed a significant declarative bias among 50% of participants with ASD (in comparison to 10% of controls) for the identification task.

Following Jiang et al. (2015) and Steffens et al. (2020), the present study used ANOVA models to inspect participants' response patterns, and found no response bias in either discrimination or identification tasks in ASD. It is worth noting that the difference in results between our study and previous studies is not due to our sample size being smaller. In fact, our ASD sample size is the largest among all these studies: 42 (our study), 31 (McCann et al., 2007; Peppé et al., 2007), 21 (Järvinen-Pasley et al., 2008a), and 17 (Jiang et al., 2015). Given the reproducibility problems in science (Begley & Ioannidis, 2015), further studies are needed to determine whether there are genuine response biases in intonation discrimination and identification in ASD.

Production of Statement-Question Intonation in ASD

In the imitation task, participants with ASD showed comparable performance to controls. Although previous research has found a deficit in intonation production in ASD, either based on subjective perceptual judgements or objective acoustic measures (Filipe et al., 2014; Fosnot & Jun, 1999; McCann et al., 2007; Peppé et al., 2007, 2011), the current study did not observe this deficit and the balance of evidence provided by our Bayesian analyses is sufficient for us to be confident in our null results. This discrepancy mainly results from the different methods used in reporting/analysing production data among these studies. Unlike previous studies using subjective judgements of the sentences produced (Filipe et al., 2014; McCann et al., 2007; Peppé et al., 2007, 2011), we explored objective measures by calculating glide sizes to verify the acoustic realisation of pitch direction in statements and questions in ASD. Thus, imitations were scored as correct only if participants shared the same sign in glide size as the models (i.e. statements imitated as statements with final falls would have negative glide sizes, and questions imitated as questions with final rises would have positive glide sizes). Additionally, while the present study was inspired by previous acoustic studies suggesting the key role pitch plays in atypical intonation production in ASD (Filipe et al., 2014; Fosnot & Jun, 1999), we used a different acoustic analysis method than those studies, in order to capture the production of pitch direction specifically. That is, when calculating imitation accuracy, we did not consider mean pitch, pitch range, or other variables, which measured the characteristics of speech production rather than imitation accuracy per se (Filipe et al., 2014; Fosnot & Jun, 1999). Rather, we focused on using objective measures to gauge the relationship between identification and imitation. The results suggested that identification abilities positively correlated with imitation abilities in both groups and that the association was stronger in the ASD group than in the control group, which were substantially supported by Bayes

factors. These findings are consistent with the finding of intact identification and production of turn-end sentences in previous studies (Järvinen-Pasley et al., 2008a; Peppé et al., 2007). Our correlation analysis further indicates that the more accurate the participants were on prosody perception and understanding, the better their performance on prosody production. Thus, an increase in receptive prosodic skills might result in amelioration of expressive prosodic disorder in ASD, and vice versa.

Perception of Pitch in Speech Versus Music in ASD

The third aim of this study was to investigate intonation processing in speech versus music in ASD. The results showed that, like the controls, participants with ASD performed better on discriminating between musical glides than on speech utterances. These findings received positive support from our Bayesian analysis. While the better performance on music than on speech in the intonation task is consistent with our hypothesis for ASD, the same perceptual pattern was also noted in the control group. It has been suggested that semantic information might hamper controls' performance due to their overly selective attention towards the content (Järvinen-Pasley & Heaton, 2007). Similarly, Bijou and Ghezzi (1999) have proposed a Behavior Interference Theory which states that typically developing children tend to focus on social stimuli (i.e. the human voice), whereas these stimuli do not easily obtain attention from children with ASD. In the present study, both groups showed semantic interference with perceptual processing and performed better on discrimination of musical analogues than natural speech. This perceptual pattern is consistent with the findings of Liu et al. (2010) for amusic participants, Cheng et al. (2017) for ASD and control participants, as well as Francis and Ciocca (2003) for typically developing English and Cantonese listeners. However, other studies reported mixed findings regarding the effect of stimulus type on intonation processing among amusic and control participants (Jiang et al., 2010; Liu et al., 2012; Patel et al., 2005, 2008). Further studies are required to tease apart the effects of stimulus type, perceptual acuity, and sensory preference on intonation processing among different participant groups.

Since no group difference was observed in discrimination of speech and musical stimuli, the current findings provide evidence for shared mechanisms of pitch processing between music and speech in both individuals with ASD and controls (Liu et al., 2010). Similar findings were also reported in Cheng et al. (2017). However, numerous previous studies have suggested enhanced musical processing in ASD compared to controls, including local music processing (Mottron et al., 2000), melodic contour identification (Jiang et al., 2015), as well as memory and labelling of musical tones and segmentation of chords (Heaton, 2003). Nevertheless,

more recent studies reported comparable or even impaired musical processing in ASD versus TD (Jamey et al., 2019; Schelinski et al., 2017; Sota et al., 2018). Thus, with mixed findings in the literature (Järvinen-Pasley & Heaton, 2007; Järvinen-Pasley et al., 2008b; Jiang et al., 2015), the domain specificity or generality of pitch processing in ASD warrants further studies.

The Development of Prosodic Abilities and its Relationship With Pitch Sensitivity

The fourth aim of our study was to examine the relationship between psychophysical pitch thresholds and intonation perception/imitation. We found that children with ASD had elevated pitch thresholds relative to their typically developing counterparts with substantial support from Bayesian factors. Adults with ASD performed worse relative to adult controls, and adolescents with ASD performed comparably to their controls in the pitch thresholds task. However, these findings received weak support from Bayesians. So it would be premature to draw strong conclusions specific to these two age cohorts. In addition, our results point to positive relationships between the pitch thresholds and intonation perception/imitation in both groups: the more sensitive to pitch, the better performance on intonation perception. This correlation was more pronounced in the control group suggested by Bayes factors. This finding is consistent with previous research showing an overall positive relationship between low-level and higher-level pitch processing (Germain et al., 2019). These findings likely reflect a bottom-up cascading in which the degree of low-level strength or impairment influences performance at later stages, such as language acquisition and communication (Bertone et al., 2010; Germain et al., 2019; Stevenson et al., 2014).

Finally, the major contribution of the current study relates to the effects of age on pitch thresholds and intonation processing. Both groups showed age-related improvements across all tasks with positive support from Bayes factors. In particular, adults consistently showed smaller pitch thresholds and better intonation perception and imitation than children, suggesting a developmental improvement in pitch perception and intonation processing. Interestingly, age-related changes across the lifespan from children to adolescents to adults were not identical across different tasks. Specifically, there were no significant differences in pitch thresholds and intonation discrimination between the adult and adolescent cohorts, who performed significantly better than the child cohort on those tasks. In terms of intonation identification and imitation, however, there was a gradual improvement from children to adolescents to adults, with no significant difference between adjoining age cohorts, but the adult cohort was significantly better than the child cohort. These findings suggest that pitch processing ability

may improve with age, and that although important developments in the understanding and use of prosody continue during the school years (Cruttenden, 1985; Lyons et al., 2014; Wells et al., 2004), it is not yet adultlike for both ASD and TD participants.

Our finding of similar developmental changes in pitch discrimination ability across the ASD and control groups is incompatible with the markedly different developmental trajectories described by Mayer et al. (2016), where pitch discrimination ability increased with age in the control group but remained stable and enhanced across age cohorts in ASD. The discrepancy between the studies may be explained in several ways. First, there were differences in the paradigms used between the studies. The present study used an adaptive-tracking pitch threshold task to measure participants' pitch sensitivity starting with a default excursion size of six semitones, while Mayer et al. (2016) used stimulus pairs with either the same pitch or at a distance corresponding to 2, 3 or 6 semitones. Thus, the pitch variability of the stimuli used in their study was coarse, resulting in the task being easier than the one in the current study. Second, in Mayer et al. (2016), participants' ages overlapped between the child cohort (between 6 years 11 months and 14 years 9 months) and the adolescent cohort (between 9 years 8 months and 16 years 5 months), with both cohorts including intellectually lower-functioning ASD individuals, while the adult groups were all intellectually high-functioning. In our study, in order to match groups for age, sex and cognitive capability, all participants with ASD were intellectually high-functioning individuals and our age cohorts were defined with adults ≥ 16 , adolescents between 12–15, and children between 7–11 years. Finally, the data in Mayer et al. (2016) were collected and combined from three separate studies, which may have also affected the results.

Furthermore, in contrast to preadolescents with ASD whose language abilities fall within the average range for their age, those with language impairment showed developmental delays in the perception and production of statements and questions (Lyons et al., 2014). Our finding of no group difference in intonation discrimination, identification and imitation across age cohorts is in line with the results from preadolescents and adolescents with ASD who had age-appropriate language in Lyons et al. (2014). While there are many different ways to categorize age cohorts (Ahmad et al., 2009; Nithyashri & Kulanthaivel, 2012), the present study followed the division methods in the Autism-Spectrum Quotient (Auyeung et al., 2008; Baron-Cohen et al., 2001, 2006). One limitation of using this three-way split is that the age differences are not fine-grained, which may not be sufficient to detect subtle developmental changes over time. Hence, while we observed developmental changes from children to adults in both ASD and TD, further studies are required to use more fine-grained classifications of age

cohorts together with larger sample sizes, in order to map detailed developmental trajectories of pitch and intonation processing abilities in both groups.

Implications of the Current Research Findings

The current study has a basic research focus with the aim of determining whether, and how, sensitivity to pitch and prosody is affected in a group of participants with ASD when compared to a matched control group. In doing so, however, the study is also directly investigating the responses of individuals with ASD to social stimuli, as speech is an inherently social act. As social stimuli can sometimes be somewhat aversive to children with ASD, possibly by virtue of their unpredictability, nonverbal children with ASD typically require hundreds of hours of generalized imitation training (e.g. object imitation, gross-motor imitation, oral-motor imitation and vocal imitation) and echoic/vocal mand training to improve their spoken language skills (Hampton & Kaiser, 2016). Even so, interventions are not always successful, not least because of other comorbidities in such children which may interfere with their ability to perceive, acquire, or reproduce speech (Tager-Flusberg & Kasari, 2013). For such children, our results have a number of consequences.

Firstly, our data confirm that, absent such comorbidities and assuming sufficient general cognitive capability, there is no *a priori* reason to suppose that speech processing and production mechanisms are impaired when actively processing speech and therefore the perceptual, cognitive and motor mechanisms are likely to be in place to support any speech-therapeutic programme which may be indicated. Secondly, there is no evidence for a dissociation between pitch processing in music and speech in our investigations. The data therefore support the use of musical stimuli (which may be less aversive to some individuals) as scaffolding for training in pitch-based discrimination and imitation for individuals who may be more attracted to music than to language. Note however, both that our sample were not representative of children with ASD and learning difficulties, and that appropriate translational studies to confirm the generalisability of training in pitch in music to perception and production of pitch in speech are beyond the scope of the current investigation. Thirdly, despite the generally equivalent performance across groups in our tasks—and the lack of a response bias which might otherwise complicate interpretations on some tasks—pitch discrimination thresholds are elevated in the ASD group relative to the control group. This identifies and highlights a particular perceptual problem in this group although, somewhat surprisingly, not one that impacted upon performance in other tasks despite the overall negative correlations between pitch threshold and intonation discrimination in both music and speech.

Conclusion

In the present study, an experimental, acoustics-based approach was used to investigate perception and production of prosody in ASD, facilitating objective comparisons between the two modalities of intonation. In addition, we examined intonation processing in ASD and TD from the perspectives of response bias, task condition (discrimination, identification, imitation), stimulus type (speech, music), developmental changes, and its association with pitch thresholds. Our study revealed that intonation discrimination (in both speech and music conditions), identification and imitation abilities may be intact in some individuals with ASD across age cohorts (children, adolescents and adults), although children with ASD tend to have elevated pitch direction discrimination thresholds than their typically developing counterparts. The ASD and control groups also showed a similar developmental improvement in pitch thresholds, intonation discrimination and identification, as well as imitation. Furthermore, intonation identification and imitation are associated in both individuals with ASD and their peers, suggesting that improvements in intonation comprehension may also contribute to intonation production, and vice versa. We also found an association between low-level pitch threshold and high-level intonation processing across all participants, which reflects that the degree of strength or impairment in low-level pitch processing may influence performance on language acquisition and/or communication skills, whereas this bottom-up effect is more pronounced in controls than in individuals with ASD. In summary, our findings provide evidence for shared mechanisms in pitch processing between speech and music, as well as associations between low- and high-level pitch processing and between perception and production of pitch in prosody in individuals with and without ASD, who also show similar developmental trajectories for these abilities. Further studies with individuals with ASD from different cultures, particularly in other languages, would be helpful in obtaining a more comprehensive understanding of these shared mechanisms in ASD.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10803-021-05220-4>.

Acknowledgments This work was supported by a European Research Council (ERC) Starting Grant, ERC-StG-2015, CAASD, 678733, to F.L. and C.J. We thank Alex Bacon, Florence Leung, Caitlin Dawson and Queenie Cheng for assistance with data collection, and Caitlin Dawson and Georgia Matthews for help with data pre-processing.

Author Contributions FL conceived the study design, provided the material, and supervised data collection in the lab. LW analysed the data and performed statistical analysis with supervision from CPB. An initial draft of the article was written by LW and rewritten by FL based upon contributions from LW, CPB and CJ. All authors commented on previous versions of the manuscript, and read and approved the final manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

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](https://doi.org/10.1016/S0160-2896(01)00066-6)
- Ahmad, O. B., Boschi-pinto, C., Lopez, A. D., Lozano, R., & Inoue, M. (2009). Age standardization of rates: A new WHO standard. Geneva: World Health Organization; 2000 (GPE Discussion Paper Series No. 31)
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. Washington: American Psychiatric Association.
- Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Allison, C. (2008). The autism spectrum quotient: children's version (AQ-Child). *Journal of Autism and Developmental Disorders*, 38(7), 1230–1240. <https://doi.org/10.1007/s10803-007-0504-z>
- Ayotte, J., Peretz, I., & Hyde, K. L. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain: A Journal of Neurology*, 125(Pt 2), 238–251.
- Baron-Cohen, S., Hoekstra, R. A., Knickmeyer, R., & Wheelwright, S. (2006). The autism-spectrum quotient (AQ)—adolescent version. *Journal of Autism and Developmental Disorders*, 36(3), 343–350. <https://doi.org/10.1007/s10803-006-0073-6>
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31(1), 5–17. <https://doi.org/10.1023/A:1005653411471>
- Begley, C. G., & Ioannidis, J. P. A. (2015). Reproducibility in science: improving the standard for basic and preclinical research. *Circulation Research*, 116(1), 116–126. <https://doi.org/10.1161/CIRCRESAHA.114.303819>
- Bertone, A., Hanck, J., Kogan, C., Chaudhuri, A., & Cornish, K. (2010). Using perceptual signatures to define and dissociate condition-specific neural etiology: autism and fragile X syndrome as model conditions. *Journal of Autism and Developmental Disorders*, 40(12), 1531–1540. <https://doi.org/10.1007/s10803-010-1109-5>
- 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>
- Bijou, S. W., & Ghezzi, P. M. (1999). The behavior interference theory of autistic behavior in young children. In P. M. Ghezzi, W. L. Williams, & J. E. Carr (Eds.), *Autism: Behavior Analytic Perspectives* (pp. 33–43). Context Press.
- Boersma, P., & Weenink, D. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5, 341–345.
- 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>
- 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*. <https://doi.org/10.1177/0301006617718715>
- Cruttenden, A. (1985). Intonation comprehension in ten-year-olds*. *Journal of Child Language*, 12(3), 643–661. <https://doi.org/10.1017/S030500090000670X>
- Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge University Press.
- Dahan, D. (2015). Prosody and language comprehension. *Wires Cognitive Science*, 6(5), 441–452. <https://doi.org/10.1002/wcs.1355>
- DePape, A.-M.R., Chen, A., Hall, G. B. C., & Trainor, L. J. (2012). Use of prosody and information structure in high functioning adults with autism in relation to language ability. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2012.00072>
- Diehl, J. J., & Paul, R. (2012). Acoustic differences in the imitation of prosodic patterns in children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 6(1), 123–134. <https://doi.org/10.1016/j.rasd.2011.03.012>
- Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2014.00781>
- Fancourt, A., Dick, F., & Stewart, L. (2013). Pitch-change detection and pitch-direction discrimination in children. *Psychomusicology: Music, Mind, and Brain*, 23(2), 73.
- 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>
- Fosnot, S. M., & Jun, S.-A. (1999). Prosodic characteristics in children with stuttering or autism during reading and imitation. *Proceedings of the 14th International Congress of Phonetic Sciences*, 1925–1928.
- Francis, A. L., & Ciocca, V. (2003). Stimulus presentation order and the perception of lexical tones in Cantonese. *The Journal of the Acoustical Society of America*, 114(3), 1611–1621.
- 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>
- Germain, E., Foster, N. E. V., Sharda, M., Chowdhury, R., Tryfon, A., Doyle-Thomas, K. A. R., Anagnostou, E., & Hyde, K. L. (2019). Pitch direction ability predicts melodic perception in autism. *Child Neuropsychology*, 25(4), 445–465. <https://doi.org/10.1080/09297049.2018.1488954>
- Hampton, L. H., & Kaiser, A. P. (2016). Intervention effects on spoken-language outcomes for children with autism: A systematic review and meta-analysis. *Journal of Intellectual Disability Research*, 60(5), 444–463. <https://doi.org/10.1111/jir.12283>
- 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. (2003). Pitch memory, labelling and disembedding in autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 44(4), 543–551.
- 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>
- Horváth, J., Czigler, I., Birkás, E., Winkler, I., & Gervai, J. (2009). Age-related differences in distraction and reorientation in an auditory task. *Neurobiology of Aging*, 30(7), 1157–1172. <https://doi.org/10.1016/j.neurobiolaging.2007.10.003>
- Howell, D. C. (2009). *Statistical methods for psychology (7th edition)*. CengageLearning.
- Hübscher, I., & Prieto, P. (2019). Gestural and prosodic development act as sister systems and jointly pave the way for children's sociopragmatic development. *Frontiers in Psychology*, 10, 1259. <https://doi.org/10.3389/fpsyg.2019.01259>
- Hutchins, S., & Peretz, I. (2012). Amusics can imitate what they cannot discriminate. *Brain and Language*, 123(3), 234–239. <https://doi.org/10.1016/j.bandl.2012.09.011>
- Jamey, K., Foster, N. E. V., Sharda, M., Tuerk, C., Nadig, A., & Hyde, K. L. (2019). Evidence for intact melodic and rhythmic perception in children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, 64, 1–12. <https://doi.org/10.1016/j.rasd.2018.11.013>
- 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. (2008a). 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>
- Järvinen-Pasley, A., Wallace, G. L., Ramus, F., Happé, F., & Heaton, P. (2008b). Enhanced perceptual processing of speech in autism. *Developmental Science*, 11(1), 109–121. <https://doi.org/10.1111/j.1467-7687.2007.00644.x>
- JASP Team. (2020). JASP (Version 0.13.1) [Computer software].
- Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2010). Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese. *Neuropsychologia*, 48(9), 2630–2639. <https://doi.org/10.1016/j.neuropsychologia.2010.05.009>
- 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>
- Kessels, R. P. C., van Zandvoort, M. J. E., Postma, A., Kappelle, L. J., & de Haan, E. H. F. (2000). The corsi block-tapping task: standardization and normative data. *Applied Neuropsychology*, 7(4), 252–258. https://doi.org/10.1207/S15324826AN0704_8
- Koelsch, S. (2011). Toward a Neural Basis of Music Perception – A Review and Updated Model. *Frontier in Psychology*. <https://doi.org/10.3389/fpsyg.2011.00110>
- Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, 9(12), 578–584. <https://doi.org/10.1016/j.tics.2005.10.001>
- Kwok, E. Y. L., Brown, H. M., Smyth, R. E., & Oram Cardy, J. (2015). Meta-analysis of receptive and expressive language skills in autism spectrum disorder. *Research in Autism Spectrum Disorders*, 9, 202–222. <https://doi.org/10.1016/j.rasd.2014.10.008>
- Levitt, A. G. (1993). The Acquisition of Prosody: Evidence from French- and English-Learning Infants. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life* (pp. 385–398). Springer.
- Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America*, 32(4), 451–454. <https://doi.org/10.1121/1.1908095>
- 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., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: Discrimination, identification and imitation. *Brain*, 133(6), 1682–1693. <https://doi.org/10.1093/brain/awq089>
- 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>
- Lloyd-Fox, S., Blasi, A., Elwell, C. E., Charman, T., Murphy, D., & Johnson, M. H. (2013). Reduced neural sensitivity to social stimuli in infants at risk for autism. *Proceedings of the Royal Society b: Biological Sciences*, 280(1758), 20123026. <https://doi.org/10.1098/rspb.2012.3026>
- Loui, P., Guenther, F. H., Mathys, C., & Schlaug, G. (2008). Action-perception mismatch in tone-deafness. *Current Biology : CB*, 18(8), R331–R332. <https://doi.org/10.1016/j.cub.2008.02.045>
- Lyons, M., Simmons, E. S., & Paul, R. (2014). Prosodic development in middle childhood and adolescence in high-functioning autism. *Autism Research*, 7(2), 181–196. <https://doi.org/10.1002/aur.1355>
- 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>
- Martin, N. A., & Brownell, R. (2011). *Receptive One-Word Picture Vocabulary Test, 4th Edition (ROWPVT-4)*. Academic Therapy Publications.
- MATLAB. (2010). *Version 7.10.0 (R2010a)*. Natick, Massachusetts: The MathWorks Inc.
- 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>
- 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(Spring), 120–132. https://doi.org/10.1044/cicsd.41_S_120
- 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>
- 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>
- McDermott, J. H., & Oxenham, A. J. (2008). Music perception, pitch, and the auditory system. *Current Opinion in Neurobiology*, 18(4), 452–463. <https://doi.org/10.1016/j.conb.2008.09.005>
- Moore, D. R., Ferguson, M. A., Halliday, L. F., & Riley, A. (2008). Frequency discrimination in children: Perception, learning and attention. *Hearing Research*, 238(1), 147–154. <https://doi.org/10.1016/j.heares.2007.11.013>
- Mottron, L., Peretz, I., & Ménard, E. (2000). Local and global processing of music in high-functioning persons with autism: Beyond central coherence? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 41(8), 1057–1065.

- 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>
- Nithyashri, J., & Kulanthaivel, G. (2012). Classification of human age based on neural network using FG-NET aging database and Wavelets. *Fourth International Conference on Advanced Computing (ICoAC)*, 2012, 1–5. <https://doi.org/10.1109/ICoAC.2012.6416855>
- Patel, A. D. (2008). *Music, Language, and the Brain*. Oxford University Press.
- 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>
- Patel, R., & Grigos, M. I. (2006). Acoustic characterization of the question–statement contrast in 4, 7 and 11 year-old children. *Speech Communication*, 48(10), 1308–1318. <https://doi.org/10.1016/j.specom.2006.06.007>
- 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>
- 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\(2007/071\)](https://doi.org/10.1044/1092-4388(2007/071))
- Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 6(7), 688–691. <https://doi.org/10.1038/nn1083>
- Pfordresher, P. Q., & Halpern, A. R. (2013). Auditory imagery and the poor-pitch singer. *Psychonomic Bulletin & Review*, 20(4), 747–753. <https://doi.org/10.3758/s13423-013-0401-8>
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111–163. <https://doi.org/10.2307/271063>
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven manual: Section 3*. Oxford Psychologists Press.
- Sammler, D., Koelsch, S., Ball, T., Brandt, A., Elger, C. E., Friederici, A. D., Grigutsch, M., Huppertz, H.-J., Knösche, T. R., Wellmer, J., Widman, G., & Schulze-Bonhage, A. (2009). Overlap of musical and linguistic syntax processing: intracranial ERP evidence. *Annals of the New York Academy of Sciences*, 1169(1), 494–498. <https://doi.org/10.1111/j.1749-6632.2009.04792.x>
- Schelinski, S., Roswadowski, C., & von Kriegstein, K. (2017). Voice identity processing in autism spectrum disorder. *Autism Research*, 10(1), 155–168. <https://doi.org/10.1002/aur.1639>
- Shriberg, L. D., Paul, R., Black, L. M., & van Santen, J. P. (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>
- 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>
- Steffens, T., Steffens, L. M., & Marcrum, S. C. (2020). Chance-level hit rates in closed-set, forced-choice audiometry and a novel utility for the significance test-based detection of malingering. *PLoS One*, 15(4), e0231715. <https://doi.org/10.1371/journal.pone.0231715>
- Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., & Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *The Journal of Neuroscience*, 34(3), 691–697. <https://doi.org/10.1523/JNEUROSCI.3615-13.2014>
- Tager-Flusberg, H., & Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum. *Autism Research: Official Journal of the International Society for Autism Research*, 6(6), 468–478. <https://doi.org/10.1002/aur.1329>
- Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Aud. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of Autism and Pervasive Developmental Disorders* (1st ed., pp. 335–364). Wiley.
- 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>
- Vuvan, D. T., Nunes-Silva, M., & Peretz, I. (2015). Meta-analytic evidence for the non-modularity of pitch processing in congenital amusia. *Cortex*, 69, 186–200. <https://doi.org/10.1016/j.cortex.2015.05.002>
- Wechsler, D. (2003). *Wechsler intelligence scale for children—Fourth Edition (WISC-IV)*. The Psychological Corporation.
- Wells, B., Peppé, S., & Goulondris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language*, 31(4), 749–778. <https://doi.org/10.1017/S030500090400652X>
- Xie, X., Buxó-Lugo, A., & Kurumada, C. (2021). Encoding and decoding of meaning through structured variability in intonational speech prosody. *Cognition*, 211, 104619. <https://doi.org/10.1016/j.cognition.2021.104619>
- Xu, Y. (2013). ProsodyPro—A tool for large-scale systematic prosody analysis. *Proceedings of Tools and Resources for the Analysis of Speech Prosody*, 7–10.
- 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>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.