

## **Facilitation of Vocabulary Acquisition by Emotional Prosody in Children with Autism**

Xing-Da Ju <sup>a b †</sup>, Han-Qian Wu <sup>c †</sup>, Bo Hu <sup>a d †</sup>, Chang Lu<sup>a b \*</sup>

† Contributed equally

<sup>a</sup> School of Psychology, Northeast Normal University, China

<sup>b</sup> Jilin Provincial Key Laboratory of Cognitive Neuroscience and Brain Development, China

<sup>c</sup> Mental Health Education Center, JiangXi Vocational College of Industry & Engineering,  
China

<sup>d</sup> School of Social and Behavioral Science, Nanjing University, China

This study has been accepted by the International Journal of Developmental Disabilities and can be accessed at the following link: <https://doi.org/10.1080/20473869.2024.2410532>

Correspondence concerning this article should be addressed to Chang Lu, School of Psychology, Northeast Normal University, Changchun, 130024, China [luc816@nenu.edu.cn](mailto:luc816@nenu.edu.cn).

### **Funding**

This work was supported by the Ministry of Education, Humanities and Social Sciences project [grant numbers 21YJCZH056].

### **Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### **Ethics approval statement**

This study was approved by the Northeast Normal University Institutional Review Board (2022025). All participants provided informed consent.

### **Abstract**

Helping children with autism gain a better grasp of vocabulary is an important way to improve their language skills. Information that incorporates emotional prosodies provide richer semantics and context and may be more conducive to vocabulary acquisition in children with autism. This study aimed to explore the facilitating effect of positive emotional prosody on vocabulary acquisition with children with autism. A total of 12 children with autism received interventions with neutral and positive emotional prosodies. The results revealed that vocabulary learned in positive emotional prosodies was more effective than neutral prosodies. The differences in the learning effects of emotional prosodies on part of speech were not significant. However, in children with higher autism severity, neutral prosodies led to diminished expressive vocabulary performance. This study improved vocabulary comprehension and language skills of children with autism through emotional prosody intervention and tentatively revealed the role of lexicality in this.

*Keywords:* vocabulary acquisition; emotional prosody; autism; part of speech

## 1 Introduction

Language disorders are prevalent in autism. Language comprehension is a central part of language disorders in children with autism, who are unable to express their thoughts and feelings appropriately and have difficulty understanding and responding to the words of others. (Rogers et al., 2006). Improving the language disorder in autism has many benefits for them, , such as enhancing life skills, facilitating social integration, and maintaining psychological health (Hardan et al., 2015; Mayer-Benarous et al., 2021; Nicoletti & Cantacini, 2021). Vocabulary acquisition is an important component of language development and the basis for the acquisition and development of language skills, which influence children's outcomes in several domains. (Haebig et al., 2017) Vocabulary is the foundation of language learning, and the process of vocabulary learning is also the process of mastering and practicing phonetics and grammar.(Hartley et al., 2019) Therefore, helping children with autism get a better grasp of vocabulary is an important way to improve their language skills, language teaching in schools should strengthen vocabulary teaching and make vocabulary teaching better combined with phonetics and grammar teaching. (Paris, 2005; Thurm et al., 2007).

The autistic population has deficits in multi-information and multi-channel integration and is less able to perceive information compared to the healthy population. However, people with autism are less sensitive to pitch and tone intensity information during the acoustic cue perception stage, and this ability contributes to some extent to the information processing bias in this group (O'Connor, 2012), which makes the group more susceptible to factors such as semantics and context.(Naigles et al., 2023). Voice is an important medium of human communication, which conveys not only semantic information, but also the emotional state of the speaker (MacFarlane et al., 2022; West et al., 2018). In the process of processing spoken information, the non-linguistic

aspects of speech signals (referring to the emotional intonation of speech, including components such as intonation, pitch, voice intensity, and rate of speech) convey emotional information beyond semantic linguistic content, and it has been found that contextually varying variations in emotional rhythms affect lexical processing in conjunction with lexical meaning, i.e., in the vocabulary learning stage, emotional prosody changes affect lexical processing. (Nygaard & Lunders, 2002). Emotional induction during the learning process may affect the allocation of attention, encoding, and consolidation of lexical semantics of new vocabulary (Nielson & Arentsen, 2012). Whether teachers are aware of it or not, the emotional messages they express are reflected in the children and further affect their learning outcomes (Harrington & Allen, 2014). Children with autism showed a psychological trait that prioritizes phonological perception when understanding vocabulary (Grossman, Bemis, Plesa Skwerer, et al., 2010). Children with autism can understand the semantic meaning conveyed and perceive the emotional condition of others through intonation and stress changes of voice (Scheerer et al., 2020; L. Wang et al., 2022).

Emotional prosody was introduced by Monrad-Kohn (1947), defined as a speaker conveying different types of emotions through changes in accent or intonation without regard to vocabulary or grammatical structure. In other words, emotional prosody is a paralinguistic message consisting of changes in acoustic cues such as pitch, intensity, and duration (Larrouy-Maestri et al., 2024). Emotional information tends to capture attention more effectively than neutral content. This heightened attention facilitates better encoding of information into memory (Hinojosa et al., 2010). Emotional words or phrases are often remembered more vividly and for longer periods, enhancing the overall processing and retention of emotional information. The emotional prosody of the sentence context also affected lexical processing. The combination of the meaning conveyed by the prosody and the semantics conveyed by the vocabulary will lead to a more lasting information

in memory (Greco et al., 2021). Emotional prosody cues in spoken language not only impact learners' vocabulary memory but also help them learn new vocabulary. Therefore, information that incorporates emotional prosody provides a richer semantic and context and may be more conducive to vocabulary acquisition in children with autism.

Vocabulary has always been of great interest in the language learning process, with nouns and verbs (i.e. the part of speech) being considered the beginning of a language learning process and the foundation upon which the human language system is built. According to statistics, there are about three times more nouns than verbs in English (Miller & Fellbaum, 1991), and nouns are also central in Chinese (Shen, 2012). Although vocabulary learning differences in children with autism have been widely documented, some of the vocabulary learning strategies observed in normal children are also reported in children with autism (Williams et al., 2013). Children with autism showed a noun tendency in the Chinese language context (Yong-Xiang et al., 2016). In picture naming and other lexical tasks, children with autism process nouns better than verbs (Mätziga et al., 2009). Understanding the role of emotional prosody in the acquisition of vocabulary with different parts of speech is relevant for the development of educational strategies.

In the field of language in children with autism, most of the literature has investigated the neurocognitive mechanisms of word learning, and few studies have considered the moderating effect of emotion on word learning. There is growing evidence of the important role of emotion in word processing, and auditory stimuli elicit stronger behavioral effects compared to visual, so it is necessary to investigate the effect between emotional information from auditory stimuli and word learning. First, there is no consensus on whether there is an emotional perceptual advantage for children with autism (Haebig et al., 2021). Validating the effectiveness of emotional prosody in vocabulary acquisition will demonstrate both the emotional perceptual characteristics of children

with autism and the importance of emotional prosody in language learning. Second, much of the existing research has focused on how emotional information affects vocabulary learning in autism, and not much research has been done on the ability of children with autism to use vocabulary. Understanding the effectiveness of emotional prosody across different part of speech can both cue the mechanisms of language acquisition for children with autism and provide guidance for the development of intervention strategies. Finally, most of the current studies have been conducted with native speakers of English, which is a non-tonal language (does not use syllables and pitch changes to differentiate words) that does not change due to changes in pronunciation, and tones indicate the tone of a sentence (e.g., to express a statement or a question). However, Chinese, as a tonal language (a language in which the meaning of words is affected by lexical tones), tones also indicate lexical meanings, and different tones express different lexical meanings (M. Zhang et al., 2022). For example, in Mandarin, the syllable /yu/ in the second tone indicates "fish", while in the third tone it indicates "rain". Few studies have been conducted on the prosodic learning of autistic individuals who use tonal languages. It needs to be further verified whether the effect of emotional prosody intervention in the Chinese language learning context has been the same as in the English context.

This study aims to find out whether interventions with different emotional prosody affect vocabulary learning in children with autism. Based on this, it will explore which part of speech (nouns or verbs) have a greater impact on vocabulary learning in children with autism. The hypotheses of this study are as follows:

**H1:** The positive emotional prosodies were more effective than neutral prosodies on vocabulary acquisition.

**H2:** The intervention effect of positive emotional prosody is more persistent than neutral

emotional prosody.

**H3:** The effects of emotional prosody interventions varied across part of speech.

## 2 Method

### 2.1 Screening of subjects

In this study, 15 children with autism were included from a special education institution in China. Subjects were screened by combining the results of the Peabody Picture Vocabulary Test (Dunn & Dunn, 2012), on-site observations, and information from teacher interviews. Subjects meeting the following conditions were included.

A post power analysis was conducted using G\*Power version 3.1.9.7 (Faul et al.2007). The effect size was 1.27 (effect size U is converted to Cohen d). With a significance criterion of  $\alpha=.05$  and sample size is 12. The statistical power was 0.82.

- (1) Children aged 2-12 who have been diagnosed with autism according to DSM-5 or ICD-11(American Psychiatric Association & Association, 2013; World Health Association, 2020);
- (2) Children without hearing or visual impairments, major physical disabilities, or genetic disorders.

The subjects were included after obtaining informed consent from parents. A total of 15 subjects were initially obtained, three were dropped out in the middle of the intervention due to the impact of the Covid-19. Data from the remaining 12 subjects who completed all interventions were analyzed. Basic characteristics of included subjects were shown in Table S1. The studies were conducted according to the guidelines in the Declaration of Helsinki and the standards of ethics in research with human beings and were approved by the Research Ethics Committee of the Principal Investigator's University.



## 2.2 Participants

Participants totaled 12 (10 males and 2 females) with an age range of 54-106 months with a mean age of 78.50 months (SD=5.65), a mean intelligence level of 46.41 (SD=5.74), a mean verbal PPVT=20.75 (SD=3.56), and a mean level of autism of 36.25 (SD=1.31).

Table 1 Participant demographics

Demographics	M	SD
Chronological age (month)	78.50	5.65
Intellectually language level	46.42	5.74
	20.75	3.56
Degree of autism	36.25	1.31

## 2.3 Intervention materials

In this paper, the intelligence score from the Peabody Picture Vocabulary Test was used to measure the IQ of children with autism and the verbal ability score to measure the level of language in autism. The test consists of two parallel books, L and M. The split-half confidence coefficients are distributed between 0.67-0.88 for the younger age samples and 0.80-0.85 for the older age samples on the L-shape; on the M-shape the split-half confidence coefficients are distributed between 0.61-0.88 for the younger age samples, 0.52-0.92 for the stability coefficients of the older age samples, and 0.73-0.91 for the coefficients of equivalence. 0.91.(Dunn & Dunn, 2012)

A total of 44 neutral vocabularies whose meanings do not involve positive or negative emotion were selected from the materials of the Peabody Picture Vocabulary Test. By performing a word picture matching task on children with autism (Harvey & Schnur, 2016), 18 vocabularies with a score of zero for all children were retained. Each vocabulary was matched with a target picture. All of the pictures were randomly selected from the word-matched picture library (Dunn & Dunn, 2012). Considering the subjects' daily teaching style and their susceptibility to the

attraction properties of electronic devices, each picture was centered and printed in black-and-white on an A4 sheet of paper ( $20.99 \times 29.7$  cm).

Voice materials were recorded by a broadcasting professional using positive and neutral emotional prosodies and saved in WAV format for a total of 36 voices (one positive voice and one neutral voice for each vocabulary). A total of 26 psychology graduate students were asked to assess the type of prosody (positive, neutral) of each voice. All voices have a recognition rate of 80% or higher.

To further ensure that the accuracy of the emotional prosodies categories of the voice materials, the pitch of voice materials was analyzed by Praat software (Boersma & Weenink, 2011). Pitch is the emotional prosody information that effectively distinguishes between different emotional states such as anger, happiness, and sadness (Pell & Baum, 1997). In general, happy speech is fast, with a high pitch average and a small fundamental frequency range; sad speech is the slowest, with a lower pitch average and a slightly larger fundamental frequency range (Banse & Scherer, 1996). An independent samples t-test was performed to test the differences in pitch across emotional prosodies. The positive emotional prosodies ( $M=326.45$ ,  $SD=49.67$ ) were significantly higher than the neutral emotional prosodies ( $M=256.88$ ,  $SD=42.55$ ),  $t(36)=4.64$ ,  $p<.001$ .

## **2.4 Measurement**

Referring to the methods of Harvey & Schnur (2016), this study examines vocabulary acquisition effects based on a word picture matching task and picture naming task. The word picture matching task was usually practiced before the picture naming task to facilitate subjects' familiarity with the cues. Both tasks are semantically mediated, and vocabulary semantic mastery can be assessed based on the scores of the two tasks. For each task, the total score is the number

of correct responses times divided by the total number of tasks. All analyses were performed with JASP version 0.18.2 (Love et al., 2019).

#### ***2.4.1 Word Picture Matching Task***

The word picture matching task examines the subject's mastery of receptive vocabulary. It involves the issues of language research such as picture comprehension, semantic judgment, and semantic interference effects. By taking advantage of the visual processing of children with autism, vocabulary acquisition can be studied using word picture matching task. The main procedure of the word picture matching task is to present the subjects with four pictures, including one target picture and three interference pictures, and play the instructional language with the related target vocabulary voice at the same time. After hearing the vocabulary voice, subjects are asked to point out the correct target picture, and the interference pictures are randomly selected from the vocabulary-matched picture library. Pointing out correctly was scored as one point and incorrectly was scored as zero point. The task was used to assess the acquisition of semantic knowledge, and over time, if the score remained stable, it indicated that the knowledge was retained.

#### ***2.4.2 Picture Naming Task***

The picture naming task examined the subjects' mastery of expressive vocabulary (vocabularies that can be used in expressive vocabulary behaviors such as speaking and writing). Picture naming task included processing stages such as conceptual preparation, lexical selection, phonological encoding, and pronunciation. The main procedure was to present subjects with a picture of the target while playing the instruction and related vocabulary speech, and then ask them to name the picture after hearing the instruction. Each question was scored one point for completely correct naming and zero points for incorrect naming or no response.

### **2.5 Intervention process**

Figure 1 displays the implementation process of our intervention. After subject selection, each subject was given one-on-one intervention in a classroom they were familiar with. The intervention was performed by authors and assisted by a teacher familiar to the subject. Each subject received nine interventions, two times per week, and each time lasting about 20 minutes. The subjects learned a total of 18 words, two words at a time. We randomized the intervention material while ensuring that the number of emotional prosody and part of speech of each subject's intervention material was consistent. After each intervention, subjects were post-tested on the vocabulary matching task and picture naming task on the selected vocabulary material. Intervention effect is a summary of all post-test results. To verify the durability of the intervention, a follow-up tests contains word picture matching task and the picture naming task were consisted three months after the end of all interventions.

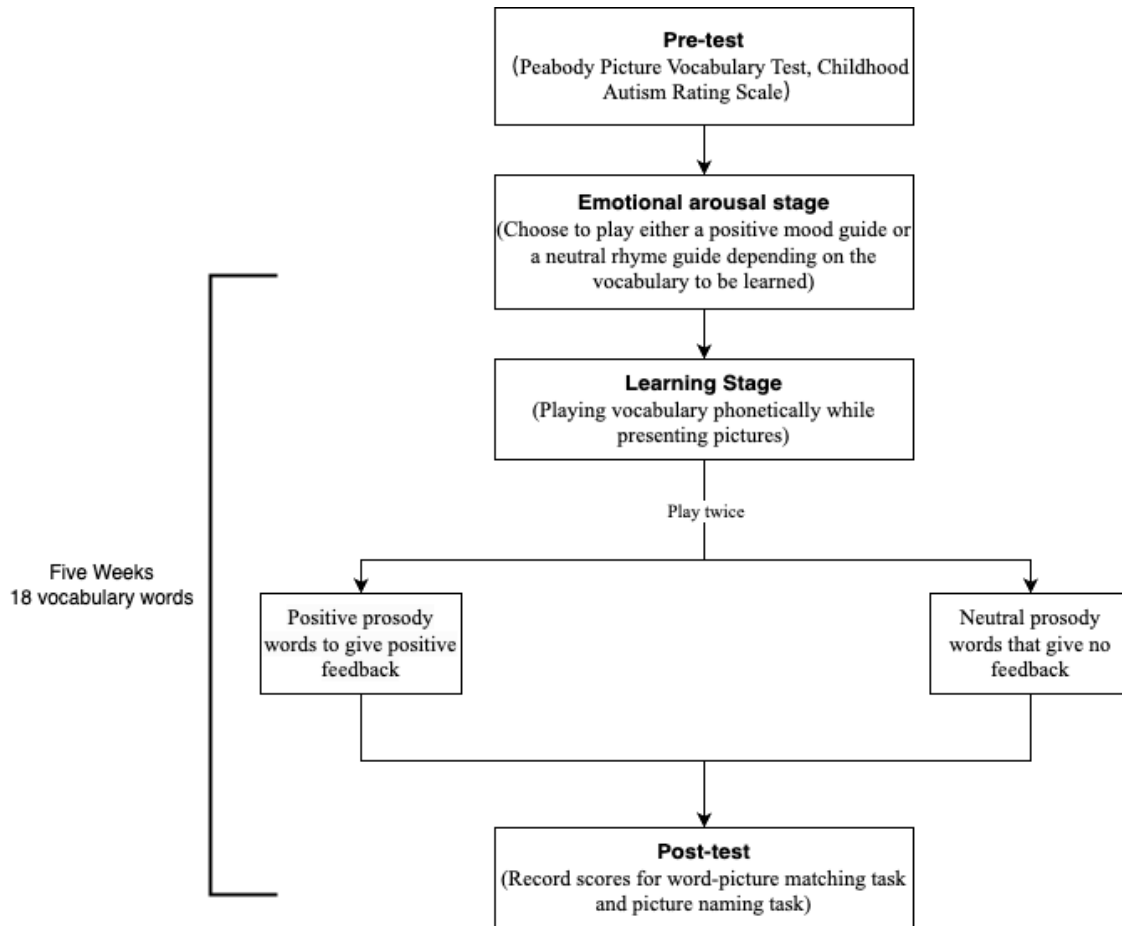
Each intervention consisted of three stages: emotional arousal, learning, and testing:

(1) Emotional arousal stage: a recorded instructional phrase with different emotions was played first to induce the emotional state before teaching new vocabulary.

(2) Learning stage: play the recorded vocabulary materials with different emotions. Since autism has difficulty concentrating, it is necessary to focus the subject's attention on the intervention materials when the subject's attention is shifted. The teacher only needs to click the play and pause buttons to adjust the attention of autism in time.

(3) Post-test: first play the emotional guidance voice, and then test the scores of the picture naming task and word picture matching task. Subjects answered and recorded their correctness and made relevant records.

**Figure 1.** *Intervention Flowchart*



### 2.6 Statistical analysis

The statistical program JASP version 0.18.2 was used (Love et al., 2019). Initially, the word-picture matching task in the vocabulary comprehension test, which represents receptive vocabulary knowledge, and the picture naming task, which represents expressive vocabulary knowledge. Each subject received positive emotional rhyming vocabulary and neutral emotional rhyming vocabulary, so emotion was grouped as the independent variable and scores on the two tasks were grouped as the dependent variable. The overall sample distribution was non-normal, as well as the sample size was small, therefore, non-parametric tests were used (Mann-Whitney U test).

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$n_1$  and  $n_2$  are the sample sizes of the two samples, respectively,  $R_1$  is the rank sum of the first

sample. A smaller U-value indicates a significant difference between the two sample distributions. (Mann & Whitney, 1947)

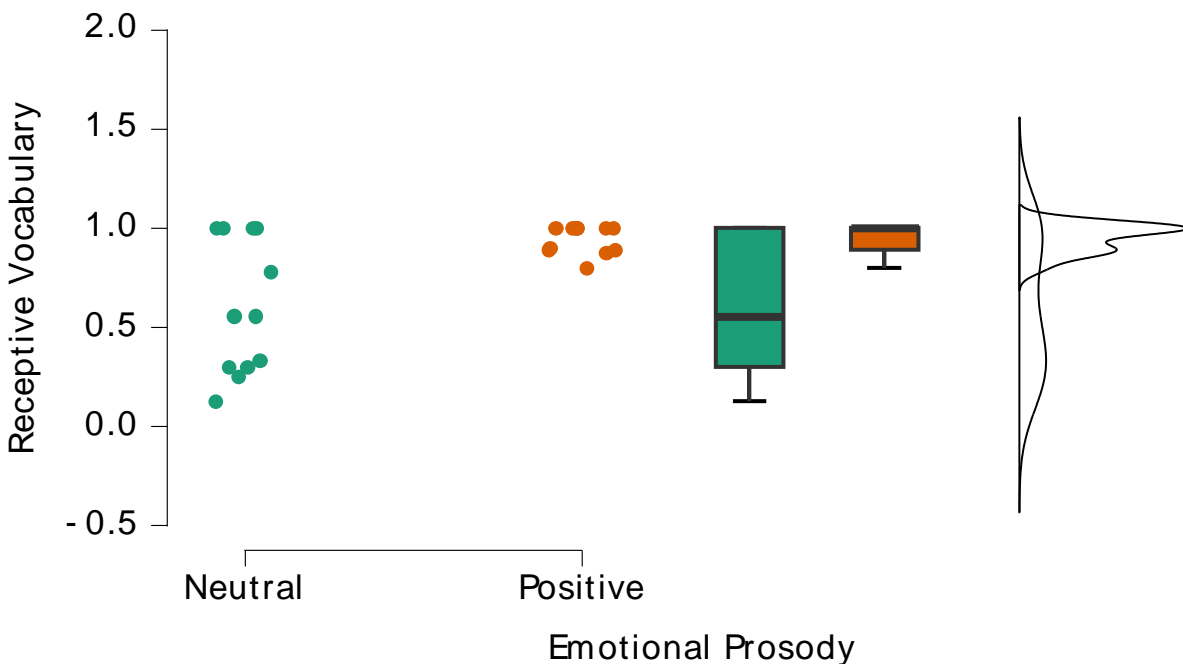
## 4 Result

### 4.1 Emotional prosody improving vocabulary acquisition

#### 4.1.1 Differences in learning effect for receptive vocabulary

The learning effect of receptive vocabulary was examined through the word picture matching task. The differences in word picture matching task scores between the positive and neutral emotional prosodies were analyzed. The analysis results are shown in Figure 2. The results showed a significant difference in vocabulary acquisition in positive emotional prosodies ( $M=0.95$ ,  $SD=0.07$ ) and neutral prosodies ( $M=0.60$ ,  $SD=0.34$ ),  $U=34.00$ ,  $p=.023$ , suggesting that children with autism are learning receptive knowledge of vocabulary better with positive emotional prosodies than with neutral emotional prosodies.

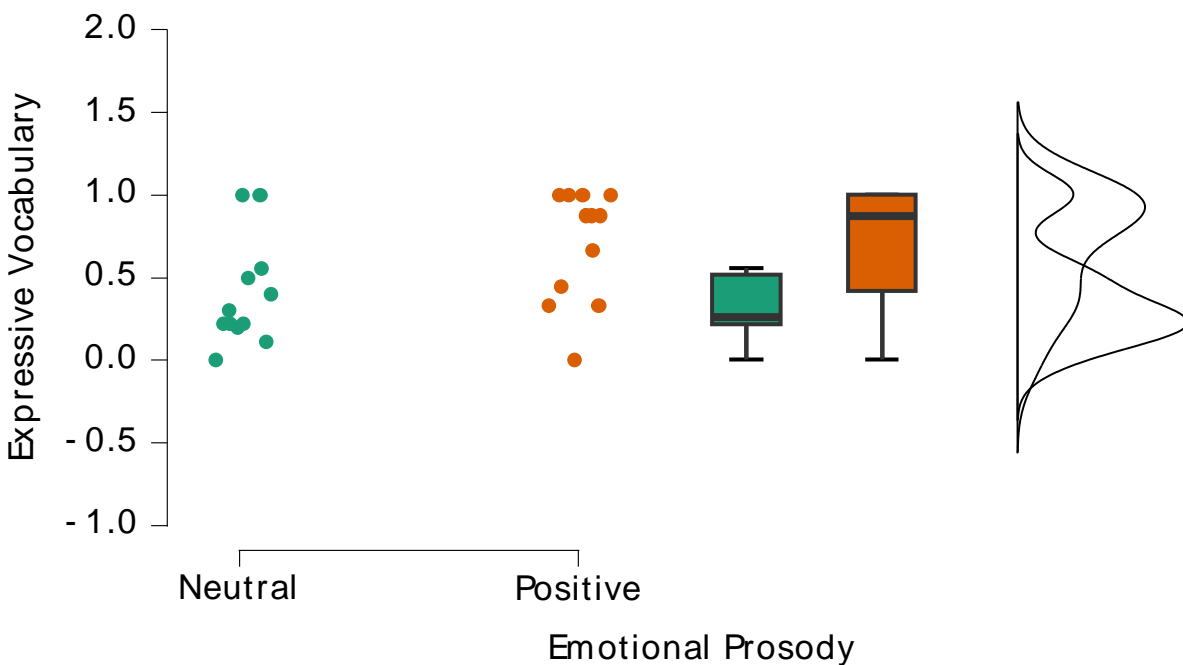
**Figure 2.** Differences in the Effectiveness of Emotional Rhyme Interventions on Receptive Vocabulary Knowledge



#### 4.1.2 Differences in learning effect for expressive vocabulary

The learning effects of expressive vocabulary were examined in a picture naming task. The differences in picture naming task scores between the positive and neutral emotional prosodies were analyzed. Observation of the post-test data showed that two subjects were reluctant to make sounds when doing the picture naming task. We excluded the two extreme values and then conducted the test of difference. The results are shown in Figure 3, there was a significant difference in the effect between the expressive vocabulary with positive emotional prosodies ( $M=0.39$ ,  $SD=0.32$ ) and neutral prosodies ( $M=0.70$ ,  $SD=0.34$ ),  $U=37.50$ ,  $p=.048$ . This indicated that positive emotional prosodies are beneficial for expressive vocabulary for children with autism.

**Figure 3.** *Differences in the Effectiveness of Emotional Rhyme Interventions on Expressive Vocabulary Knowledge*

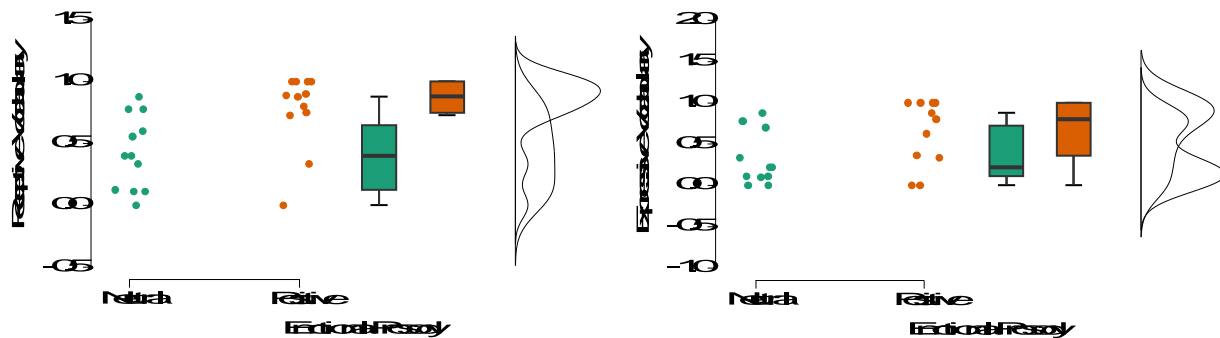


#### 4.2 Long-term effect of emotional prosody on vocabulary acquisition

The present study further validated the durability of the effect of emotional prosody on vocabulary acquisition after three months, results are shown in Figure 4. In receptive vocabulary,

the positive emotional prosodies ( $M=0.77$ ,  $SD=0.31$ ) are significantly higher than neutral prosodies ( $M=0.42$ ,  $SD=0.30$ ),  $U=26.50$ ,  $p=.009$ . In expressive vocabulary, the positive emotional prosodies ( $M=0.65$ ,  $SD=0.38$ ) are significantly higher than neutral prosodies ( $M=0.35$ ,  $SD=0.33$ ),  $U=37.00$ ,  $p=.045$ . That means that positive emotional prosodies are more persistence than neutral prosodies in vocabulary acquisition.

**Figure 4.** Long-term effects of emotional Prosody on receptive and expressive vocabulary knowledge



### 4.3 Differences in emotional prosodies on part of speech

Finally, this study compared the differences between receptive vocabulary and expressive vocabulary in different part of speech (nouns and verbs). The results are shown in Table 1. There is no significant difference in scores on the expressive vocabulary and receptive vocabulary between nouns and verbs indicated for children with autism. The difference in the follow-up test results for the receptive vocabulary as well as the expressive vocabulary is shown in Table 2, which shows that there is still no significant difference in part of speech.

**Table 2.** Differences in part of speech of post test

	Emotional Prosody	Noun (M±SD)	Verb (M±SD)	U	p
Receptive Vocabulary	Positive	0.79±0.06	0.76±0.08	82.00	.552



	Neutral	0.55±0.32	0.43±0.33	86.50	.415
Expressive Vocabulary	Positive	0.59±0.32	0.54±0.30	86.50	.400
	Neutral	0.30±0.29	0.34±0.28	62.50	.597

**Table 3.** *Differences in part of speech of follow-up test*

	Emotional Prosody	Noun (M±SD)	Verb (M±SD)	U	p
Receptive Vocabulary	Positive	0.72±0.13	0.67±0.15	86.50	.403
	Neutral	0.44±0.27	0.37±0.31	80.00	.661
Expressive Vocabulary	Positive	0.56±0.30	0.50±0.31	85.00	.463
	Neutral	0.33±0.33	0.35±0.28	67.50	.815

#### 4.4 Exploratory analysis

Given the differences among the subjects, we further analyzed the role of heterogeneity factors in emotional prosody intervention, including gender (Table S2), intellectual age (Table S3), and autism level (Table S4). The results showed no significant differences in the effects of emotional prosody across different genders and ages. However, there was heterogeneity in autism severity. Children with higher levels of autism (higher CARS scores) exhibited lower expressive vocabulary performance than moderate levels on both the posttest ( $U=2.00$ ,  $p=.011$ ) and follow-up test ( $U=2.50$ ,  $p=.016$ ) of the neutral emotional prosody intervention. In contrast, positive emotional prosody did not show any differences.

### 5 Discussion

This study used word picture matching tasks and picture naming tasks to examine the effect of emotional prosody on vocabulary acquisition in children with autism. The results showed that positive emotional prosodies were more effective than neutral prosodies on both receptive and

expressive vocabulary for vocabulary acquisition. Emotional prosodies had no significant effect on different part of speech (nouns and verbs).

### **5.1 Effects of emotional prosodies on vocabulary acquisition**

The present study found that emotional prosody significantly improved the learning effect of receptive and expressive vocabulary in Chinese context. This is consistent with the findings of native English speakers. Children with autism often show differences in the processing of emotional cues (McCann et al., 2007). The intervention process effectively utilizes the role of emotional induction in the learning process of children with autism, which can affect the allocation of attention, encoding, and consolidation of lexical semantics of new words. (Nielson & Arentsen, 2012), empowering them to better navigate and internalize vocabulary acquisition. Emotional prosody, by conveying emotions through speech, can capture and hold the attention of individuals with autism more effectively than neutral speech (Le Sourn-Bissaoui et al., 2013). This enhanced attention to emotionally charged stimuli aids in the more efficient encoding of new information (Savekar et al., 2021; West et al., 2017). The study suggests that emotional prosodies, especially those conveying positive emotions, improve the engagement and motivation of learners, leading to better learning outcomes.

Three months after the intervention, the memory maintenance effect of vocabulary presented in positive emotional prosodies was superior to that of vocabulary presented in neutral prosodies in both receptive and expressive vocabularies. This suggests that during the learning process, the vocabulary presented in positive prosodies is learned more persistently than neutral vocabulary, and the memory maintenance effect is also better than that of the vocabulary presented in neutral prosodies. Emotional information can influence cognitive processes (Matsumoto et al., 2016). This information activates key emotional and memory-processing regions of the brain, such as the

hippocampus, amygdala, and prefrontal cortex (LaBar & Cabeza, 2006). These activities enhance the encoding, storage, and retrieval of memories, allowing them to leave a deeper and more lasting impression on the brain (Qasim et al., 2023).

Notably, children with higher levels of autism exhibited lower expressive vocabulary performance following neutral emotional prosody intervention. This emphasizes the importance of emotional prosody in vocabulary acquisition in autism. For children with higher autism severity, their impaired processing of neutral emotional prosody may make it difficult for them to extract linguistic information from these tones (Grossman, Bemis, Plesa, et al., 2010; J.-E. Wang & Tsao, 2015). On the other hand, positive emotional prosody, due to its heightened arousal properties, may create a better learning environment for children with autism in different levels of severity (Icht et al., 2021; Loveland et al., 1997).

## **5.2 Differences in learning performance across part of speech**

We observed no significant differences in the impact of emotional prosody on the learning of nouns and verbs, nor were there notable differences in follow-up assessments. Contrary to findings by Haebig et al. (2021), which suggested a noun preference among children with autism, our research did not reflect this tendency. This discrepancy hints at potential cultural variations in lexical preferences among children with autism (Norbury & Sparks, 2013). There are significant phonological, syntactic, semantic and pragmatic differences between Chinese and English (Fang et al., 2022; D. Zhang, 2012). The process of recognizing Chinese characters necessitates intricate cognitive and linguistic efforts pointing to the unique challenges and learning mechanisms in Chinese (Yin et al., 2023). The distinctive features of the Chinese language, its culture, and its writing system are likely to shape language and literacy acquisition in Chinese children with autism in ways that differ from their peers exposed to alphabetic scripts, as proposed by Naigles

et al. (2023). This underscores the importance of considering linguistic and cultural contexts in understanding the language learning experiences of children with autism.

### **5.3 Implication**

Based on the outcomes of this research, incorporating positive emotional intonations is recommended when introducing new vocabulary, as the emotional tone of spoken language significantly influences the effectiveness of language acquisition. Given the unique traits of children with autism, educators often find themselves needing to repeat instructions continuously, which can pose challenges in sustaining a consistently positive teaching demeanor. Creating audio materials can prove beneficial, offering children with autism consistent exposure to positive emotional cues throughout the learning experience and fostering a conducive environment for positive learning outcomes.

### **5.4 Limitation**

The present study found that positive emotional prosodies would have an effect on vocabulary acquisition in children with autism, but there are still some limitations that can be further expanded in subsequent studies. First, the sample size was small. Nonetheless, further investigation with a greater sample size is desirable. And we would also like to take into account more factors affecting the development of language learning in subsequent studies in order to isolate more clearly the effects of the intervention. Secondly, this study only assessed the effects of emotional prosody on vocabulary acquisition in children with autism. Future research could explore the contribution of emotional prosody to other language disorder issues in autism. Thirdly, whether the heterogeneity of children found in the present study to exist in children with autism persists in a larger group needs to be further investigated. Finally, although the present study provided the effects of an emotional prosody intervention in the context of Chinese culture, it did not make

direct cross-cultural comparisons. Future studies could assess the differences of the same intervention across cultures.

### **Data Availability Statement**

Details of data and analyses script are available via Open Science Framework ([https://osf.io/m9tkn/?view\\_only=601461414cee486ca87a3485091c4a0d](https://osf.io/m9tkn/?view_only=601461414cee486ca87a3485091c4a0d)).

### Reference

- American Psychiatric Association, D., & Association, A. P. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (Vol. 5, Issue 5). American psychiatric association Washington, DC.
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614–636. <https://doi.org/10.1037//0022-3514.70.3.614>
- Boersma, P., & Weenink, D. (2011). Praat: Doing Phonetics by Computer. *Ear & Hearing*, *32*(2), 266. <https://doi.org/10.1097/AUD.0b013e31821473f7>
- Dunn, L. M., & Dunn, D. M. (2012). *Peabody Picture Vocabulary Test—Fourth Edition* [Dataset]. <https://doi.org/10.1037/t15144-000>
- Fang, P., Liu, Q., Fang, P., & Liu, Q. (2022). A semantic contrastive study of Chinese and English verb “跑/run” from the perspective of cognitive semantics. *Forum for Linguistic Studies*, *4*(1), Article 1. <https://doi.org/10.18063/fls.v4i1.1458>
- Greco, C., Romani, M., Berardi, A., De Vita, G., Galeoto, G., Giovannone, F., Vigliante, M., & Sogos, C. (2021). Morphing Task: The Emotion Recognition Process in Children with Attention Deficit Hyperactivity Disorder and Autism Spectrum Disorder. *International Journal of Environmental Research and Public Health*, *18*(24), 13273. <https://doi.org/10.3390/ijerph182413273>
- Greenslade, K. J., Utter, E. A., & Landa, R. J. (2019). Predictors of Pragmatic Communication in School-Age Siblings of Children with ASD and Low-Risk Controls. *Journal of Autism and Developmental Disorders*, *49*(4), 1352–1365. <https://doi.org/10.1007/s10803-018-3837-x>
- Grossman, R. B., Bemis, R. H., Plesa, S. D., & Tager, -Flusberg Helen. (2010). Lexical and

- Affective Prosody in Children With High-Functioning Autism. *Journal of Speech, Language, and Hearing Research*, 53(3), 778–793. [https://doi.org/10.1044/1092-4388\(2009/08-0127\)](https://doi.org/10.1044/1092-4388(2009/08-0127))
- Grossman, R. B., Bemis, R. H., Plesa Skwerer, D., & Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research: JSLHR*, 53(3), 778–793. [https://doi.org/10.1044/1092-4388\(2009/08-0127\)](https://doi.org/10.1044/1092-4388(2009/08-0127))
- Haebig, E., Jiménez, E., Cox, C. R., & Hills, T. T. (2021). Characterizing the early vocabulary profiles of preverbal and minimally verbal children with autism spectrum disorder. *Autism*, 25(4), 958–970. <https://doi.org/10.1177/1362361320973799>
- Haebig, E., Saffran, J. R., & Ellis Weismer, S. (2017). Statistical word learning in children with autism spectrum disorder and specific language impairment. *Journal of Child Psychology and Psychiatry*, 58(11), 1251–1263. <https://doi.org/10.1111/jcpp.12734>
- Hardan, A. Y., Gengoux, G. W., Berquist, K. L., Libove, R. A., Ardel, C. M., Phillips, J., Frazier, T. W., & Minjarez, M. B. (2015). A randomized controlled trial of Pivotal Response Treatment Group for parents of children with autism. *Journal of Child Psychology and Psychiatry*, 56(8), 884–892. <https://doi.org/10.1111/jcpp.12354>
- Harrington, J. W., & Allen, K. (2014). The clinician's guide to autism. *Pediatrics in Review*, 35(2), 62–78; quiz 78. <https://doi.org/10.1542/pir.35-2-62>
- Hartley, C., Bird, L.-A., & Monaghan, P. (2019). Investigating the relationship between fast mapping, retention, and generalisation of words in children with autism spectrum disorder and typical development. *Cognition*, 187, 126–138. <https://doi.org/10.1016/j.cognition.2019.03.001>
- Harvey, D. Y., & Schnur, T. T. (2016). Different Loci of Semantic Interference in Picture Naming

- vs. Word-Picture Matching Tasks. *Frontiers in Psychology*, 7. <https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00710>
- Hinojosa, J. A., Méndez-Bértolo, C., & Pozo, M. A. (2010). Looking at emotional words is not the same as reading emotional words: Behavioral and neural correlates. *Psychophysiology*, 47(4), 748–757. <https://doi.org/10.1111/j.1469-8986.2010.00982.x>
- Icht, M., Zukerman, G., Ben-Itzhak, E., & Ben-David, B. M. (2021). Keep it simple: Identification of basic versus complex emotions in spoken language in individuals with autism spectrum disorder without intellectual disability: A meta-analysis study. *Autism Research*, 14(9), 1948–1964. <https://doi.org/10.1002/aur.2551>
- LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience*, 7(1), Article 1. <https://doi.org/10.1038/nrn1825>
- Larrouy-Maestri, P., Poeppel, D., & Pell, M. D. (2024). The Sound of Emotional Prosody: Nearly 3 Decades of Research and Future Directions. *Perspectives on Psychological Science*. <https://doi.org/10.1177/17456916231217722>
- Le Sourn-Bissaoui, S., Aguert, M., Girard, P., Chevreuril, C., & Laval, V. (2013). Emotional speech comprehension in children and adolescents with autism spectrum disorders. *Journal of Communication Disorders*, 46(4), 309–320. <https://doi.org/10.1016/j.jcomdis.2013.03.002>
- Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q. F., Šmíra, M., & Epskamp, S. (2019). JASP: Graphical statistical software for common statistical designs. *Journal of Statistical Software*, 88, 1–17. <https://doi.org/10.18637/jss.v088.i02>
- Loveland, K. A., Tunali-Kotoski, B., Chen, Y. R., Ortegon, J., Pearson, D. A., Brelsford, K. A., & Gibbs, M. C. (1997). Emotion recognition in autism: Verbal and nonverbal information. *Development and Psychopathology*, 9(3), 579–593.



<https://doi.org/10.1017/S0954579497001351>

- MacFarlane, H., Salem, A. C., Chen, L., Asgari, M., & Fombonne, E. (2022). Combining voice and language features improves automated autism detection. *Autism Research, 15*(7), 1288–1300. <https://doi.org/10.1002/aur.2733>
- Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics, 18*(1), 50–60. <https://doi.org/10.1214/aoms/1177730491>
- Matsumoto, K., Sugiyama, T., Saito, C., Kato, S., Kuriyama, K., Kanemoto, K., & Nakamura, A. (2016). Behavioral Study on Emotional Voice Perception in Children with Autism Spectrum Disorder. *Journal of Pediatric Neuropsychology, 2*(3–4), 108–118. <https://doi.org/10.1007/s40817-016-0021-0>
- Mätziga, S., Druks, J., Masterson, J., & Vigliocco, G. (2009). Noun and verb differences in picture naming: Past studies and new evidence. *Cortex, 45*(6), 738–758. <https://doi.org/10.1016/j.cortex.2008.10.003>
- Mayer-Benarous, H., Benarous, X., Vonthron, F., & Cohen, D. (2021). Music Therapy for Children With Autistic Spectrum Disorder and/or Other Neurodevelopmental Disorders: A Systematic Review. *Frontiers in Psychiatry, 12*. <https://www.frontiersin.org/articles/10.3389/fpsy.2021.643234>
- 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>
- Miller, G. A., & Fellbaum, C. (1991). Semantic networks of english. *Cognition, 41*(1), 197–229.

[https://doi.org/10.1016/0010-0277\(91\)90036-4](https://doi.org/10.1016/0010-0277(91)90036-4)

Monrad-Krohn, G. H. (1947). Dysprosody or altered melody of language. *Brain: A Journal of Neurology*, 70(Pt 4), 405–415. <https://doi.org/10.1093/brain/70.4.405>

Naigles, L. R., & Tek, S. (2017). “Form is easy, meaning is hard” revisited: (Re) characterizing the strengths and weaknesses of language in children with autism spectrum disorder. *Wiley Interdisciplinary Reviews. Cognitive Science*, 8(4). <https://doi.org/10.1002/wcs.1438>

Naigles, L. R., Yi, L., & Zhao, J. (2023). Editorial for the Special Issue - Language and Reading in Autism: Perspectives from Chinese. *Reading and Writing*, 36(6), 1359–1368. <https://doi.org/10.1007/s11145-023-10445-z>

Nicoletti, D. C. da C. S., & Cantacini, E. (2021). The importance of speech therapy rehabilitation in the search for the independence of persons with autistic spectrum disorders. *Health and Society*, 1(06), Article 06. <https://doi.org/10.51249/hs.v1i06.540>

Nielson, K. A., & Arentsen, T. J. (2012). Memory modulation in the classroom: Selective enhancement of college examination performance by arousal induced after lecture. *Neurobiology of Learning and Memory*, 98(1), 12–16. <https://doi.org/10.1016/j.nlm.2012.04.002>

Norbury, C. F., & Sparks, A. (2013). Difference or disorder? Cultural issues in understanding neurodevelopmental disorders. *Developmental Psychology*, 49(1), 45–58. <https://doi.org/10.1037/a0027446>

Nygaard, L. C., & Lunders, E. R. (2002). Resolution of lexical ambiguity by emotional tone of voice. *Memory & Cognition*, 30(4), 583–593. <https://doi.org/10.3758/bf03194959>

O’Connor, K. (2012). Auditory processing in autism spectrum disorder: A review. *Neuroscience and Biobehavioral Reviews*, 36(2), 836–854.

<https://doi.org/10.1016/j.neubiorev.2011.11.008>

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, 40(2), 184–202. <https://doi.org/10.1598/RRQ.40.2.3>

Pell, M. D., & Baum, S. R. (1997). The ability to perceive and comprehend intonation in linguistic and affective contexts by brain-damaged adults. *Brain and Language*, 57(1), 80–99. <https://doi.org/10.1006/brln.1997.1638>

Qasim, S. E., Mohan, U. R., Stein, J. M., & Jacobs, J. (2023). Neuronal activity in the human amygdala and hippocampus enhances emotional memory encoding. *Nature Human Behaviour*, 7(5), Article 5. <https://doi.org/10.1038/s41562-022-01502-8>

Rogers, S. J., Hayden, D., Hepburn, S., Charlifue-Smith, R., Hall, T., & Hayes, A. (2006). Teaching young nonverbal children with autism useful speech: A pilot study of the Denver Model and PROMPT interventions. *Journal of Autism and Developmental Disorders*, 36(8), 1007–1024. <https://doi.org/10.1007/s10803-006-0142-x>

Savekar, A., Tarai, S., Singh, M., & Raag, R. (2021). Impact of Prosocial and Positive Emotional Languages of Teacher on Habits of Students: Some Quantitative Empirical Evidence. *Journal of Language and Linguistic Studies*, 17(0), Article 0. <https://www.jlls.org/index.php/jlls/article/view/1952>

Schafer, G., Williams, T. I., & Smith, P. T. (2013). Which Words are Hard for Autistic Children to Learn? *Mind & Language*, 28(5), 661–698. <https://doi.org/10.1111/mila.12038>

Scheerer, N. E., Shafai, F., Stevenson, R. A., & Iarocci, G. (2020). Affective Prosody Perception and the Relation to Social Competence in Autistic and Typically Developing Children. *Journal of Abnormal Child Psychology*, 48(7), 965–975. <https://doi.org/10.1007/s10802-020-00644-5>

- Thurm, A., Lord, C., Lee, L.-C., & Newschaffer, C. (2007). Predictors of Language Acquisition in Preschool Children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *37*(9), 1721–1734. <https://doi.org/10.1007/s10803-006-0300-1>
- Wang, J.-E., & Tsao, F.-M. (2015). Emotional prosody perception and its association with pragmatic language in school-aged children with high-function autism. *Research in Developmental Disabilities*, *37*, 162–170. <https://doi.org/10.1016/j.ridd.2014.11.013>
- Wang, L., Beaman, C. P., Jiang, C., & Liu, F. (2022). Perception and Production of Statement-Question Intonation in Autism Spectrum Disorder: A Developmental Investigation. *Journal of Autism and Developmental Disorders*, *52*(8), 3456–3472. <https://doi.org/10.1007/s10803-021-05220-4>
- West, M. J., Copland, D. A., Arnott, W. L., Nelson, N. L., & Angwin, A. J. (2017). Effects of emotional prosody on novel word learning in relation to autism-like traits. *Motivation and Emotion*, *41*(6), 749–759. <https://doi.org/10.1007/s11031-017-9642-6>
- West, M. J., Copland, D. A., Arnott, W. L., Nelson, N. L., & Angwin, A. J. (2018). Effects of Prosodic and Semantic Cues on Facial Emotion Recognition in Relation to Autism-Like Traits. *Journal of Autism and Developmental Disorders*, *48*(8), 2611–2618. <https://doi.org/10.1007/s10803-018-3522-0>
- World Health Association. (2020). *International Classification of Diseases (ICD-11)*.
- Yin, L., Lai, J., Zhang, S., Bao, C., & Zhao, J. (2023). Early reading skills in Chinese children with autism spectrum disorder. *Reading and Writing*, *36*(6), 1515–1537. <https://doi.org/10.1007/s11145-022-10300-7>
- Yong-Xiang, C., Jie, N. I. U., & Li-Qi, Z. H. U. (2016). Is there a Noun Bias for Mandarin-speaking Infants: From the Perspective of Vocabulary Size. *Journal of Psychological Science*,

39(3), 600.

- Zeroual, A., Harrou, F., Dairi, A., & Sun, Y. (2020). Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study. *Chaos, Solitons, and Fractals*, 140, 110121. <https://doi.org/10.1016/j.chaos.2020.110121>
- Zhang, D. (2012). A Contrastive Study of Verb+Nouns in Chinese and English. *Journal of Tangshan Teachers College*.
- Zhang, M., Xu, S., Chen, Y., Lin, Y., Ding, H., & Zhang, Y. (2022). Recognition of affective prosody in autism spectrum conditions: A systematic review and meta-analysis. *Autism: The International Journal of Research and Practice*, 26(4), 798–813. <https://doi.org/10.1177/1362361321995725>
- Shen. (2012). Rethinking "nominal verbs": problems and responses. *Chinese Language Teaching in the World*, 26(1), 3–17. <https://doi.org/10.13724/j.cnki.ctiw.2012.01.001>

## Appendix

**Table S1**

*Basic characteristics of included subjects*

Subject	Age	Gender	PPVT	CARS	Intellectual Age
Subject G	6	Male	45	33 Moderate	4
Subject A	5	Female	13	34 Moderate	<3
Subject K	5	Male	14	38 High	<3
Subject H	5	Male	17	33 Moderate	<3
Subject B	8	Male	44	31 Moderate	4
Subject C	9	Male	13	43 High	<3
Subject E	7	Male	28	31 Moderate	<3
Subject I	4	Male	16	36 High	<3
Subject D	5	Male	18	33 Moderate	<3
Subject J	8	Male	23	44 High	<3
Subject F	7	Female	9	41 High	<3
Subject L	8	Male	9	38 High	<3

*Note:* PPVT means Peabody Picture Vocabulary Test, CARS means Childhood Autism Rating Scale.

**Table S2***Differences in Gender of emotional prosody intervention*

	Emotional Prosody	U	p
Post Test			
Receptive Vocabulary	Positive	9.50	1.000
	Neutral	9.00	.912
Expressive Vocabulary	Positive	9.50	1.000
	Neutral	10.50	1.000
Follow-up Test			
Receptive Vocabulary	Positive	7.00	.583
	Neutral	10.00	1.000
Expressive Vocabulary	Positive	10.00	1.000
	Neutral	12.00	.746

Note: Male (n=10), Female (n=2).

**Table S3***Differences in CARS of emotional prosody intervention*

	Emotional Prosody	U	p
Post Test			
Receptive Vocabulary	Positive	8.50	.107
	Neutral	10.50	.250
Expressive Vocabulary	Positive	6.50	.073
	Neutral	2.00	.011
Follow-up Test			
Receptive Vocabulary	Positive	10.50	.252
	Neutral	8.00	.127
Expressive Vocabulary	Positive	7.50	.107
	Neutral	2.50	.016

Note: High (n=6), Moderator (n=6), CARS means Childhood Autism Rating Scale.



**Table S4***Differences in intellectual age of emotional prosody intervention*

	Emotional Prosody	U	p
Post Test			
Receptive Vocabulary	Positive	9.50	1.000
	Neutral	14.50	.377
Expressive Vocabulary	Positive	13.50	.512
	Neutral	NA	NA
Follow-up Test			
Receptive Vocabulary	Positive	12.50	.661
	Neutral	13.00	.590
Expressive Vocabulary	Positive	11.00	.914
	Neutral	14.00	.449

Note: intellectual age equal to four (n=2), intellectual age less than three (n=12). NA: one of the groups has an SD of zero and cannot continue the calculations.