Research in Developmental Disabilities 37 (2015) 162–170

Contents lists available at ScienceDirect

Research in Developmental Disabilities

Emotional prosody perception and its association with pragmatic language in school-aged children with high-function autism

Jia-En Wang [a,b], Feng-Ming Tsao [a,*]

[a] Department of Psychology, National Taiwan University, Taiwan
[b] Department of Psychiatry, Mackay Memorial Hospital, Taiwan

ARTICLE INFO

Article history:
Received 28 October 2014
Received in revised form 12 November 2014
Accepted 15 November 2014
Available online

Keywords:
High function autism
Prosody perception
Pragmatics
Social adaptation
School-aged children

ABSTRACT

Emotional prosody perception is essential for social communication, but it is still an open issue whether children with high-function autism (HFA) exhibit any prosodic perception deficits or experience selective impairments in recognizing the prosody of positive emotions. Moreover, the associations between prosody perception, pragmatic language, and social adaptation in children with HFA have not been fully explored. This study investigated whether emotional prosody perception for words and sentences in children with HFA ($n = 25$, 6–11 years of age) differed from age-matched, typically developing children (TD, $n = 25$) when presented with an emotional prosody identification task. The Children's Communication Checklist and Vineland Adaptive Behavior Scale were used to assess pragmatic and social adaption abilities. Results show that children with HFA performed poorer than TD children in identifying happy prosody in both emotionally neutral and relevant utterances. In contrast, children with HFA did not exhibit any deficits in identifying sad and angry prosody. Results of correlation analyses revealed a positive association between happy prosody identification and pragmatic function. The findings indicate that school-aged children with HFA experience difficulties in recognizing happy prosody, and that this limitation in prosody perception is associated with their pragmatic and social adaption performances.

1. Introduction

Social communication deficit is one of the core symptoms in children with autistic spectrum disorder (DSM-V, 2013). The ability to correctly and effectively recognize a speaker's emotions from their utterances is essential to social communication (Lindner & Rosén, 2006; Shriberg et al., 2001). In addition to semantic meaning, emotional prosody of utterances also manifests a speaker's emotions during social communication. Deficits in producing appropriate emotional prosody in storytelling and conversations are well documented for children with autism (e.g., DSM-V, 2013), but it is still an open issue whether perceptual deficits in emotional prosody play any role in precipitating these deficits in producing intact emotional prosody.

* Corresponding author at: Department of Psychology, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei, Taiwan. Tel.: +886 2 33664467.
E-mail address: tsaosph@ntu.edu.tw (F.-M. Tsao).

http://dx.doi.org/10.1016/j.ridd.2014.11.013
0891-4222/

1.1. Emotional prosody perception in high function autism

Prosody is a linguistic term that includes intonation, stress, and rhythm (McCann, Peppé, Gibbon, O'Hare, & Rutherford, 2007). It plays an important role in linguistic functions as well as in emotional expression and comprehension (Lindner & Rosén, 2006; Paul, Augustyn, Klin, & Volkmar, 2005). Deficits in prosody perception might result in atypical prosody production in autism, and also associate with difficulties in decoding emotional signals during social communication (Paul, Augustyn, et al., 2005). During social interactions, the ability to integrate emotional prosody and emotional semantics of a speaker's utterances is crucial to understanding the emotional state and communicative intentions of the speaker as it allows the listener to reciprocate with appropriate responses. Thus, the ability to comprehend emotional prosody in speech is a critical function of human social life (Wittfoth et al., 2010). Many studies have examined autistic children's abilities to recognize emotions through facial expressions; however, relatively few studies have been conducted to examine autistic children's abilities to understand emotions through prosody (Lindner & Rosén, 2006; McCann & Peppé, 2003). Further, while there is extensive evidence in the literature of impairment in the expressive prosodic abilities of autistic individuals, the state of receptive prosodic abilities has been neglected (Le Sourn-Bissaoui, Aguert, Girard, Chevreuil, & Laval, 2013; McCann & Peppé, 2003). Among studies on prosody perception in adults, adolescents, or children with high-function autism (HFA) or Asperger syndrome (AS), findings of deficits in prosody perception have been equivocal.

In the past decade, many studies have reported that adults or children with autism experienced difficulties in using emotional prosody to identify the emotions of others (Golan, Baron-Cohen, Hill, & Rutherford, 2007; Hubbard & Trauner, 2007; Järvinen-Pasley, Peppé, King-Smith, & Heaton, 2008; Lindner & Rosén, 2006; McCann et al., 2007; Peppé, McCann, Gibbon, O'Hare, & Rutherford, 2007; Rutherford, Baron-Cohen, & Wheelwright, 2002). In contrast, several studies have reported that the emotional prosody perception in autistic individuals is similar to that in normal controls (Brennand, Schepman, & Rodway, 2011; Brooks & Ploog, 2013; Chevallier, Noveck, Happe, & Wilson, 2011; Grossman, Bemis, Skwerer, & Tager-Flusberg, 2010; Paul, Augustyn, et al., 2005). Therefore, it remains an open issue whether autistic children exhibit any difficulties in recognizing emotions through speech prosody.

1.2. Methodological and age differences in previous studies

Inconsistent findings may occur as a consequence of small participant sample sizes (Hubbard & Trauner, 2007; Lindner & Rosén, 2006). In addition, the use of relatively large age ranges, e.g., from early childhood to adolescence, make it difficult to compare results between different studies to determine whether people with autism show any developmental differences in prosodic processing at certain ages (Brooks & Ploog, 2013; Golan et al., 2007; Grossman & Tager-Flusberg, 2012; Grossman et al., 2010; Hubbard & Trauner, 2007; Järvinen-Pasley et al., 2008). The ability to perceive emotional prosody starts from infancy and continues to develop during middle childhood. Sakkalou and Gattis (2012) showed that 14- to 18-month-old infants infer the intentions of other people using prosodic cues. Sensitivity to emotional prosody continues to improve in typically developing (TD) school-aged children. In a study by Friend and Bryant (2000), TD children aged 4, 7, and 10 were presented with utterances containing semantic content of positive or negative emotion in combination with happy or angry prosody, and asked to judge whether the speaker was happy or angry. Their results showed that 4- and 7-year-old children weighted semantic meaning more than emotional prosody, whereas 10-year-old children valued prosodic information more than semantic meaning. Another study, which compared the relative contribution of emotional prosody versus semantic content in judging a speaker's emotion, found that adults relied exclusively on prosody, whereas 4-year-old children responded primarily to semantic content (Morton & Trehub, 2001). These studies showed a developmental trend where children between 5 and 10 years of age gradually reduced their reliance on the semantic content of utterances to identify emotions. In other words, emotional prosody sensitivity improves in elementary school-aged children; this period is critical for the development of relative weighting between prosody and semantic content in emotion identification. Thus, assessing emotional prosody perception in elementary school-aged children with HFA would be essential to determine whether children with HFA exhibit any emotional prosody perception deficits when typically developing children are still developing their emotional prosody perception during the middle childhood.

The variety of emotional prosody tasks employed in previous studies is another reason for inconsistent findings (Golan et al., 2007). Previous studies have presented emotional prosody to people with autism in different linguistic units, such as words and sentences. In addition to stimulus duration, these stimuli also varied in many aspects, such as sufficiency of contextual information, that would influence emotion identification (Paulmann & Pell, 2010; Peppé et al., 2007). For example, the profiling elements of prosody in speech-communication (PEPS-C) test requires children to distinguish a food item that a person likes or dislikes, based on the emotional prosody of a *single word* (Peppé et al., 2007). Studies adopting the PEPS-C to assess emotional prosody perception in children with HFA have demonstrated impairments in this population (McCann et al., 2007; Peppé et al., 2007). In contrast, studies presenting *sentences* as stimuli to assess emotional prosody perception in children with HFA have reported that it is similar to that in TD children (Bertrand & Priego-Valverde, 2011; Chevallier et al., 2011; Grossman et al., 2010; Paul, Augustyn, et al., 2005).

Further, studies have also utilized variations in emotional semantics of speech stimuli, i.e., using stimuli with either emotionally neutral content (e.g., *fruit name* or *I have a pencil*) or emotionally relevant content (e.g., *hate* or *I am very glad to see you*). When a listener identifies emotions from speech stimuli, emotionally neutral and emotionally relevant utterances provide different emotional semantic cues that require different levels of cognitive processing (e.g., execution function). For

emotionally neutral utterances, speech prosody would be the only valid cue used to identify emotions. In contrast, for emotionally relevant utterances, both semantic and prosodic cues are valid cues to identify emotions. When both cues are incongruent in recognizing emotions, for instance when the word *happy* is presented with the *sad* prosody, the processing of cognitive inhibition is mobilized to correctly recognize emotions through emotional prosody. Thus, variation in the stimuli used by studies on emotional prosody perception in children with HFA—some studies using only emotionally neutral stimuli (Grossman & Tager-Flusberg, 2012; Hubbard & Trauner, 2007; Järvinen-Pasley et al., 2008; McCann et al., 2007), some studies using only emotionally relevant stimuli (Golan et al., 2007; Rutherford et al., 2002), and some studies using a mixture (Le Sourn-Bissaoui et al., 2013; Singh & Harrow, 2014)—might be another reason for inconsistent findings.

1.3. The relationships between emotional prosody, pragmatics, and social adaptation

Atypical prosody perception may explain social communication deficits in children with autism (Shriberg et al., 2001). Pragmatic impairment in social communication might be defined as a mismatch between language and the situation in which it is used, making the language inappropriate to situational demands (Vonden & Lord, 1991). People with pragmatic impairment experience difficulties with using verbal and non-verbal language. Pragmatic deficits are specifically and universally impaired in autism (Young, Diehl, Morris, Hyman, & Bennetto, 2005), and the ability to identify emotional intonation in utterances is essential to pragmatics and social communication (O'Connor, 2012; Wilson & Wharton, 2006). However, only few studies have examined the relationship between emotional prosody perception and pragmatic function in children with HFA (McCann et al., 2007; Paul, Shriberg, et al., 2005; Volden, Coolican, Garon, White, & Bryson, 2009; Wang, Lee, Sigman, & Dapretto, 2006). In brief, the relationships between prosody perception, pragmatics, and social adaptation have rarely been explored in children with autism. Therefore, this study also attempted to assess the relationship between emotional prosody perception and linguistic pragmatics in children with HFA.

1.4. Specific aims

To control the effects of methodological differences on assessing prosody perception in children with HFA, both words and short sentences were used in this study; further, these speech stimuli also varied in their emotional meaning and prosody. This design would lead to a better understanding of emotional prosody perception in elementary school-aged children with HFA.

This study had two goals. First, it aimed to investigate whether school-aged children with HFA were less accurate than TD children in using emotional prosody information to judge emotions when emotionally neutral and emotionally relevant stimuli were presented in words and short sentences. If children with HFA exhibited atypical emotional prosody perception, then they were anticipated to be less accurate than TD children in judging emotions. The other goal of this study was to assess the association between emotional prosody perception and social communication skills in children with HFA. If deficits in emotional prosody perception adversely affected social interaction skills in children with HFA, reduced perceptual performance was expected to associate with poor social interaction skills.

2. Materials and methods

2.1. Participants

Participants consisted of 25 boys with HFA between 6 and 11 years of age (mean age = 8 years, 2 mo.) and 25 TD boys between 6 and 11 years of age (mean age = 8 years, 2 mo.). The Chinese version of the Wechsler Intelligence Scale for Children, 4th ed. (WISC-IV; Wechsler, 2003) was used to assess the general cognitive abilities of the participants. Children with HFA were diagnosed according to DSM-IV-TR criteria (APA, 2000) and were recruited from child psychiatry clinics at a medical center in Taiwan. A senior child psychiatrist and child psychologist independently confirmed the clinical diagnosis of HFA. The Childhood Asperger Syndrome Test (CAST) (Scott, Baron-Cohen, Bolton, & Brayne, 2002) was also used to confirm the diagnosis of autism. The inclusion criteria for children with HFA or AS were a CAST score $\geq$15, a Full Scale IQ (FSIQ) of WISC-IV $>$90, and a Verbal Comprehension Index (VCI) of WISC-IV $>$80.

The 25 TD children were recruited as the control group from elementary schools in the Taipei metropolitan area in Taiwan. They were screened for autism by using the CAST, and no children scored above the cut-off score for autism. The TD children in the control group were also matched to the chronological ages and FSIQ of WISC-IV of autistic children. Table 1 lists the background information of both TD children and children with autism spectrum disorder (ASD).

2.2. Materials

2.2.1. Emotional-prosody identification task

The experiment tested three basic emotions: happy, sad, and angry. The stimuli for emotional prosody identification were delivered using two types of linguistic stimuli—words and short sentences—and with variation in emotional semantics: emotionally relevant (i.e., happy, sad, and angry) and neutral meaning. A professional voice actress recorded the speech

Table 1
Descriptive characteristics of the participants.

	HFA ($n = 25$)		TD ($n = 25$)		F	p
	Mean	(SD)	Mean	(SD)		
CA (mo.)	97.84	14.01	98.36	12.49	.019	.890
CAST	19.96	4.01	5.30	2.79	212.938	.000
FSIQ	107.12	11.14	112.96	9.91	3.835	.056
VCI	107.08	10.57	113.68	8.47	5.936	.019
PRI	11.72	14.15	113.08	15.05	.108	.743
WMI	108.50	17.31	107.16	10.77	.098	.756
PSI	92.44	18.10	103.72	10.64	6.851	.014
CCC-Pragmatic index	113.68	8.78	150.70	6.78	263.778	.000
VABS-Communication	33.63	18.29	69.52	21.95	37.232	.000
VABS-Socialization	18.13	15.45	67.17	26.66	58.261	.000

stimuli in a quiet room with a digital recorder (Marantz PMD66). Each speech stimulus recorded (sampling rate = 44 kHz, resolution = 16 bits) was then selected and equalized to the RMS amplitude using Sound Forge 7.0 sound editing software (Sony Pictures Digital Inc., 2004).

We conducted a pilot study to confirm whether the emotional prosody of speech stimuli was perceptually salient and allowed listeners to correctly identify the emotions of the speaker. The speech stimuli in the pilot study used 15 emotionally relevant words (e.g., *glad*, *sad*, and *hate*) and 5 emotionally neutral words (i.e., names of fruits). In addition, the speech stimuli also included 9 short emotional sentences that contained six syllables (e.g., *wǒ jué de bù shū fú/I don't feel well*) and 5 emotionally neutral short sentences. Each word and sentence was spoken with three emotional prosodies (a happy, sad, or angry tone) and randomly presented three times. The pilot experiment included 306 testing trials (20 spoken words × 3 emotional tones × 3 repetitions and 14 sentences × emotional tones × 3 repetitions) that required approximately 30–40 min to complete. The undergraduate students ($n = 40$) participating in the pilot study wore earphones (Philips shp-1900) and were instructed to identify emotions through emotional prosody by listening to speech stimuli played out from a laptop computer (HP Pavilion dv2000). Participants also rated the salience of each stimulus on a 7-point scale, where 1 represented an obscure prosody feature and 7 represented the most obvious prosody feature. The pilot experiment consisted of a practice phase and a test phase that contained six sessions (three words and three short sentences).

The emotional semantics and prosody of speech stimuli can be either congruent (e.g., *happy* with happy prosody) or incongruent (e.g., *happy* with sad prosody). In our study, in order for a speech stimulus to be selected for the subsequent study on emotional prosody perception in children with HFA, it needed to satisfy the conditions of emotion identification accuracy of ≥80% (for prosodic/semantic congruent stimuli) or ≥70% (for incongruent stimuli). The average salience rating of each stimulus needed to be ≥4.

Based on results of the adult pilot study, the experimental stimuli for children consisted of 12 words describing happy, sad, angry, and neutral situations and 10 short sentences describing happy ($n = 2$), sad ($n = 2$), angry ($n = 2$), and neutral ($n = 4$) situations. Each stimulus was spoken with three emotional prosodies (a happy, sad, or angry tone) and randomly presented three times. In total, 198 trials were conducted in this study (12 words × 3 emotional tones × 3 repetitions and 10 sentences × 3 emotional tones × 3 repetition). The presentation sequence of each word or sentence block was randomized across participants.

2.2.2. Pragmatic and social adaptive abilities

Parents of children completed the Chinese version of the Vineland Adaptive Behavior Scale (VABS; Sparrow, Balla, & Cicchetti, 1984; Wu, Zhang, Lu, & Qiu, 2004) and the Children's Communication Checklist (Bishop, 1998; Bishop & Baird, 2001) to rate the pragmatic functions of the children. The CCC consists of 70 items grouped into 9 sub-scales by which teachers or parents rate child behavior. A composite score of pragmatic language behavior was then derived from the checklist.

2.2.3. Language abilities

Language abilities were assessed using the Chinese Language Development Test for School-aged Children (CLDSC; Lin, Huang, Huang, & Xuan, 2009). The CLDSC was designed to assess both language production and comprehension abilities in children between 6 and 12 years of age. The 32-item receptive language sub-test assesses vocabulary, syntax, and short story comprehension. The 23-item expressive language sub-test assesses naming, sentence repetition, sentence construction, and the narrative ability to tell short stories. Approximately 30–40 min were required to complete the language assessment.

2.3. Procedure

Parents were asked to complete the CAST, CCC, and VABS, and children were assessed with a battery of cognitive and language tests, including the WISC-IV and CLDSC. The emotional-prosody identification task was run after the children had completed their psychological tests.

During the prosody identification task, children were tested individually in a quiet room. They wore headphones and listened to speech stimuli played out through a laptop computer. They sat facing the monitor while the experimenter sat beside the child. E-prime 1.1 software (2004) was used to control the presentation of the speech stimuli and to record the child's responses. Before the experiment began, the experimenter instructed the children that they were going to play a computer game by listening to words and short sentences. The children were asked to judge the emotions of the speaker based on prosody, and then press a key corresponding to their judgment. The test included six blocks, with words in the first three blocks and short sentences in the last three blocks. The children could take a break between each block if necessary. In each practice trial, three cartoon figures with emotional faces (happy, sad, and angry) were displayed to remind children to identify the emotion of the speaker. Before the practice phase started, the experimenter asked the children whether they were able to identify correctly the emotions of each cartoon figure to ensure that the children could correctly recognize emotions from the faces displayed. After completing the practice trials, the test trials started, which required approximately 20–30 min to complete. Children each received about US$ 15 for their participation.

3. Results

This study aimed to: (a) investigate the emotional prosody perception of 6- to 11-year-old children with HFA; and (b) assess the relationships between prosody perception, pragmatics, and social adaptation.

3.1. Emotional prosody perception in the HFA and TD groups

A two-way ANOVA was used with Group (HFA, Control) as a between-subject factor and Stimulus Type (word, sentence) as a within-subject factor to assess emotional prosody perception in children. The results showed the effect of Stimulus Type to be significant [$F(1, 47) = 42.145$, $p < .001$], but neither the effect of Group [$F(1, 47) = 3.641$, $p = .062$] nor the interaction of Group × Stimulus Type [$F(1, 47) = .171$, $p = .681$] was. A paired t-test was performed and the results indicated that the percentage of correct responses for short sentences was significantly higher than that for words in both the HFA [$t(24) = -4.263$, $p < .001$] and TD groups [$t(23) = -4.932$, $p < .001$].

Because emotional prosody is easier for children to identify in sentences than in words, separated mixed three-way ANOVAs were conducted with Group (HFA, Control) as a between-subject factor and Emotional Prosody (Happy, Sad, Angry) and Emotional Semantic (Neutral, Happy, Sad, Angry) as within-subject factors for the different linguistic stimuli employed (words, sentences, and combined (words + sentences)). The results showed a significant effect of Prosody in three conditions: word [$F(2, 46) = 47.996$, $p < .001$], sentence [$F(2, 46) = 29.200$, $p < .001$], and combined [$F(2, 46) = 51.533$, $p < .001$]. The effect of Emotional Semantics was significant under all conditions: word [$F(3, 45) = 10.313$, $p < .001$], sentence [$F(3, 45) = 5.062$, $p < .002$], and combined [$F(3, 45) = 9.105$, $p < .001$]. The effect of Group was not significant in the word [$F(1, 47) = 1.791$, $p = .187$] and combined conditions [$F(1, 47) = 3.641$, $p = .062$], however, it was significant in the sentence condition [$F(1, 47) = 4.529$, $p = .039$]. The interaction of Group × Prosody was significant in the word [$F(2, 46) = 3.461$, $p = .035$] and combined conditions [$F(2, 46) = 4.035$, $p = .021$], but not significant in the sentence condition [$F(2, 46) = 1.866$, $p = .16$]. The interaction of Group × Semantics was not significant over all linguistic stimulus conditions: word [$F(3, 45) = 1.027$, $p = .383$], sentence [$F(3, 45) = .371$, $p = .774$], and combined [$F(3, 45) = .853$, $p = .476$]; likewise, neither was the interaction of Group × Prosody × Semantics: word [$F(6, 42) = .137$, $p = .991$], sentence [$F(6, 42) = 1.989$, $p = .067$], and combined [$F(6, 42) = .776$, $p = .589$].

To further examine the interaction between emotional semantics and speech prosody in the identification of emotions in the two groups of children, separate one-way ANOVAs were conducted to determine in which emotional prosody condition(s) children with HFA were less accurate than TD children in judging the speaker's emotions. The results indicated that the HFA group performed poorer than the TD group in correctly judging emotions for emotionally neutral sentences with happy prosody [$F(1, 48) = 5.473$, $p = .024$], but not for emotionally neutral words [$F(1, 48) = 1.329$, $p = .255$]. Although children with HFA performed poorer than TD children when using the happy prosody to judge emotions for short sentences, the effect of Group was not significant when linguistic stimuli were present with sad and angry prosodies (Fig. 1).

Fig. 2 shows the percentages of correct emotional judgments when children listened to emotionally relevant utterances with various emotional prosodies. As shown in Fig. 2, the HFA group experienced difficulty identifying happy prosody for emotionally relevant words [$F(1, 48) = 7.166$, $p = .010$], but not for short sentences [$F(1, 48) = 1.835$, $p = .182$]. There were no significant differences by Group when children identified emotionally relevant utterances with sad and angry prosody.

Follow-up analyses showed that the HFA group judged emotions less accurately than the TD group when listening to both happy words with happy prosody [$F(1, 48) = 5.150$, $p = .028$] and sad words with happy prosody [$F(1, 48) = 13.098$, $p = .001$]. In addition, an effect of Group was also observed for angry sentences with happy prosody [$F(1, 48) = 6.086$, $p = .017$]. These results revealed that children with HFA experienced more difficulties than TD children in recognizing happy emotions from both emotional-prosody congruent and incongruent linguistic stimuli when they listened to words and sentences with happy prosody. Fig. 3 displays the results from the HFA and TD groups in identifying happy emotions from emotionally relevant words and sentences with happy prosody.

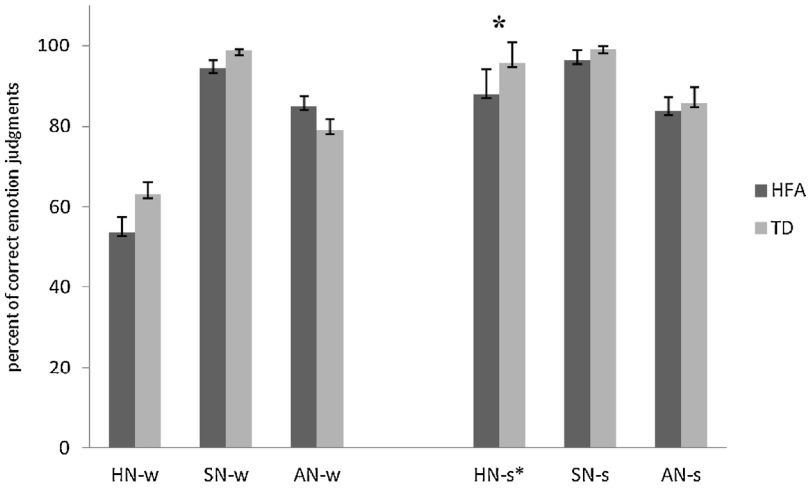

Fig. 1. Percentages of correct emotion judgments based on emotional prosody in the *neutral semantic* condition for the HFA and TD groups. *Note*: "HN" means happy prosody with neutral semantic; "SN" means sad prosody with neutral semantic; "AN" means angry prosody with neutral semantic; "w" means word stimuli; "s" means short sentence stimuli; "*" denotes a significant difference between groups.

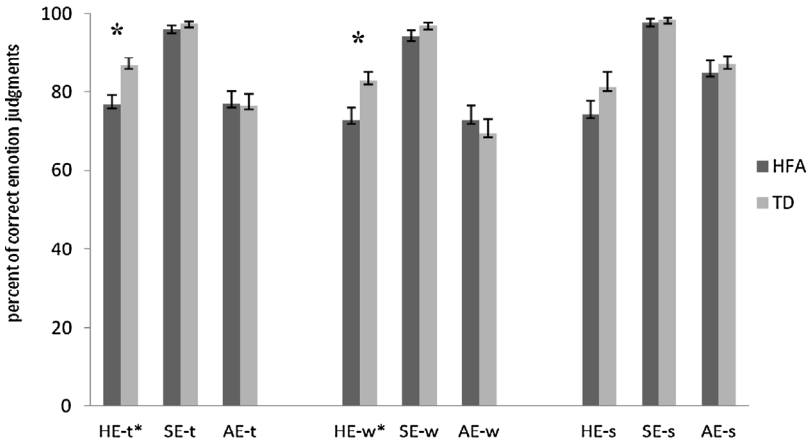

Fig. 2. Percentages of correct emotion judgments based on emotional prosody in the *emotionally relevant semantic* condition for the HFA and TD groups. *Note*: "HE" means happy prosody with emotionally relevant semantic; "SE" means sad prosody with emotionally relevant semantic; "AE" means happy prosody with emotionally relevant semantic; "t" means word + sentence stimuli; "w" means word stimuli; "s" means short sentence stimuli; "*" denotes a significant difference between groups.

3.2. Associations between prosody perception, pragmatics, and social adaption

The second goal of this study was to assess whether emotional prosody perception played any role in reduced social communication function in children with HFA by examining the associations between prosody perception, pragmatic language, and social adaptation. Combining data from TD children and children with HFA, correlation analyses were run on prosody, autistic symptoms (i.e., CAST score), pragmatic index (i.e., CCC), and the VABS indices (communication and socialization). Results showed that emotion judgment accuracies for emotionally neutral utterances with happy [$r(48) = -.322$, $p = .028$] and sad [$r(48) = -.381$, $p = .007$] prosody and for emotionally relevant utterances with happy [$r(48) = -.450$, $p = .002$] and sad [$r(48) = -.314$, $p = .030$] prosody were significantly correlated with CAST score. These results suggest that poor prosody perception is associated with severity of autistic symptoms. For pragmatic functions, significant correlations were shown between CCC score and emotionally neutral sentences [$r(48) = .326$, $p = .025$], happy sentences [$r(48) = .393$, $p = .006$], and angry sentences [$r(48) = .332$, $p = .022$] with happy prosody. The association between CCC score and happy words with happy prosody was also significant [$r(48) = .295$, $p = .042$]. These results suggest that children's accuracy in perceiving happy prosody is positively associated with their pragmatic ability. For social adaption ability, the results showed that sad sentences [$r(48) = .315$, $p = .033$] and angry sentences [$r(48) = .416$, $p = .004$] with happy prosody were significantly correlated with the VASB socialization index, suggesting that children performing better in perceiving happy prosody from emotional-prosody incongruent sentences are capable of better interpersonal relationships and coping skills.

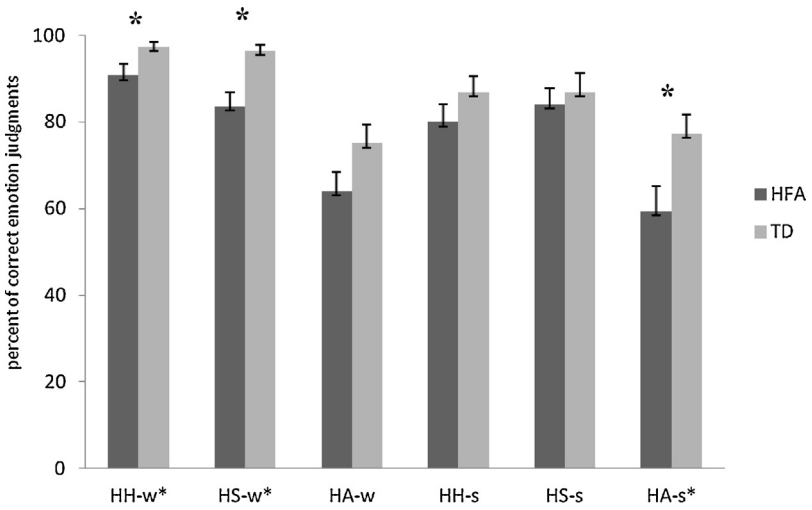

Fig. 3. Percentages of correct emotion judgments of *happy prosody* according to emotionally relevant utterance types for the HFA and TD groups. *Note*: "HH" means happy prosody with happy semantic; "HS" means happy prosody with sad semantic; "HA" means happy prosody with angry semantic; "w" means word stimuli; "s" means short sentence stimuli; "*" denotes significant difference between groups.

4. Discussion

This study used emotionally relevant and emotionally neutral speech stimuli to examine whether school-aged children with HFA displayed any deficits in using prosodic cues to identify emotions. The results of this study revealed that children with HFA generally performed similar to TD children in perceiving emotional prosody. However, children with HFA showed a selective impairment for perceiving specific emotions. When presented with prosodies of emotionally neutral and emotionally relevant utterances, children with HFA performed similar to TD children in recognizing prosodic cues of negative emotions (sad and angry), but they experienced more difficulty in recognizing positive emotions (happy).

Studies on emotional prosody perception of autistic individuals have yielded inconsistent findings: some have reported prosody deficits (Järvinen-Pasley et al., 2008; McCann et al., 2007; Rutherford et al., 2002), whereas other studies have reported intact prosody perception (Grossman & Tager-Flusberg, 2012; Paul, Augustyn, et al., 2005). Presumptive emotional prosody processing impairments in children with HFA might be task-dependent (Singh & Harrow, 2014). As previously discussed, methodological differences—such as different perceptual salience of prosody present with different speech materials (words vs. sentences)—and variation in participant age could result in inconsistent findings. This study controlled task demands by deliberately manipulating congruence between emotional prosody and semantics, as well as by changing the perceptual salience of emotional prosody through the use of both words and short sentences. In addition, we explored prosodic perception of children with HFA at an age range where TD children are still developing their ability to accurately recognize emotions through emotional prosody. Our results show that children with HFA performed poorer than TD children in categorizing happy emotions based on prosodic cues. In addition, children with HFA were more accurate in judging emotions when presented with sentences than with words.

Would the presentation sequence of words and sentences account for the poor performance of children with HFA in categorizing happy prosody? The performance differences between words and sentences in children with HFA might reflect that prosodic cues of sentences are perceptually more salient than those of words, and this condition in turn reduces the task demand for recognizing emotions. However, an alternative interpretation is that the experimental procedure took advantage of the sentential facilitation effect. To familiarize children with the procedure of focusing on the emotional prosody to identify emotions, the words were presented first in this study, and then followed by the sentences. However, results from paired *t*-tests showed that while the HFA group processed sentences more easily than words with sad [$t(24) = -2.341$, $p = .028$] and angry prosody [$t(24) = -4.226$, $p < .001$], this was not the case with happy prosody [$t(24) = -.378$, $p = .709$]. Thus, it appears that the presentation sequence of the words and sentences cannot serve as an alternative interpretation for the difficulty in processing happy prosody by children with HFA.

Recent studies have demonstrated similar findings of autistic children processing emotional prosody of positive emotions differently from TD children. Brooks and Ploog (2013) developed a video game to examine emotional prosody perception in children with low-function autism (LFA). They reported that children with LFA lacked preference for positive emotional tone compared with TD children. Le Sourn-Bissaoui et al. (2013) showed that when emotional prosody was the only cue to identify emotions and presented without any situational contexts, children and adolescents with HFA performed just as well as controls. However, when prosody was embedded in a discrepant context, the HFA groups relied less on the positive emotional prosody than TD groups to identify emotions.

Several alternative interpretations would account for the deficit of happy prosody perception in children with HFA. First, the happy prosody is perceptually less salient than sad and angry prosodies. For example, Dimitrovsky (1964) suggested that happy prosody is not perceptually distinct even for TD children. In that study, children sometimes confused happy prosody with angry prosody, but they seldom showed the reverse pattern. The reduced perceptual salience of happy prosody might have larger effects on children with HFA than on TD children when they judge the prosody of positive emotions.

Secondly, the social interaction perspective of emotional development offers another explanation. Children with HFA would confront more negative daily events than TD children when interacting with people. For example, Peterson, Wellman, and Liu (2005) suggested that negative life experiences, such as being a victim of bullying, were more common among ASD children than TD children, and such life experiences were correlated with the theory of mind deficit in ASD children. Negative life events, such as being bullied or blamed, are usually accompanied by prosody of negative emotions (including sadness and anger). Therefore, compared with TD children, children with HFA would be less responsive to prosody of positive emotions and would show selective deficits in processing of happy prosody.

The results of this study revealed that identifying emotions from sentences was easier than from words for both children with HFA and TD children, suggesting that prosodic cues serve as a more useful basis for emotional judgment in sentences than in words. In other words, children with HFA more accurately perceive emotions for sentences with embedded prosodic cues than for words of the same kind. Interestingly, it seems that extended utterance length was also able to assist the children to correctly recognize emotions though speech prosody. However, the effect of utterance length on emotional prosody might interact with the congruence between emotion and semantics. The results from this study revealed that children with HFA experienced greater difficulty in recognizing happy prosody than TD children for both emotionally neutral sentences and emotionally relevant words.

Further, the results from this study revealed associations among happy prosody perception, pragmatic language, and social adaptation. Moreover, poor prosody perception is associated with severity of autistic symptoms. These findings support the hypothesis that limitations in perceiving emotional prosody are related to social communication impairments in school-aged children with ASD. The results of this study support the hypothesis that ASD children's difficulties in perceiving the prosody of positive emotions underlie their poor performance in recognizing the emotions and communication intent of a speaker.

This study has identified several clinical implications for designing intervention programs to alleviate pragmatic impairments in children with ASD. Emotional prosody perception is often neglected when designing intervention programs for children with autism in clinical settings (Matsuda and Yamamoto, 2013). Strong intonation cues, such as a lower pitch, longer tempo, and louder amplitude, facilitates the interpretation of ironic utterances in TD children (Wang et al., 2006). Intervention programs might more effectively improve pragmatic skills (such as emotional understanding or irony comprehension) when utterances addressing children with ASD are delivered with prosody cues they find more perceptually salient when recognizing emotions.

This study assessed the role of emotional prosody perception on recognizing emotional states. However, in real-life social interactions, ASD children typically encounter multiple sensory inputs, such as visual and auditory inputs, in various contexts, and these cues are crucial for recognizing a speaker's emotional state (Adolphs, Sears, & Piven, 2001; Bal et al., 2010; Brosch, Grandjean, Sander, & Scherer, 2009). A meta-analysis article indicated that recognition of negative emotions from visual stimuli was significantly impaired in autistic children, whereas that of happiness was only marginally impaired (Uljarevic & Hamilton, 2013). ASD children might weight the perception of auditory and visual stimuli differently in recognizing positive versus negative emotions. Future studies might explore the relative contributions of prosodic and facial information to the impairments of emotional classification and social communication in school-aged children with HFA.

This current study extends our understanding of emotional prosody perception in children with HFA by using emotional prosody presented with words and sentences. In conclusion, school-aged children with HFA are less accurate than TD children in using prosody information to recognize the positive emotion of happiness; however, they perform as well as TD children in perceiving prosody of the negative emotions of sadness and anger. Additionally, children with HFA with greater accuracy in happy prosody perception also had higher functional abilities in pragmatic language and social adaptation, suggesting that increasing prosody perception abilities in children with HFA can reduce their social communication and adaptation difficulties.

Acknowledgements

The authors would like to express their warmest thanks to all the children who participated in this study, and to their parents for their kind co-operation. We also thank Hui-Ching Liu, M.D., Wei-Chih Hung, Wan-Hsuan Chen, Chih-Hung Chang, Shu-Ting Lu, and Chih-Hsang Chang for their assistance with data collection. This research was supported by the National Science Council at Taiwan (NSC100-2628-H-002-002-MY3, NSC102-2420-H-002-005-MY2).

References

Adolphs, R., Sears, L., & Piven, J. (2001). Abnormal processing of social information from faces in autism. *Journal of Cognitive Neuroscience, 13*(2), 232–240.
American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. Arlington, VA: American Psychiatric Association.

Bal, E., Harden, E., Lamb, D., Van Hecke, A. V., Denver, J. W., & Porges, S. W. (2010). Emotion recognition in children with autism spectrum disorders: Relations to eye gaze and autonomic state. *Journal of Autism and Developmental Disorders, 40*(3), 358–370.
Bertrand, R., & Priego-Valverde, B. (2011). Does prosody play a specific role in conversational humor? *Pragmatics and Cognition, 19*(2), 333–356.
Bishop, D. V. M. (1998). Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry, 39*(6), 879–891.
Bishop, D. V. M., & Baird, G. (2001). Parent and teacher report of pragmatic aspects of communication: Use of the Children's Communication Checklist in a clinical setting. *Developmental Medicine and Child Neurology, 43*(12), 809–818.
Brennand, R., Schepman, A., & Rodway, P. (2011). Vocal emotion perception in pseudo-sentences by secondary-school children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 5*(4), 1567–1573.
Brooks, P. J., & Ploog, B. O. (2013). Attention to emotional tone of voice in speech perception in children with autism. *Research in Autism Spectrum Disorders, 7*(7), 845–857.
Brosch, T., Grandjean, D., Sander, D., & Scherer, K. R. (2009). Cross-modal emotional attention: Emotional voices modulate early stages of visual processing. *Journal of Cognitive Neuroscience, 21*(9), 1670–1679.
Chevallier, C., Noveck, I., Happe, F., & Wilson, D. (2011). What's in a voice? Prosody as a test case for the theory of mind account of autism. *Neuropsychologia, 49*(3), 507–517.
Dimitrovsky, L. (1964). The ability to identify the emotional meaning of vocal expression at successive age levels. In J. Davitz (Ed.), *The communication of emotional meaning* (pp. 69–86). New York: McGraw-Hill.
E-Prime (Version 1.1) (2004). *Psychology Software Tools, Pittsburgh, PA*. http://www.pstnet.com (Computer program).
Friend, M., & Bryant, J. B. (2000). A developmental lexical bias in the interpretation of discrepant messages. *Merrill-Palmer Quarterly (1982), 46*(2), 342–369.
Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. D. (2007). The 'Reading the Mind in the Voice' test-revised: A study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders, 37*(6), 1096–1106.
Grossman, R. B., & Tager-Flusberg, H. (2012). Quality matters! Differences between expressive and receptive non-verbal communication skills in adolescents with ASD *Research in Autism Spectrum Disorders, 6*(3), 1150–1155.
Grossman, R. B., Bemis, R. H., Skwerer, D. P., & Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research, 53*(3), 778–793.
Hubbard, K., & Trauner, D. A. (2007). Intonation and emotion in autistic spectrum disorders. *Journal of Psycholinguistic Research, 36*(2), 159–173.
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.
Le Sourn-Bissaoui, S., Aguert, M., Girard, P., Chevreuil, C., & Laval, V. (2013). Emotional speech comprehension in children and adolescents with autism spectrum disorders. *Journal of Communication Disorders, 46*(4), 309–320.
Lin, B. Q., Huang, Y. Q., Huang, G. J., & Xuan, C. H. (2009). *Chinese language development test for school-aged children*. Taipei, TX: Special Education Center in National Taiwan Normal University.
Lindner, J. L., & Rosén, L. A. (2006). Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger's syndrome. *Journal of Autism and Developmental Disorders, 36*(6), 769–777.
Matsuda, S., & Yamamoto, J. (2013). Intervention for increasing the comprehension of affective prosody in children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 7*(8), 938–946.
McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: A critical review. *International Journal of Language and Communication Disorders, 38*(4), 325–350.
McCann, J., Peppé, S., Gibbon, F. E., O'Hare, A., & Rutherford, M. (2007). Prosody and its relationship to language in school-aged children with high-functioning autism. *International Journal of Language and Communication Disorders, 42*(6), 682–702.
Morton, J. B., & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Development, 72*(3), 834–843.
O'Connor, K. (2012). Auditory processing in autism spectrum disorder: A review. *Neuroscience & Biobehavioral Reviews, 36*(2), 836–854.
Paul, R., Augustyn, A., Klin, A., & Volkmar, F. R. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 35*(2), 205–220.
Paul, R., Shriberg, L. D., McSweeny, J., Cicchetti, D., Klin, A., & Volkmar, F. (2005). Brief report: Relations between prosodic performance and communication and socialization ratings in high functioning speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 35*(6), 861–869.
Paulmann, S., & Pell, M. (2010). Contextual influences of emotional speech prosody on face processing: How much is enough? *Cognitive, Affective, & Behavioral Neuroscience, 10*(2), 230–242.
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.
Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development, 76*(2), 502–517.
Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders, 32*(3), 189–194.
Sakkalou, E., & Gattis, M. (2012). Infants infer intentions from prosody. *Cognitive Development, 27*(1), 1–16.
Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002). The CAST (Childhood Asperger Syndrome Test): Preliminary development of a UK screen for mainstream primary-school age children. *Autism, 6*(1), 9–31.
Shriberg, L. D., Paul, R., McSweeny, J. L., Klin, A., Cohen, D. J., & Volkmar, F. 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.
Singh, L., & Harrow, M. S. (2014). Influences of semantic and prosodic cues on word repetition and categorization in autism. *Journal of Speech, Language, and Hearing Research, 57*(5), 1764–1788.
Sony Pictures Digital Inc. (2004). *Sound forge (version 7.0) (computer software)*. Madison, WI: Sony Pictures Digital Inc.
Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland adaptive behavior scales: Interview edition, survey form manual*. Circle Pines, MN: American Guidance Service.
Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders, 43*(7), 1517–1526.
Vonden, J., & Lord, C. (1991). Neologisms and idiosyncratic language in autistic speakers. *Journal of Autism and Developmental Disorders, 21*(2), 109–130.
Volden, J., Coolican, J., Garon, N., White, J., & Bryson, S. (2009). Brief report: Pragmatic language in autism spectrum disorder: Relationships to measures of ability and disability. *Journal of Autism and Developmental Disorders, 39*(2), 388–393.
Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2006). Neural basis of irony comprehension in children with autism: The role of prosody and context. *Brain, 129*(4), 932–943.
Wechsler, D. (2003). *Wechsler intelligence scale for children—Fourth edition (WISC-IV)*. San Antonio, TX: The Psychological Corporation.
Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics, 38*(10), 1559–1579.
Wittfoth, M., Schröder, C., Schardt, D. M., Dengler, R., Heinze, H. J., & Kotz, S. A. (2010). On emotional conflict: Interference resolution of happy and angry prosody reveals valence-specific effects. *Cerebral Cortex, 20*(2), 383–392.
Wu, W. D., Zhang, Z. F., Lu, T. H., & Qiu, S. C. (2004). *Vineland adaptive behavior scale-Chinese version (Sparrow, S.S, Balla, D.A., & Cicchetti, D. V., 1984)*. Taipei, Taiwan, R.O.C Psychological Publishing.
Young, E., Diehl, J., Morris, D., Hyman, S., & Bennetto, L. (2005). The use of two language tests to identify pragmatic language problems in children with autism spectrum disorders. *Language Speech and Hearing Services in Schools, 36*(1), 62–72.