Article

Perceptual and Motor Skills
0(0) 1–18

DOI: 10.1177/0031512516663164
pms.sagepub.com

Auditory Perception, Suprasegmental Speech Processing, and Vocabulary Development in Chinese Preschoolers

Hsiao-Lan S. Wang, I-Chen Chen, and Chun-Han Chiang
Department of Special Education, National Taiwan Normal University, Taipei, Taiwan

Ying-Hui Lai
Department of Electrical Engineering, Yuan Ze University, Taoyuan, Taiwan

Yu Tsao
Research Center for Information Technology Innovation (CITI) at Academia Sinica, Taipei, Taiwan

Abstract

The current study examined the associations between basic auditory perception, speech prosodic processing, and vocabulary development in Chinese kindergartners, specifically, whether early basic auditory perception may be related to linguistic prosodic processing in Chinese Mandarin vocabulary acquisition. A series of language, auditory, and linguistic prosodic tests were given to 100 preschool children who had not yet learned how to read Chinese characters. The results suggested that lexical tone sensitivity and intonation production were significantly correlated with children's general vocabulary abilities. In particular, tone awareness was associated with comprehensive language development, whereas intonation production was associated with both comprehensive and expressive language development. Regression analyses revealed that tone sensitivity accounted for 36% of the unique variance in vocabulary development, whereas intonation production accounted for 6% of the variance in vocabulary development. Moreover, auditory frequency

Corresponding Author:
Hsiao-Lan Wang, Department of Special Education, National Taiwan Normal University, 162, Heping East Road Section 1, Taipei 106, Taiwan.
Email: hlw36@ntnu.edu.tw

discrimination was significantly correlated with lexical tone sensitivity, syllable duration discrimination, and intonation production in Mandarin Chinese. Also it provided significant contributions to tone sensitivity and intonation production. Auditory frequency discrimination may indirectly affect early vocabulary development through Chinese speech prosody.

Keywords
Auditory perception, perception, speech, listening, processing, cognition, learning and memory, Chinese

Introduction

A well-established research finding is that phonological awareness underlying the segmental qualities of language is an important correlate of vocabulary development. Children with inferior expressive vocabulary (e.g., late talkers) usually present delayed phonological development (Williams & Elbert, 2003), and, conversely, children with high expressive vocabulary skills demonstrate superior phonological ability (Smith, McGregor, & Demille, 2006). Although previous research has primarily focused on processing the segmental aspects of speech (e.g., phoneme awareness), recent studies indicate that the ability to process prosody is also important for the development of language skills (Lindfield, Wingfield, & Goodglass, 1999; Nooteboom, 1997; Wood, 2006). Unlike segmental phonology, prosody refers to properties of syllables and speech over phonetic segments, so called suprasegmental phonology, and encompasses the rhythm, tempo, stress, tone, and intonation of a language. Particularly in a tonal and syllable-based language, such as Chinese Mandarin, multiple variations in lexical tones and speech intonation directly represent different meanings at the syllable/word level. Prosodic information may therefore provide a template that facilitates the access of lexical representations in Chinese spoken word recognition.

Given the established interrelationships between vocabulary and phonological awareness in many developmental studies across languages (e.g., Chinese: Huang & Hanley, 1997; English: Carroll, Snowling, Stevenson, & Hulme, 2003; French: Genesee & Jared, 2008; Japanese: Seki, Kassai, Uchiyama, & Koeda, 2008; Spanish: Manrique & Signorini, 1994; Wise, Sevcik, Morris, Lovett, & Wolf, 2007), it is also important to consider prosodic performance in studies of children's language development. Several English investigations have demonstrated the link between stress perception and speech recognition in children (Jusczyk, Cutler, & Redanz, 1993; Werker & Yeung, 2005; Wood, 2006) and adults (Fernald & Mazzie, 1991; McGrath & Summerfield, 1985); yet, very few studies have examined similar relationship in children speaking Chinese Mandarin. Different from English, the prosodic components that convey lexical information are primarily through lexical tone and intonation, not stress. Mandarin has four to five tones (depending on whether the "neutral" tone is included as a separate tone). There is a high-level tone, a high-rising tone, a fall-rising tone, a high-falling tone, and a flat tone. Typically, every syllable in Mandarin is assigned a lexical tone which determines semantic meaning. For instance, in Mandarin Chinese, when the syllable "ma" is pronounced with a high-level tone, it means *mother* [媽]; however, it refers to *horse* [馬] when pronounced with a falling-rising tone. Another aspect of prosody in Mandarin may be that there is a broader range of pitch alterations in a word/sentence than the range that is typically utilized in English (Chen, Robb, Gilbert, & Lerman, 2001). Speech intonation interacts commonly with lexical tones to express grammatical contrasts in the sentence context. For instance, it is common that questions tend to have high or rising intonation on the final syllables with a neutral tone whereas statements tend to have low or falling intonation with various tones (Liu, Tsao, & Kuhl, 2007). It has been shown that Mandarin speakers have not only a broader range of pitch fluctuation but also a higher frequency and sharpness in their pattern of pitch change. This stands in contrast to the English pattern in which there is a more shallow pitch range and which demonstrates a more gradient pattern of pitch transitions (Yuan & Liberman, 2014).

Notably, there was a recent study longitudinally exploring the influential factors upon vocabulary development in 264 Chinese children aged 4 to 10 years (Song et al., 2015). The findings showed that preschool children's (5–6 years old) tone detection and syllable deletion were significant correlates of their vocabulary size and growth rate. It seems that better tone detection and syllable deletion at early age could predict better vocabulary size and progress later on. On the other hand, Zhou, Su, Crain, Gao, and Zhan (2012) reported that Chinese children as young as four years old were able to be attentive to intonational cues in sentence comprehension to resolve ambiguities. For example, in Mandarin Chinese, a *wh*-word can be presented as a question-marker within an inquiry or an indefinite noun phrase within a statement. A rising intonation with a *wh*-word usually refers to the question expression; whereas a flat intonation indicates the statement purpose. Zhou et al. (2012) tested 38 preschool children and found that these children can reach about 97% and 87% accuracy in perceiving whether sentences were produced with a rising intonation or a flat intonation. The authors suggested that the processing of some prosodic information may be particularly crucial for language and communication development in Chinese. It is worth future research in discussion of the correlation between children's prosodic processing skills and early communication acquisition.

Moreover, another issue may be raised concerning children's acquisition of prosodic features of Mandarin in relation to language learning. From the perceptual point of view, tone awareness may be studied in relation to pitch perception because fundamental frequency (F_0) is the primary and sufficient acoustic cue. Examples from the second-language learners of Chinese have highlighted the effect of language learning experience on these learners' sensitivities to pitch variation (Polich, 2007; Wang & Kuhl, 2003; Zhao & Kuhl, 2015). On the other hand, the training of suprasegmental contrasts of Chinese Mandarin to English listeners may help mastery of acoustic analyses of pitch height and pitch contour in these nonnative learners (Wang, Jongman, & Sereno, 2003). A study by Alexander, Wong, and Bradlow (2005) proposed that Chinese children's tone acquisition may be related to their sensitivity of fundamental frequency complexity. It is thus interesting to collect some acoustic data on children's acquirement of lexical tones and intonations. The findings might provide insightful information regarding the suprasegmental speech processing and vocabulary development in Chinese Mandarin.

To extend the previous findings, the purpose of this study seeks to understand the role of prosody in the development of Chinese vocabulary knowledge. Specifically, we wonder how perceptual processing of prosodic characteristics in Chinese Mandarin is related with vocabulary development in preschool children. By studying children without formal reading instruction, one can exclude the influential factor that the results are due to schooling rather than skills that emerge previous to engaging in learning how to read. Furthermore, there may be some basic auditory factors that affect young children's acquisition of prosodic sensitivity. In this exploratory study, the focus was children's auditory discrimination of sound frequency and duration because both signals appear salient in lexical tones and intonations. The data may help us to elucidate the sensory basis of prosodic sensitivity in Chinese Mandarin.

Method

Participants and procedure

One hundred children from Taipei City were recruited (age range = 55–65 months; M age = 60 months, $SD = 4.2$; 50 females) and they had no reported hearing impairment, specific language disorders, or developmental delays. Our research information was announced to the public from July 2013 to July 2014. Children and parents received the message and voluntarily joined the study. All participant children were monolingual Mandarin Chinese speakers, and they were measured at the end of their final semester in kindergarten. Informed consent was obtained from the participants' parents and the researchers obtained verbal and written consents from the participants before measuring. The informed consent procedure was approved by the Research Ethics Committee at National Taiwan University (NTU-REC No.: 201306EM020).

The prereading levels of all participants were self-reported by their kindergarten teachers. Children who could correctly recognize 37 Zhuyin or 26 Pinyin symbols and who can use phonetic spelling system were excluded. Three test sessions, each taking about 30 minutes, were administered one-on-one by the researcher in a quiet room so as to have minimal distraction. There was a sufficient break between each test session. The testing batteries were administered at a time most convenient to the teachers. All tasks had practice trials, and no task was administered until the researcher was sure that the participant fully understood the instructions.

Measures

Vocabulary assessment: Chinese expressive and receptive vocabulary development. The Receptive and Expressive Vocabulary Test (REVT; Huang, Jian, Zhu, & Lu, 2010) was used to examine preschoolers' language development. This is a common test used in Taiwan to understand children who obviously fall behind their peers in oral vocabulary proficiency and to assess receptive and expressive vocabulary strengths and weaknesses. The assessment is designed for children aged three to six years old. The assessment is divided into two parts to examine children's receptive and expressive language abilities. Within each part, there are four subtests, including naming, category, definition, and reasoning tests. Children are shown pictures and are read a question by the administrator. In the receptive part, children are then asked to point to the picture that is related to the question. For instance, the administrator may read a word (or a sentence) to the child, "Can you tell me where the cat is?" and the child needs to point out which picture shows a cat. While in the expressive part, children need to answer the question using spoken language. For example, the administrator may ask the child "What is this?" (indicating a rainbow photo), and the child needs to say "Rainbow" aloud. This evaluation had Cronbach's alphas of .80–.96 and test–retest reliability of .80–.97 in various samples.

Linguistic prosodic assessment: Lexical tone sensitivity. This test adapted an oddity task from Lin, Wang, and Shu (2013). In our test, there were four practice trials and 10 testing trials. Each trial contained one monosyllable with two different lexical tones. In each trial, children heard three syllables sharing the same onset and rime and only two had the same tones (e.g., /hua2/, /hua4/, and /hua2/). Meanwhile, children saw cards of 1, 2, and 3 on the desk that corresponded to each syllable, respectively. They were asked to choose the target syllable with the odd tone they heard from the administrator. The following was the instruction:

> You will hear three syllables, the first one matches with the number '1' marked on the card, the second one matches with the number '2' marked on the card, the third one matches with the number '3' marked on the card. Within these three syllables, one has a different tone compared with the other two. Please pick out the syllable that has a different tone, and indicate the corresponding card on the desk. You have to listen carefully and answer as accurately as possible. Ready?

Half of the testing trials were real spoken syllables, while the rest were nonsense syllables. One point was scored if the participant selected the target syllable. The coefficient of internal consistency Cronbach's alpha coefficient was .63 in this sample.

Linguistic prosodic assessment: Syllable duration discrimination. This task assessed children's prosodic sensitivity at the word level. In this task, children were asked to carefully listen to a pair of identical bisyllabic words and to tell whether two words sound the same or different regarding the last syllable length. Simple consonant–vowel syllables of synthesized voices were used for all testing trials, while the vowel length of [a] was variously controlled with initial voiceless stops. The only variable in this test was the duration of the last syllable in a pair of words. In each word, while the control syllable within each word was set with the duration of 500 seconds, the last lengthened syllable was adjusted by 375 ms or 500 ms resulting in an 875 ms or 1000 ms duration, respectively. The test consisted of 20 word pairs, including 10 the same and 10 the different pairs. Children were asked to tell which pairs were the same and which were different by saying "same" or "different" after each pair was presented, and the instructor recorded all answers. The instruction was as follows:

> Please listen for a pair of words with two syllables that are each about one half second. Your task is to decide whether or not the words are equal in duration. Some word-pairs are equal in duration and some are not. After you hear a pair of words, tell me if you think them to be equal in duration by saying SAME, or not equal in duration by saying DIFFERENT. The word-pair are separated by about two seconds of silence. We now have two trials to practice and there will be twenty trials for testing. Let's start the task and please listen carefully!

Correct answers were determined by accurate discrimination between the two words. Cronbach's alpha coefficient was .79 in this sample.

Linguistic prosodic assessment: Intonation production imitation. This task assessed children's prosodic performance at the sentence level. There were eight questions and eight statements in a set of 16 pictures for each participant. The vocabularies and sentences involved in the task were familiar to the kindergarteners, and the target names (locations or objects) were demonstrated in the pictures. All utterances were controlled for length from four to eight words and were syntactically simple. The eight pictures were arranged in pairs demonstrating 10 utterance pairs. One sentence of the pair was produced with a rising final intonation (i.e., a declarative question) and the other sentence was produced with a falling intonation (i.e., a statement). The utterance stimuli for this task were recorded by a native adult female Chinese speaker from Taipei City. The speaker read each utterance pair aloud and the utterances were recorded onto a laptop computer.

The sentence pairs were then randomized so that alternate forms of the same utterance were not presented sequentially. Children listened to the sentences through headphones. They were instructed to imitate the sentences they heard from the laptop and their utterances were recorded with a microphone directly onto the laptop computer. The instruction was:

> You will see a picture on the computer and hear two expressions. Please repeat the expression you hear and I will record your voice on the computer.

Following the instructions, a practice session was provided and then the actual task began. Children's intonation contours produced from the task were recorded digitally and analyzed with Praat Version 4508 (Boersma & Weenink, 2006). The direction fundamental frequencies of the final pitch change of the intonation contours were examined to determine whether children could perform questions or statements with corresponding intonations. Two trained research assistants recalculated this measurement for the interrater reliability, .91.

Psychoacoustic assessment. All the acoustic stimuli were presented biaurally. The study was updated from the computerized program used by our previous research (see Wang, Huss, Hämäläinen, & Goswami, 2012). Two kinds of experimental paradigms were applied. In the AXB paradigm, children were required to listen to three auditory stimuli sequentially and to choose the sound that was different from the two other sounds. In the two-interval forced-choice paradigm, children were asked to identify the expected answer from two alternatives, according to the instruction. The interstimulus interval was 500 ms. The sounds were set to be presented at 75 dB level to the children.

All children were given five practice trials for each auditory task. Feedback was provided after their answer. The analysis replicated research of Wang et al. (2012) and used an adaptive staircase procedure (Levitt, 1971) with a combined two-down one-up and three-down one-up procedure. A test run may terminate after eight response reversals or after the maximum 40 trials. The final estimated threshold value reflects an individual's auditory sensitivity at around the 79% correct level. The threshold value of 40 was set as the maximum value.

Duration discrimination (AXB). Three cat cartoon pictures in the computer game were presented with three pure tones, respectively. Children had to point out the cat that produced the sound with a different duration from the standard sound. The middle one always stands for the standard sound. Each stimulus was 500-Hz sinusoid with 50-ms rise and fall times. Forty auditory stimuli were linearly generated with different durations ranged from 400 ms to 600 ms.

Frequency discrimination (two-interval forced-choice paradigm). In this test, there were two birds singing rhythmic sounds with five tones on the screen. The target bird

Table 1. Mean participant characteristics.

Variable	*M*	*SD*	Range	
			Min	Max
Age (m)	68.93	4.22	61	78
IQ (TONI-3)	101.04	10.99	75	126
REVT full scale (standard score)	109.74	11.85	79	133
Comprehensive language	108.31	11.91	77	131
Expressive language	109.52	13.23	83	133
Tone sensitivity (Max:10)	5.96	1.15	4	8
Syllable duration (Max: 20)	12.93	2.42	7	19
Intonation production (Max: 16)	9.96	2.43	5	16
Duration discrimination	29.61	7.20	12	37
Frequency discrimination	29.01	8.14	11	39

Note. REVT: Receptive and Expressive Vocabulary Test; TONI-3: Test of Nonverbal Intelligence-Third Edition.

sang with alternately high and low pitches and the other one sang with one identical pitch. The participants were asked to distinguish the target bird that sang high and low. The standard pitch was created with 500 Hz and 50 ms rise and fall times in 200 ms duration. The standard pitch was changed into a continuum of 40 stimuli that had ranged from 500 Hz to 510 Hz in a logarithmic scale. The five-tone songs in both formats had the same interstimulus interval of 50 ms.

Results

Descriptive statistics

Histograms, quintile–quintile plots, and scatterplots were adapted to evaluate and confirm the assumptions of univariate normality and linearity. The assumption of multivariate normality was explored using SPSS. No participant was classified as an outlier. For the natural distribution of individual differences in reading, we included all participants' performance for further statistical analyses. The descriptive statistics of the measures are given in Table 1. No any ceiling or floor effect was found in these measures.

Partial correlations between variables

Partial correlations, controlling for age and IQ, vocabulary development, tone awareness, prosodic performance, and auditory perception across all participants, are presented in Table 2.

Table 2. Partial correlations between vocabulary development, prosodic performance, and auditory perception, controlling for age and IQ.

Variable	2	3	4	5	6	7	8
1. REVT full scale	.63***	.88***	.37**	.03	.25*	.08	−.11
2. Comprehensive language		.26*	.44***	.00	.15	.08	−.13
3. Expressive language			.27*	.05	.26*	.07	−.09
4. Tone sensitivity				−.01	.48***	.19	−.43***
5. Syllable duration					.25*	−.14	−.38***
6. Intonation production						−.06	−.44***
7. Duration discrimination							.17
8. Frequency discrimination							

Note. REVT: Receptive and Expressive Vocabulary Test. *$p < .05$. ***$p < .001$.

The vocabulary development measures, the REVT and its subtests, were significantly intercorrelated. Consistent with other studies in Chinese, children's performance on lexical tone discrimination was associated with vocabulary development. In the present study, tone awareness was significantly correlated with the REVT full scale score ($r = .37$, $p = .002$) and with both receptive vocabulary ($r = .44$, $p < .001$) and expressive vocabulary ($r = .27$, $p = .02$). Moreover, intonation production was significantly related to expressive vocabulary ($r = .26$, $p = .01$) and the REVT full scale score ($r = .25$, $p = .01$). The better children were in intonation production, the greater their vocabulary expression and vocabulary development. On the other hand, the relationship between syllable duration discrimination and vocabulary development did not reach statistical significance. Interestingly, there was a clear correlation between tone awareness and intonation production ($r = .48$, $p < .001$), which may indicate a common sensory mechanism for the processing of speech sounds. In our auditory analyses, we found a significant correlation between tone awareness and auditory frequency discrimination ($r = -.43$, $p < .001$), as well as between intonation production and frequency discrimination ($r = -.44$, $p < .001$). Further, auditory frequency discrimination was significantly associated with the syllable duration test results ($r = -.38$, $p < .001$). It seems that basic auditory frequency perception was related to both children's phonological awareness and linguistic prosodic processing in Mandarin Chinese particularly.

Multiple regression analyses

For the vocabulary analyses shown in Table 3, data from all participants were entered into a series of three-step fixed-entry multiple regression equations, entering IQ at Step 1 and then prosodic measures in turn at Step 2. As expected, tone sensitivity was the most powerful predictor of vocabulary development in Chinese, $\Delta R^2 = .36$, $p < .01$. Further, when intonation production was controlled, the

Table 3. R^2 change in separate linear regression equations examining the contribution of linguistic prosodic performance to vocabulary development.

Step		β	ΔR^2
DV: REVT full scale			
1	IQ	.22	.05
2	Tone sensitivity	.36	.36**
2	Syllable duration	.04	.00
2	Intonation production	.25	.06*

Note. DV: Dependent Variable; REVT: Receptive and Expressive Vocabulary Test.
$^{*}p < .05$. $^{**}p < .01$.

contribution of tone sensitivity to vocabulary development decreased but remained statistically significant, $\Delta R^2 = .75$, $p < .05$. Additionally, intonation production accounted for a substantial amount of the unique variance (6%) in vocabulary development. Lexical tone sensitivity appears to be the strongest predictor of Chinese vocabulary development in preschoolers.

Regarding the contribution of auditory perception, similar three-step multiple regression analyses were run using all the children and taking the total number of correct responses as the dependent variable. Each equation again entered age at Step 1 and IQ at Step 2 and then one of the auditory processing measures at Step 3. The results are shown in Table 4. As to the unique variance in tone sensitivity and intonation production accounted for by each auditory test, the results showed that frequency discrimination provided the most significant contribution to tone awareness, $\Delta R^2 = .16$, $p < .001$, as well as in intonation production, $\Delta R^2 = .19$, $p < .001$. In contrast, duration discrimination explained little of the variance in both tasks.

Discussion

In examining the contribution of prosodic processing to children's vocabulary development, this study showed the evidence that intonation production is related to vocabulary expression particularly in children speaking Mandarin Chinese. Additionally, children's lexical tone sensitivity was associated with vocabulary development, both with receptive and expressive vocabularies. Further, tone sensitivity and intonation production were significantly intercorrelated, suggesting there might be a mutual perceptual process underlying both tests. Auditory frequency discrimination was significantly related to both tests, indicating the perceptual importance of sound frequency discrimination in early Chinese speech learning and development.

Table 4. R^2 change in separate linear regression equations examining the contribution of auditory perception to tone sensitivity and intonation production.

Step		β	ΔR^2
DV: Tone sensitivity			
1	Age	.28	.08*
2	IQ	.15	.02
3	Duration discrimination	.19	.03
3	Frequency discrimination	−.41	.16***
DV: Intonation production			
1	Age	.03	.00
2	IQ	.16	.03
3	Duration discrimination	−.06	.00
3	Frequency discrimination	−.44	.19***

*$p < .05$. ***$p < .001$.

The strong connection between lexical tone sensitivity and vocabulary development in this study supports some claims that lexical tone is crucial and significantly accounts for individual variance in kindergarteners' word acquisition in Chinese, both with spoken or written forms (Walley, 1993). For example, in a recent investigation by Tong, Tong, and McBride-Chang (2015), lexical tone identification was found to be significantly correlated with expressive vocabulary development and Chinese word recognition in 199 five- to six-year-old kindergarten children speaking Cantonese. Lexical tone identification was significantly associated with vocabulary development ($r = .32$) and accounted for 5% of the variance in their vocabulary acquisition and word recognition, respectively, even when age, IQ, segmental phonological awareness, and morphological awareness were controlled. Clearly, tone awareness of young Cantonese-speaking children could contribute uniquely to early expressive vocabulary development and also predict early word recognition after the influence of vocabulary knowledge was controlled. In the present sample, children's lexical tone sensitivity was significantly related to both comprehensive language ($r = .44$) and expressive language abilities ($r = .27$). Tone sensitivity accounted 36% of the variance in their general vocabulary development. It also uniquely contributed 7.5% of the variance in vocabulary development even after accounting the intonation production. These values seem very close to the data from Tong et al. (2015).

Concerning children's lexical development and word learning, speech perception, articulation, and vocabulary development usually occur simultaneously and may rely on each other. It is well known that early phonetic abilities are important for later language development. For instance, a study by Tsao, Liu, and Kuhl (2004) indicated that early speech discrimination in six-month-old infants could predict vocabulary size in their second year of life. In typical language development, the first words usually emerge around nine months of age. The emergence requires a lexical representation not only for an approximation of the word meaning but also an encoding of the sound shape of the word, to be able to produce it consistently across contexts. This production depends not only on the articulatory abilities but also on the phonological details of the early lexical representations. Research has shown that children with higher phonological skills can learn both novel nonnative and native names more accurately, compared with children with lower phonological skills (Hu & Schuele, 2005). On the other hand, studies of second language learning indicate that training of lexical tone identification abilities may significantly improve nonnative speakers' word identification proficiency in a tonal language. For example, Wang et al. (2003) reported that training American speakers to perceive and identify Mandarin tones can improve their pitch perception and Mandarin word production. This effect was even retained when tested six months after training. Similar findings from Cooper and Wang (2013) also revealed that short-term Cantonese tone identification training with native English nonmusicians had a significant effect on word identification proficiency as well, supporting phonetic-phonological-lexical continuity in word learning. This implies that lower level auditory and perceptual sensitivity of phonetic contrasts may contribute to attainment of a higher level lexical task.

Meanwhile, an interesting question was raised as to whether fundamental auditory processing abilities to distinguish pitch would be related to learning words' meaning distinguished by tonal cues. As tone is characterized mainly in terms of its acoustic properties, fundamental frequency heights, and contours, we examined children's auditory sensitivity to pure sound frequency and duration. Our data extended previous findings and pinpointed the involvement of bottom-up processes in the spoken-word learning domain by showing a substantive relationship between auditory frequency perception, tone awareness, and prosodic processing. Auditory frequency discrimination was significantly correlated with lexical tone sensitivity ($r = -.43$), syllable duration discrimination ($r = -.38$), and intonation production ($r = -.44$). Frequency discrimination accounted for a significant amount of the variance in tone sensitivity (16.2%) and intonation production (18.6%). The present findings highlight the role of bottom-up processes during phonological and prosodic processing, as featured in some models of speech perception (McClelland & Elman, 1986; Norris, McQueen, & Cutler, 2000), in that lower level sensory information can be critical during the early stages of language acquisition. Indeed, a study by Bradlow and Alexander (2007) also reported that when lower level acoustic information was highly degraded, listeners became unable to effectively utilize higher level contextual information in sentence recognition.

In addition to lexical tone, intonation is another important suprasegmental component of language, and it has been characterized as the “music” or “tune” of speech. Intonation is formed by changes in fundamental frequency (F_0) across speech units, and it is recognized perceptually as changes in pitch too. Typically in a grammatical role, intonation functions when prosodic features serve to indicate a contrast, such as that between a statement and a question. Studies have suggested that infants demonstrate a preference for intonation patterns they have heard prenatally (DeCasper & Spence, 1986) and can discriminate intonation contour direction within days after birth (Nazzi, Floccia, & Bertoncini, 1998). On the other hand, skills in production of intonation have also been shown to follow a developmental pattern of growth beginning in early infancy. The intonation contours of the cries of young infants have been reported to reflect the contours of the language to which the infant has been exposed (Mampe, Friederici, Christophe, & Wermke, 2009). Similarly, infants have been observed to use prosody that is characteristic of their native language during the period of transition from babbling to single-word production (Halle, Boysson-Bardies, & Vihman, 1991).

Development in ability to produce intonation has been reported across the preschool to school age range (Van Der Meulen, Janssen, & den Os, 1997; Wells, Peppe, & Goulandris, 2004). In the present study, intonation production of Chinese preschool children was positively correlated with language development, particularly expressive language development. This is consistent with some research in English suggesting that children’s comprehension of intonation seems to occur in line with their receptive and expressive language development. For example, Wells et al. (2004) reported that children’s intonation skills (i.e., Profiling Elements of Prosodic Systems-Child Assessment; Peppe & McCann, 2003) were positively correlated with children’s sentence-level comprehension on the Test for the Reception of Grammar (Bishop, 1989) and their sentence-level production on the Clinical Evaluation of Language Fundamentals-Revised (Semel, Wiig, & Secord, 1987). Stojanovik, Setter, and van Ewijk (2007) also reported a significant correlation between typically developing children’s intonation skills and their receptive and expressive language development (i.e., Assessment of Comprehension and Expression battery; Adams, Cooke, Crutchley, Hesketh, & Reeves, 2001). Our data regarding Mandarin Chinese therefore provide crosslinguistic support for this result. Moreover, the regression analysis reveals that both tone awareness and intonation production significantly contribute to language development in Chinese. Tone sensitivity accounted for a greater amount of the unique variance in vocabulary development. Perhaps young children’s capability to process contrasts in prosodic patterns of speech exerts a greater effect in the early stage of language development when they are not ready to demonstrate systematic use of intonation in their vocalization.

As noted above, variation in F_0 is important for the recognition of contrasts in several suprasegmental components of speech, including lexical tones and intonations.

Although it is unlikely that F_0 variation serves as the only acoustic property that contributes to the perception of speech prosody, F_0 variation dominates. This is consistent with our data that auditory frequency discrimination was significantly associated with tone sensitivity, syllable duration discrimination, and intonation production. While F_0 contour furnishes the listener with the most important information for recognizing speech intonation contrasts that mark declarative or interrogative utterances (Cooper & Sorensen, 1981; Ladd, 1996; Lehiste, 1970, 1976), variation in F_0 contours may take place in conjunction with other acoustic correlates, such as intensity and duration patterns. However, we did not find this relationship between auditory duration perception and the linguistic prosodic tests. In Mandarin Chinese, frequency variation is still the primary cue for the recognition and production of speech intonation.

Limitations and conclusions

However, there are limitations to the study. First, in the current study, children's perception or comprehension of speech intonation itself was not assessed. In addition, linguistic prosodic processing was assessed using self-designed tasks, and no assessment has yet been made using standardized assessments in Chinese. Our measures might not be reliable and valid sufficiently across all children. A more comprehensive and accurate assessment that measures various processing aspects of speech prosody in Mandarin Chinese is needed to draw a more representative picture of children's language skills. Additionally, longitudinal studies evaluating perception and production of speech prosody beginning in kindergarten are recommended to provide further insights into the trajectories of development in intonation, oral language skills, and literate language skills. These could potentially lead to early identification and/or intervention strategies, thereby reducing the effect of some language difficulties.

The current study provided support for the consideration of intonation production as one of the language skills that may influence developing expressive vocabulary in Chinese preschool children. Consistent with previous studies, we also found a strong correlation between lexical tone sensitivity and general vocabulary development in Chinese (Meng et al., 2005). Furthermore, children's tone sensitivity and intonation production were associated. Fundamentally, auditory perception of sound frequency significantly explained substantial variance in both tone sensitivity and intonation production and appeared to be a critical perceptual basis for Chinese language development in young children.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This article was supported by the Minister of Science and Technology (102-2410-H-003-001-MY2) in Taiwan.

References

Adams, C., Cooke, R., Crutchley, A., Hesketh, A., & Reeves, D. (2001). *Assessment of comprehension and expression 6–11 (ACE 6–11)*. Windsor, Canada: NFER-Nelson.

Alexander, J. A., Wong, P. C., & Bradlow, A. R. (2005, September). Lexical tone perception in musicians and non-musicians. In *Interspeech* (pp. 397–400). Lisbon, Portugal.

Bishop, D. V. (1989). *Test for the reception of grammar: (TROG)*. London, England: Medical Research Council.

Boersma, P., & Weenink, D. (2006). *Praat (Version 4.5) [Computer software]*. Amsterdam, Netherlands: Institute of Phonetic Sciences.

Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America*, *121*(4), 2339–2349.

Carroll, J. M., Snowling, M. J., Stevenson, J., & Hulme, C. (2003). The development of phonological awareness in preschool children. *Developmental Psychology*, *39*(5), 913.

Chen, Y., Robb, M., Gilbert, H., & Lerman, J. (2001). Vowel production by Mandarin speakers of English. *Clinical Linguistics & Phonetics*, *15*(6), 427–440.

Cooper, W. E., & Sorenson, J. (1981). *Fundamental frequency in speech production*. New York, NY: Springer-Verlag.

Cooper, A., & Wang, Y. (2013). Effects of tone training on Cantonese tone-word learning. *The Journal of the Acoustical Society of America*, *134*(2), EL133–EL139.

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, *9*(2), 133–150.

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, *27*(2), 209.

Genesee, F., & Jared, D. (2008). Literacy development in early French immersion programs. *Canadian Psychology/Psychologie canadienne*, *49*(2), 140.

Halle, P. A., Boysson-Bardies, B. D., & Vihman, M. M. (1991). Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants. *Language and Speech*, *34*, 299–318.

Hu, C. F., & Schuele, C. M. (2005). Learning nonnative names: The effect of poor native phonological awareness. *Applied Psycholinguistics*, *26*(3), 343–362.

Huang, H. S., & Hanley, J. R. (1997). A longitudinal study of phonological awareness, visual skills, and Chinese reading acquisition among first-graders in Taiwan. *International Journal of Behavioral Development*, *20*(2), 249–268.

Huang, R., Jian, X., Zhu, L., & Lu, L. (2010). *Receptive and expressive vocabulary test*. Taipei, Taiwan: Psychological Publishing.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*(3), 675–687.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Lehiste, I. (1976). Influence of fundamental frequency pattern on the perception of duration. *Journal of Phonetics*, *4*, 113–117.

Levitt, H. C. C. H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2B), 467–477.

Ladd, D. R. (1996). *Intonational phonology*. Cambridge, England: Cambridge University Press.

Lindfield, K. C., Wingfield, A., & Goodglass, H. (1999). The role of prosody in the mental lexicon. *Brain and Language*, *68*(1), 312–317.

Lin, C. Y., Wang, M., & Shu, H. (2013). The processing of lexical tones by young Chinese children. *Journal of Child Language*, *40*(04), 885–899.

Liu, H. M., Tsao, F. M., & Kuhl, P. K. (2007). Acoustic analysis of lexical tone in Mandarin infant-directed speech. *Developmental Psychology*, *43*(4), 912–917.

Mampe, B., Friederici, A. D., Christophe, A., & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current Biology*, *19*(23), 1994–1997.

Manrique, A. M. B., & Signorini, A. (1994). Phonological awareness, spelling and reading abilities in Spanish-speaking children. *British Journal of Educational Psychology*, *64*(3), 429–439.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86.

McGrath, M., & Summerfield, Q. (1985). Intermodal timing relations and audio-visual speech recognition by normal-hearing adults. *The Journal of the Acoustical Society of America*, *77*(2), 678–685.

Meng, X., Sai, X., Wang, C., Wang, J., Sha, S., & Zhou, X. (2005). Auditory and speech processing and reading development in Chinese school children: Behavioural and ERP evidence. *Dyslexia*, *11*(4), 292–310.

Nazzi, T., Floccia, C., & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior and Development*, *21*(4), 779–784.

Nooteboom, S. G. (1997). The prosody of speech: Melody and rhythm. In W. J. Hardcastle & J. Laver (Eds.), *The handbook of phonetic sciences* (pp. 640–673). Oxford, England: Blackwell Publishers.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, *23*(3), 299–325.

Peppe, S., & McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: The PEPS-C test and the revised version. *Clinical Linguistics & Phonetics*, *17*, 345–354.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148.

Seki, A., Kassai, K., Uchiyama, H., & Koeda, T. (2008). Reading ability and phonological awareness in Japanese children with dyslexia. *Brain and Development*, *30*(3), 179–188.

Semel, E., Wiig, E., & Secord, W. (1987). *Clinical evaluation of language fundamentals—Revised*. Cleveland, OH: The Psychological Corporation/Harcourt Brace Jovanovich.

Smith, B. L., McGregor, K. K., & Demille, D. (2006). Phonological development in lexically precocious 2-year-olds. *Applied Psycholinguistics*, *27*(3), 355–375.

Song, S., Su, M., Kang, C., Liu, H., Zhang, Y., McBride-Chang, C., ... Shu, H. (2015). Tracing children's vocabulary development from preschool through the school-age years: An 8-year longitudinal study. *Developmental Science*, *18*(1), 119–131.

Stojanovik, V., Setter, J., & van Ewijk, L. (2007). Intonation abilities of children with Williams syndrome: A preliminary investigation. *Journal of Speech, Language, and Hearing Research*, *50*, 1606–1617.

Tong, X., Tong, X., & McBride-Chang, C. (2015). Tune in to the tone: Lexical tone identification is associated with vocabulary and word recognition abilities in young Chinese children. *Language and Speech*, *58*(4), 441–458.

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, *75*, 1067–1084.

Van Der Meulen, S., Janssen, P., & Den Os, E. (1997). Prosodic abilities in children with specific language impairment. *Journal of Communication Disorders*, *30*(3), 155–170.

Walley, A. C. (1993). The role of vocabulary development in children′ s spoken word recognition and segmentation ability. *Developmental Review*, *13*(3), 286–350.

Wang, H. L. S., Huss, M., Hämäläinen, J. A., & Goswami, U. (2012). Basic auditory processing and developmental dyslexia in Chinese. *Reading and Writing*, *25*(2), 509–536.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America*, *106*, 3649–3658.

Wang, Y., & Kuhl, P. K. (2003). Evaluating the 'critical period' hypothesis: perceptual learning of Mandarin tones in American adults and American children at 6, 10 and 14 years of age. *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona*, pp. 1537–1540.

Wells, B., Peppe, S., & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language*, *31*(4), 749–778.

Werker, J. F., & Yeung, H. H. (2005). Infant speech perception bootstraps word learning. *Trends in cognitive sciences*, *9*(11), 519–527.

Williams, A. L., & Elbert, M. (2003). A prospective longitudinal study of phonological development in late talkers. *Language, Speech, and Hearing Services in School*, *34*, 138–153.

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., & Wolf, M. (2007). The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *Journal of Speech, Language, and Hearing Research*, *50*(4), 1093–1109.

Wood, D. (2006). Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. *Canadian Modern Language Review*, *63*, 13–33.

Yuan, J., & Liberman, M. (2014). F_0 declination in English and Mandarin Broadcast News Speech. *Speech Communication*, *65*, 67–74.

Zhao, T. C., & Kuhl, P. K. (2015). Effect of musical experience on learning lexical tone categories. *The Journal of the Acoustical Society of America*, *137*(3), 1452–1463.

Zhou, P., Su, Y. E., Crain, S., Gao, L., & Zhan, L. (2012). Children's use of phonological information in ambiguity resolution: A view from Mandarin Chinese. *Journal of Child Language*, *39*(04), 687–730.

Author Biographies

Hsiao-Lan S. Wang is an assistant professor at the National Taiwan Normal University, Taipei, Taiwan. Hsiao-Lan researches in the field of educational neuroscience and literacy development. Specifically, her research focuses upon reading difficulties and the application of neuroscientific techniques to the study of learning disabilities or mental disorders. She is also a clinically certified Occupational Therapist and a Special Education Teacher at the Secondary School Level in Taiwan. Her recent work mainly uses behavioral and experimental tools to study the identification of children with learning difficulties or mental disorders in Taiwan.

I-Chen Chen is a research assistant in the Department of Special Education in National Taiwan Normal University (NTNU), Taipei, Taiwan. Her research interests involve children language development, reading difficulties, and phonology.

Chun-Han Chiang is a PhD student in the Department of Special Education in National Taiwan Normal University (NTNU). His research interests involve reading disabilities, assistive technology, and augmentative and alternative communication. He recently moves his research area into the context of neuroscience in auditory perception and dyslexia.

Ying-Hui Lai is currently an assistant professor at Department of Electrical Engineering, Yuan Ze University. His research interests focus on hearing aids, cochlear implants, noise reduction, pattern recognition, and source separation.

Yu Tsao is currently an associate research fellow of the Research Center for Information Technology Innovation (CITI) at Academia Sinica, Taiwan. His research interests include speech recognition, audio-coding, deep neural networks, bio-signals, and acoustic modeling. He is currently a member of Institute of Electrical and Electronics Engineers (IEEE) and International Speech Communication Association (ISCA).