CrossMark

Auditory processing, linguistic prosody awareness, and word reading in Mandarin-speaking children learning English

Wei-Lun Chung[1] · Linda Jarmulowicz[2] ·
Gavin M. Bidelman[2,3]

Abstract This study examined language-specific links among auditory processing, linguistic prosody awareness, and Mandarin (L1) and English (L2) word reading in 61 Mandarin-speaking, English-learning children. Three auditory discrimination abilities were measured: pitch contour, pitch interval, and rise time (rate of intensity change at tone onset). Linguistic prosody awareness was measured three ways: Mandarin tone perception, English stress perception, and English stress production. A Chinese character recognition task was the Mandarin L1 reading metric. English L2 word reading was assessed by English real word reading and nonword decoding tasks. The importance of the auditory processing measures to reading was different in the two languages. Pitch contour discrimination predicted Mandarin L1 word reading and rise time discrimination predicted English L2 word reading, after controlling for age and nonverbal IQ. For the prosodic and phonological measures, Mandarin tone perception, but not rhyme awareness, predicted Chinese character recognition after controlling for age and nonverbal IQ. In contrast, English rhyme awareness predicted more unique variance in English real word reading and non-word decoding than did English stress perception and production. Linguistic prosody awareness appears to play a more important role in L1 word reading than phonological awareness; while the reverse seems true for English L2 word reading in Mandarin-speaking children. Taken together, auditory processing, language-specific linguistic prosody awareness, and phonological awareness play different

✉ Wei-Lun Chung
  weilun.chung@gmail.com

[1] Department of Applied Chinese Language and Culture, National Taiwan Normal University, Taipei 10610, Taiwan

[2] School of Communication Sciences and Disorders, Institute for Intelligent Systems, University of Memphis, Memphis, TN 38152, USA

[3] Univeristy of Tennessee Health Sciences Center, Department of Anatomy and Neurobiology, Memphis, TN, USA

⌂ Springer

roles in L1/L2 reading, reflecting different prosodic and segmental structures between Mandarin and English.

Keywords Pitch perception · Amplitude envelope perception · Suprasegmental phonology · Lexical tone · Lexical stress · Literacy

Introduction

In English, word reading relies on the ability to analyze the sounds of language (i.e., phonological awareness) which facilitates the link between sounds and printed words. Children learn the sound-letter mapping, which in turn fosters the ability to decode letters into sounds and sounds into words. Not only is phonological awareness (PA) is a predictor of early reading ability in English (Muter, Hulme, Snowling, & Taylor, 1998), it has become a benchmark in educational curricula in the United States (Common Core State Standard Initiative, 2015). The importance of PA is not unique to English. PA is also a predictor of learning to read in many languages, including Chinese, which has an orthographic system in which printed words cannot be decoded into sounds directly (Hu & Catts, 1998; Siok & Fletcher, 2001). Although PA is well established as a precursor to reading, recent studies have revealed that linguistic prosody awareness (i.e., identifying sound patterns spanning across syllables) and auditory processing of acoustic features signaling prosodic patterns are important to reading acquisition (Zhang & McBride-Chang, 2010, 2014). Hence, the current study examined contributions of auditory processing, linguistic prosody awareness, and PA to word reading in Mandarin as a first language and English as a second language.

Linguistic prosody and reading ability

Some developmental phonology theories suggest that children's phonological representation includes both individual sounds and prosodic patterns (Pierrehumbert, 2003; Vihman & Croft, 2007). According to Cutler's (1996) rhythmic segmentation hypothesis, listeners use the rhythmic structure specific to their native phonological system to segment the speech stream. Developmentally, stressed syllables are perceptually salient cues for English speech segmentation (Echols, 1996). In English, individual sounds are more easily detected in stressed syllables than in unstressed syllables (Mehta & Cutler, 1988; Wood & Terrell, 1998), supporting the hypothesis that linguistic prosody might trigger or at least enhance the perception of individual sounds (Chiat, 1983; Pierrehumbert, 2003). Taken together, awareness of prosodic patterns (e.g., English stress and Mandarin tone) may be important to reading acquisition because individuals might use such patterns as a segmentation cue to sound out words.

In studies with children, English stress perception explained unique variance in word reading independent of phonological awareness (Goswami, Gerson, & Astruc, 2010; Whalley & Hansen, 2006). English stress production also predicted significant and unique variance in word reading (Holliman, Wood, & Sheehy, 2008;

Jarmulowicz, Taran, & Hay, 2007) after controlling for phonological awareness. Additionally, native and nonnative English speakers are expected to share the same learning mechanism(s) for English reading ability. Chung and Jarmulowicz (2017) reported that adult Mandarin speakers could use English stress and suffixes as cues to sound out words. However, it is still unclear whether Mandarin-speaking children in non-English-dominant contexts are sensitive to acoustic cues associated with lexical stress and are using it to gain access to words.

Mandarin speakers tend to use fundamental frequency, a perceptually salient feature in their native tone language, as a cue for English stress production (Zhang, Nissen, & Francis, 2008) and perception (Ou, 2010; Yu & Andruski, 2010). Furthermore, Chung and Bidelman (2016) provided neurological evidence that adult Mandarin speakers' auditory brain activity is poorer at tracking intensity variations in English stress patterns relative to native English speakers. Mandarin speakers appear not to use intensity as a cue to perceive English stress as efficiently as native English speakers. It is possible that Mandarin-speaking children who have limited exposure to English may develop awareness of English stress patterns slowly. If this is true, Mandarin-speaking children learning English in a Mandarin-dominant context might have difficulties using English stress as an anchor to group phonemes into syllables and subsequent difficulty blending syllables into words.

While considerable attention has been paid to potential links between linguistic prosody awareness and reading in English, similar empirical evidence in tone languages is scarce. Of the many spoken dialects of Chinese, Mandarin and Cantonese are the most represented in the literature. Cantonese- and Mandarin-speaking children need to associate Chinese characters (i.e., printed words) with language-specific sound patterns for word reading. Cantonese tone perception was found to be associated with Chinese character recognition in Cantonese-speaking children at kindergarten (McBride-Chang et al., 2008) and elementary school (So & Siegel, 1997). In Mandarin, tone perception had significant association with Chinese character recognition in fourth-grade children with and without dyslexia (Wang, Huss, Hämäläinen, & Goswami, 2012). Zhang and McBride-Chang (2014) also revealed that Cantonese tone perception had an indirect effect on Chinese two-character word reading through phonological awareness. It appears that tone language speakers map individual phonemes and lexical tones to characters and then sound out Chinese characters in Mandarin and Cantonese. However, it remains unclear whether or not Mandarin tone perception independent of phonological awareness predicts Chinese character recognition.

Auditory processing and reading ability

Auditory processing is the perceptual ability to process general acoustic features (e.g., intensity, fundamental frequency, and duration). Those acoustic features are the limited set with which linguistic prosody (sound patterns beyond individual syllables) is realized. For example, English stress is a relative phenomenon in which a minimum of two syllables is required to determine which one is stressed. A stressed syllable may have higher fundamental frequency, higher intensity, and longer duration relative to an unstressed syllable (Fry, 1958; Kehoe, Stoel-Gammon,

& Buder, 1995; Morton & Jassem, 1965). In the case of Mandarin, lexical tone is physically represented by fundamental frequency height and contour (Howie, 1976). Thus, regardless of language, acoustic features must be processed as linguistically meaningful in order to acquire language normally. Auditory processing is the cornerstone of linguistic prosody and several attributes of this domain have been investigated in relation to language skills.

Amplitude envelope rise time is the rate of intensity change (i.e., loudness) at sound onset. The modulation of rise time signals speech rhythm or syllable stress—sharper or steeper rise times reflect stressed syllables while slower rise times reflect unstressed syllables. For example, the spoken word *seven* has sharper rise time for the stressed first vowel and slower rise time for the unstressed second vowel. Rise time discrimination is a significant predictor of L1 word reading in English (Corriveau, Pasquini, & Goswami, 2007; Goswami et al., 2010, 2011, 2013) and Mandarin (Wang et al., 2012). These findings suggest that rise time discrimination (i.e., detecting the percept of each individual beat within the rhythm), might trigger or underlie awareness of linguistic prosody and phonemes, which might further facilitate reading acquisition.

Additionally, pitch (the perceptual correlate of fundamental frequency) also plays a role in linguistic prosody. Melodies represented by tone sequences are often said to have two types of pitch structure: "contour" and "interval". Pitch contour refers to directional changes in the rises and falls of a sequence whereas pitch interval signals the distance between two adjacent tones (Dowling, 1982). Pitch contour is important to music and prosody (Patel, Peretz, Tramo, & Labreque, 1998), whereas pitch interval is specific to music. Foxton et al. (2003) found that pitch contour discrimination but not pitch interval discrimination was associated with English L1 word reading in adult monolingual readers. This suggests that contour processing, which requires monitoring more abstract and global auditory structure, might be a requisite of linguistic prosody and reading. Similar to pitch contour discrimination, pitch interval discrimination contributed to language ability (i.e., story retelling) in children speaking a tone language like Cantonese (Antoniou, To, & Wong, 2015). Hence, we aimed to explore whether pitch contour and interval are important for word reading in a tone language like Mandarin.

Pitch contour discrimination was found to be important to English L1 word reading (Foxton et al., 2003), although research has yet to determine whether pitch contour discrimination contributes to Mandarin L1 and English L2 word reading. Presumably, Mandarin speakers who are able to discriminate pitch contour patterns at the phrase level might also be good at the perception of language-specific pitch patterns at the syllable level (i.e., lexical tone), which in turn could help segment multisyllabic words into syllables or distinguish homophones differing in tone (e.g., $tang^1$ 'soup', $tang^2$ 'sugar', $tang^3$ 'lie down', $tang^4$ 'hot'). Once the linguistic connection between lexical tone and permissible phoneme combinations is formed, then it may become easier to sound out Chinese characters in Mandarin (each Chinese character corresponds to one syllable). Recently, we have shown that Mandarin-speakers pre-attentively detect subtle contour changes in pitch, suggesting that tone-language experience enhances the processing of global auditory pitch structure (Bidelman & Chung, 2015). Other neurophysiological studies indicate that

the neural representation of pitch is tuned by one's native prosodic system (Bidelman, Gandour, & Krishnan, 2011; Krishnan, Gandour, & Bidelman, 2010), suggesting that listeners are most sensitive at discriminating pitch variations specific to their native prosodic patterns. Hence, we hypothesized that pitch contour discrimination might be important to Mandarin L1 word reading but not English L2 word reading.

Although the English and Mandarin prosodic systems share the acoustic feature of fundamental frequency, it is less important for English stress perception relative to intensity and duration cues (Choi, Hasegawa-Johnson, & Cole, 2005; Greenberg, 1999; Kochanski, Grabe, Coleman, & Rosner, 2005). Hence, it remains unclear whether amplitude rise time, pitch contour, and pitch interval discrimination play different roles in Mandarin L1 and English L2 word reading. According to Antoniou et al. (2015) language-specific auditory cue hypothesis, only auditory cues specific to a language are important to learning that language. For Mandarin-speaking children, we hypothesized that pitch contour would be more important to Mandarin L1 word reading than is rise time, whereas the reverse would be true for English L2 word reading. This is because Mandarin uses pitch variations to signal different tonal patterns and syllable boundaries, and English uses rise time to signal boundaries of stressed and unstressed syllables.

Specific aims of the present study

Auditory processing and linguistic prosody awareness have been found to be important predictors of English L1 word reading in English monolingual children. The current study examined the relationships between auditory processing, linguistic prosody awareness, phonological awareness, and Mandarin L1 and English L2 word reading in Taiwanese children. The specific aims were as follows:

1. We aimed to determine the contributions of separate auditory processing abilities (pitch and amplitude rise time) to Mandarin L1 and English L2 word reading. We hypothesized that pitch contour discrimination would contribute to Mandarin L1 word reading (Foxton et al., 2003); whereas rise time discrimination would contribute to both Mandarin L1 (Wang et al., 2012) and English L2 word reading (Goswami et al., 2010).
2. Our next purpose was to compare the relative contributions of linguistic prosody awareness and phonological awareness to Mandarin L1 and English L2 word reading after controlling for age and nonverbal IQ. It was expected that linguistic prosody awareness would account for more variance in Mandarin L1 and English L2 word reading in comparison with phonological awareness (Goswami et al., 2010).
3. Our third goal was to compare the total contributions of linguistic prosody awareness and phonological awareness to Mandarin L1 and English L2 word reading after controlling for age and nonverbal IQ. Because Chinese is a logographic language and English is an alphabetic language, it was expected that linguistic prosody awareness and phonological awareness would account for more variance in English L2 word reading than in Mandarin L1 word reading.

Methods

Participants

Sixty-three fourth graders in Taipei, Taiwan participated in this study. Normal hearing (<25 dB HL) was confirmed in both ears for all children via an audiometric hearing screening conducted at octave frequencies between 1 and 4 kHz. Two of the children failed to pass the hearing screening and were excluded from the study. Sixty-one children remained in the current study (29 boys, 32 girls; age: $M = 9.82$ years, $SD = 0.25$). The children had no speech, language, emotional, or physical problems reported by classroom teachers.

The children were native Mandarin speakers and seldom had any opportunity to speak English in daily conversation except the classroom setting. Moreover, they rarely spoke to native English speakers in their home environment. In Taipei, the compulsory education begins formal literacy instruction in Mandarin (L1) and English (L2) from first grade, at the age of six. The instruction medium is Mandarin. However, the children's mean onset age of English learning was around four ($M = 4.87$ years, $SD = 1.13$) because some children began learning English through tutoring programs.

Materials

Nonverbal intelligence

Raven's Standard Progressive Matrices (RSPM; Chen & Chen, 2006) were used to assess children's nonverbal intelligence. The RSPM consists of 60 black-and-white test items. In each test item, children were required to select from six to eight choices the missing element that completed a pattern. The RSPM scores were obtained from the schools, as all children were given the RSPM test and the subsequent testing materials in the same semester.

Auditory processing

The auditory processing tasks were presented using custom routines coded in MATLAB via a graphical user interface (GUI). The testing Mac laptop output was calibrated (70 dB SPL) and stimuli were presented binaurally through Sennheiser HD 280 headphones. Each auditory processing task included five practice trials and 40 experimental trials. During practice and experimental trials, visual feedback was presented by the MATLAB GUI signaling the correctness of each trial. Each child received extra verbal explanation and reinforcement in the five practice trials.

One rise time task and two pitch tasks were used to tap children's auditory processing abilities along multiple perceptual dimensions. In the rise time task, each trial had three tones varying in rise time (rate of intensity change at tone onset) presented in a three interval forced choice task (3IFC). The parameters of rise time stimuli were based on those of Goswami et al. (2013). Two of the intervals

contained standard tones with a 300 ms rise time; the third contained a comparison which had a shorter rise time (e.g., 150 ms). The duration of rise time was adaptively varied according the child's response in a 2-down and 1-up procedure, tracking 71% correct performance (Levitt, 1971). That is, the duration of rise time decreased (i.e., made more difficult) following two consecutive correct responses and increased (made easier) following each incorrect response. The task was to decide which interval sounded different (i.e., "odd-one-out"). Using this procedure, differential thresholds were measured as the smallest difference in rise time that children could reliably detect. Smaller discrimination thresholds represent a higher sensitivity to intensity changes of tone onset.

Pitch contour and interval discrimination were measured using tasks initially developed by Foxton et al. (2003). Both pitch tasks consisted of 40 pairs of six-tone sequences. Half of the pairs contained identical tone sequences; the other half contained standard tone sequences and deviations in which a random tone mid-sequence was altered (see asterisks, Fig. 1). Pitch interval discrimination required children to discriminate the standard tone sequence from one that maintained the contour structure of melody but changed the precise pitch distance between adjacent tones. In contrast, pitch contour discrimination asked children to discriminate the standard tone sequence from a deviant which violated the contour pattern of pitch rises and falls (e.g., the random tone went down instead of up). The pitch contour and pitch interval tasks were presented in a same-different (2IFC) paradigm. The task was to decide whether the pairs of six-tone sequences were the same or different. Responses were quantified via d' [i.e., $d' = z(H)–z(FA)$, where H and FA are the hit and false alarm rates, respectively]. A higher d' signals better discrimination of interval/contour information.

Linguistic prosody awareness measures

Four linguistic prosody awareness tasks were used to measure how children process English stress and Mandarin tone across two modalities (perception, production) and three types (monosyllable, disyllable, phrase). The stimuli in English and Mandarin linguistic prosody tasks were produced by English and Mandarin native speakers, respectively.
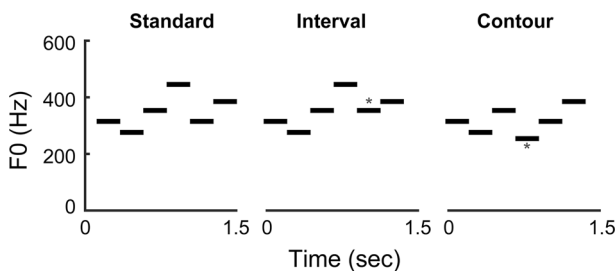

Fig. 1 Schematic spectrogram of pitch interval and contour stimuli. Stimuli are shown for a standard six-tone sequence and deviant conditions, which altered the interval and contour structure of the repeating pitch pattern by altering one random tone in the mid-sequence marked by an *asterisk*

Three "DEEdee" tasks were used to assess children's English stress perception and production, and Mandarin disyllabic tone perception. The DEEdee task was first used in Kitzen's (2001) dissertation. In the DEEdee task, the phonemic information of each syllable was eliminated and replaced by the syllable 'dee', but the stress or tone patterns were retained in each word. The DEEdee task has been used to measure English stress perception in English monolingual children (Goswami et al., 2010; Whalley & Hansen, 2006). In the current study, the DEEdee task was adapted to tap English stress production and Mandarin tone perception. The receptive English DEEdee task included four practice trials and 15 experimental trials. Each child heard a digitally recorded target English phrase (e.g., *Humpty Dumpty*) and then chose from two choices the DEEdee phrase that matched in stress (e.g., DEEdee DEEdee corresponds with HUMPty DUMPty). Its Cronbach's alpha was .537.

An expressive English DEEdee task measured Mandarin-speaking children's English stress production in disyllabic words. In this task, there were four practice trials and 12 experimental trials for disyllabic words. These disyllabic targets were low-frequency words selected from Arciuli and Cupples' (2006) study so that children could not use their lexical knowledge as a cue to determine which syllable in a disyllabic word was stressed. Children were auditorily presented a word and then asked to produce its stress pattern with each syllable replaced by the syllable 'dee' (DEEdee for PENcil and deeDEE for diVIDE). This task was intended to isolate Mandarin-speaking children's stress production ability from phonemic awareness. Its Cronbach's alpha was .718.

A receptive Mandarin DEEDEE task was created for this study and consisted of four practice trials and 15 experimental trials. Just as in English, each child heard pre-recorded DEEDEE sequences, however for Mandarin the tone remained (e.g., DEE^4DEE^1—the superscript numbers indicate tone patterns). The child heard a target Mandarin word (e.g., qi^4che^1 "car") and then selected from two choices the DEEDEE phrase with the same tone pattern as the target Mandarin word (e.g., DEE^4DEE^1 for qi^4che^1 "car"). The two choice DEEDEE phrases consisted of tone patterns that matched the target word and a distractor tone sequence (e.g., DEE^4DEE^1 for a target word qi^4che^1 "car" and DEE^3DEE^1 for a distractor word lao^3shi^1 "teacher"). The targets and distractors were both high frequency disyllabic words commonly used by elementary school children (Ministry of Education, 2002). The target words ($M = 0.045$, $SD = 0.012$) and the distractor words ($M = 0.045$, $SD = 0.016$) were not significantly different from each other in word frequency [$t(38) = 0.95$, $p > .05$].

Several constraints were also imposed on the composition of the stimuli. First, in this disyllabic tone perception task, it was never the case that the first syllable contained a third tone, because the third tone preceding other tones results in tone sandhi or tone change (Li & Thompson, 1981). Additionally, for the target and distractor words, the first and second syllables of the disyllabic words did not share the same tone. This was to control for the possibility that it might be easier to identify words with two same tones than those with different tones in each syllable (e.g., lao^3shi^1 "teacher" vs. $dong^1xi^1$ "thing"). Its Cronbach's alpha was .651.

As a second tone perception task, Liu and Hu's (2010) tone matching task was used to assess monosyllabic tone perception. In this task, there were two practice trials and 20 experimental trials. Children were auditorily presented three monosyllables and then required to select from the second (e.g., gao^4) or the third syllable (e.g., gan^3) the one that has the same tone as the first syllable (e.g., gei^4). The monosyllables were permissible sound combinations in Mandarin, and included both low-frequency words and nonwords. Compared with the Mandarin DEEDEE perception task, the monosyllabic tone perception task retained phonetic information in each syllable. Its Cronbach's alpha was .713

Phonological awareness (PA)

Sound oddity tests for rhyme and final phoneme contrasts were used to assess children's PA in Mandarin (Chan, Hu, & Wan, 2005; Hu & Catts, 1998) and in English (Bowey, Cain, & Ryan, 1992). The stimuli in the Mandarin and English PA tasks were produced by Mandarin and English native speakers, respectively. Each of the Mandarin PA tests consisted of two practice trials and 10 experimental trials and each of English PA tests consisted of three practice trials and 12 experimental trials. In each trial, three pre-recorded words were presented twice through speakers to children in a class group. The presentation of Mandarin and English PA stimuli followed Hu and Schuele's (2005) procedure. The experimenter pointed to the numbers 1, 2, and 3 on the black board which corresponded with the three words they would hear. The child's job was to choose the relative order of the spoken word that sounded different from the others (e.g., for onsets, which word has a different first sound *sing, bus,* or *sun*) by circling the number on an answer sheet that represented the odd spoken word. Their Cronbach alphas were .763 (Mandarin rhyme), .601(Mandarin final phoneme), .750 (English rhyme), and .745 (English final phoneme).

Reading measures

The *Graded Chinese Character Recognition Test* (Huang, 2004) was used to assess Chinese character recognition ability. Children sounded out each Chinese character in Mandarin until they made 20 consecutive errors. The task is a standardized test, which has been adopted in several published studies (Chung & Hu, 2007; Goswami et al., 2011). It has an internal consistency of 0.99 and test–retest reliability ranging from 0.81 to 0.95. No pseudo-character recognition task was used because Chinese readers cannot decode pseudo-characters into sounds without memorization of character-sound patterns.

For English, the sight word efficiency and the phonemic decoding efficiency subtests of the *Test of Word Reading Efficiency-II* (*TOWRE-II*; Torgesen, Wagner, & Rashotte, 2012) were used to assess real word reading and nonword decoding skills. Each child read aloud a list of English real words within 45 s and decoded a list of English nonwords within 45 s. The two English reading tasks were treated as separate dependent variables for three reasons: (1) real word reading typically requires stored sound patterns for printed words and less application of grapheme-

phoneme correspondence knowledge than does nonword decoding, thus, poor performance in nonword decoding could be due to emerging and incomplete knowledge of grapheme and phoneme correspondence, but real word reading might be a relative strength; (2) stress perception and production made more contributions to real word reading than to nonword decoding by adults (Chung & Jarmulowicz, 2017); and (3) this approach the current study adopted was found in previous studies (Goswami et al., 2010; Whalley & Hansen, 2006).

Procedures

The study was approved by the Institutional Review Board at the University of Memphis. After obtaining informed consent forms and questionnaires from children's parents, the first author met each child individually and in a class group setting through four sessions, lasting 40–50 min. All measures were given to children after assent had been obtained from children. Each child was randomly assigned a number when he or she agreed to participate in the study. Children with odd numbers and those with even numbers were given tasks in two different sequences in order to diminish the effect of inattention on the performance of tasks given at the end of each session. Children's oral responses were recorded with a SONY ICD-UX543F digital voice recorder and on the answer sheets scored by the first author.

Reliability

The first author administered all tasks to the 61 children. All tasks except the stress production task were either forced-choice or discrimination (same/different) and were scored at the time of testing. The English stress production task was recorded, transcribed, and scored by the first author, and a subset of the recordings were transcribed and scored by a second trained transcriber. Hence, inter-rater reliability was provided to ensure reliable judgments for English stress production. Inter-rater scoring reliability was examined using a two-way mixed, absolute agreement, single-measures intra-class correlation (ICC) (Hallgren, 2012). The degree that two coders provided absolute values in their ratings of English stress production was assessed in about 25% of participants (15 children). The resulting ICC was 0.968 and appeared in the excellent range between 0.75 and 1.0 (Cicchetti, 1994), indicating that English stress production was rated similarly across coders.

Results

Raw scores are reported for children's performance on all tasks except the auditory processing measures. The maximum scores, means, and standard deviations for all of the measures are shown in Table 1. Performance on the two pitch tasks is reported as d'. The d' scores were calculated by taking into consideration children's correct (hits) and incorrect (false alarm) responses. Performance on the rise time

Table 1 Descriptive statistics for all measures (N = 61)

Measures	Maximum	Mean	SD
Age (in years)	–	9.82	.25
Nonverbal IQ	60	43.57	6.26
Auditory processing			
Pitch contour discrimination (*d'*)	–	1.92	1.06
Pitch interval discrimination (*d'*)	–	1.11	.78
Rise time discrimination (ms)	–	122.29	56.56
Mandarin measures			
Disyllabic tone perception	15	13.49	1.64
Monosyllabic tone perception	20	15.39	3.00
Rhyme awareness	10	8.76	1.90
Final phoneme awareness	10	5.80	2.20
Chinese character recognition	200	98.49	21.33
English measures			
Stress perception	15	9.95	2.50
Stress production	12	8.34	2.77
Rhyme awareness	12	9.80	2.32
Final phoneme awareness	12	8.66	2.63
English real word reading	108	32.87	17.25
English nonword decoding	66	13.02	8.73

task is reported as a threshold, i.e., the smallest difference in amplitude onset (in ms) that children were able to detect.

Pearson's correlations among age, nonverbal IQ, auditory processing, Mandarin measures, and English measures are shown in Table 2. Several points are worth noting from this correlational matrix. The two pitch discrimination tasks were strongly correlated to each other ($r = 0.69$), but neither was correlated with the rise time task, suggesting an independence between amplitude and pitch processing measures. The pitch discrimination tasks were associated with Mandarin final phoneme awareness ($r = 0.26$ for pitch interval) and Chinese character recognition ($r = 0.38$ for pitch contour), but not with any of the English measures. Rise time discrimination was significantly correlated with final phoneme awareness (Mandarin: $r = -0.25$; English: $r = -0.32$), real word reading (Mandarin: $r = -0.29$; English: $r = -0.41$), rhyme awareness in English ($r = -0.26$; but not in Mandarin), and English nonword reading ($r = -0.26$).

Contributions of auditory processing to word reading

Several regression models were used to examine the contributions of separate auditory processing abilities to different word reading abilities (i.e., Chinese character recognition, English word reading and nonword reading) after controlling for subject factors (i.e., age and nonverbal IQ). For each 3-step fixed entry hierarchical regression model, age was entered at Step 1 followed by nonverbal IQ

Table 2 Correlations between auditory processing, linguistic prosody awareness, phonological awareness, and word reading per language (N = 61)

Age, IQ, & auditory processing	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Age	–															
2. Nonverbal IQ	**.461**	–														
3. Pitch contour discrimination	.027	**.336**	–													
4. Pitch interval discrimination	.161	.237	**.694**	–												
5. Rise time discrimination	.008	**−.335**	−.192	−.109	–											
6. Disyllabic tone perception	.212	.227	.131	.108	−.082	–										
7. Monosyllabic tone perception	.188	**.324**	.060	.138	−.208	**.350**	–									
8. Mandarin rhyme awareness	**.311**	**.500**	.241	.168	−.041	.201	.171	–								
9. Mandarin final phoneme	−.011	.172	.218	**.261**	**−.253**	.178	**.369**	.123	–							
10. Chinese character recognition	−.037	**.263**	**.376**	.218	**−.294**	**.362**	.213	.212	**.266**	–						
11. Stress perception	.187	**.263**	−.062	−.050	−.016	.236	**.414**	.150	.234	.076	–					
12. Stress production	−.038	.125	.116	−.074	−.196	**.308**	**.341**	−.022	.101	**.269**	**.453**	–				
13. English rhyme awareness	.019	.158	.046	.197	**−.268**	**.330**	.243	.008	**.262**	**.461**	.119	**.269**	–			
14. English final phoneme	−.029	.219	.082	.175	**−.323**	**.289**	−.052	−.073	.020	**.381**	.033	.176	**.689**	–		
15. English real word	.124	**.310**	.047	.155	**−.409**	**.425**	**.395**	**.307**	.250	**.330**	.233	**.332**	**.535**	**.437**	–	
16. English nonword	.110	**.273**	.015	.157	**−.265**	**.370**	**.374**	**.345**	.279	.234	.190	**.284**	**.510**	**.380**	**.839**	–

Significant values ($p < .05$) are marked in boldface. Rise time discrimination has negative associations with other variables because its smaller threshold score signals higher sensitivity to rise time

Table 3 Hierarchical regressions showing the variance in Mandarin L1 and English L2 word reading accounted for by separate auditory processing abilities after controlling for age and nonverbal IQ

Step	Chinese character		English real word		English nonword	
	Final β	R^2 change	Final β	R^2 change	Final β	R^2 change
1. Age	–	.001	–	.015	–	.012
2. Nonverbal IQ	–	.099*	–	.081*	–	.062*
3. Auditory processing						
Pitch contour	.302*	.079*	−.069	.004	−.091	.007
Pitch interval	.177	.029	.088	.007	.100	.009
Rise time	−.203	.035	−.353**	.107**	−.199	.034

The final beta values of age and nonverbal IQ were different when each auditory processing skill was entered at Step 3. Hence, their final beta values were omitted in this table

** $p < .01$; * $p \leq .05$

at Step 2. Each auditory processing ability was entered separately at Step 3 in three successive models. No regression model was found to predict English nonword reading. As shown in Table 3, pitch contour discrimination accounted for an additional 7.9% of the variance in Chinese character recognition after entering age and nonverbal IQ. Rise time discrimination explained 10.7% of the variance in English real word reading after entering age and nonverbal IQ. However, none of the three auditory processing abilities predicted English nonword decoding.

Relative contributions of prosody and PA to word reading

In several studies with English monolingual children, linguistic prosody awareness independent of phonological awareness predicted English L1 reading abilities (Goswami et al., 2010; Holliman et al., 2008; Whalley & Hansen, 2006). Hence, the unique contributions of linguistic prosody awareness and phonological awareness to Mandarin L1 and English L2 word reading were examined using two, 4-step fixed-entry hierarchical regression equations. In both models, age was entered at Step 1 and nonverbal IQ at Step 2. Steps 3 and 4 were either prosodic perception/production or phonological awareness. Because the current study aimed to examine whether phonological awareness was more important to English L2 word reading than were prosodic perception and production, the entry steps of prosodic perception/production and phonological awareness were reversed in the two models (for similar analysis approach, see Goswami et al., 2010; Whalley & Hansen, 2006).

In our hierarchical regression analyses, several criteria guided the selection of independent variables. First, disyllabic tone perception (i.e., Mandarin DEEdee task), but not monosyllabic tone perception, was entered as an independent variable based on the correlation results (See Table 2). Second, only rhyme awareness was used as an observed variable for the construct "phonological awareness" because phonological awareness at the rhyme level has been extensively examined with prosodic awareness in previous studies on English L1 and Mandarin L1 reading acquisition (Goswami et al., 2010; Wang et al., 2012). Third, English stress

perception and production were used as different observed variables for the construct "prosodic awareness" because (1) English stress perception and Mandarin disyllabic tone perception were designed based on the same DEEdee task (Kitzen, 2001), and (2) stress perception and production differ in activation of stress knowledge (production: more active; perception: more passive) and retention of sound sequences in working memory (production: more; perception: less) (Chung & Jarmulowicz, 2017).

The Mandarin data are presented in Table 4. Given that Mandarin final phoneme awareness and Chinese character recognition were significantly correlated, phonological awareness at the final phoneme level was also entered in hierarchical regression analyses (See Table 4). Overall, the four variables (age, nonverbal IQ, disyllabic tone perception, and Mandarin rhyme/final phoneme awareness) explained about 22–24% of variance in Mandarin word reading. Mandarin disyllabic tone perception made a significant contribution to Chinese character recognition after controlling for age, nonverbal IQ, and Mandarin rhyme/final phoneme awareness. In fact, disyllabic tone perception was the most important predictor in the model, uniquely explaining about 11% of Chinese character reading. In contrast, Mandarin rhyme/final phoneme awareness was not a significant addition to the model, irrespective of entry steps.

In Table 5, the English data are illustrated. In contrast to the Chinese results, English stress perception did not alter the models of either English real word reading or nonword decoding, irrespective of entry steps. However, English rhyme awareness significantly explained additional variance, ~ 24 and 22%, in English real word reading and nonword decoding, respectively (after accounting for age and nonverbal IQ). For English, rhyme awareness was the most important predictor for word reading ($\beta = .49$) and non-word reading ($\beta = .47$).

Table 4 Hierarchical regressions showing the variance in Mandarin L1 word reading accounted for by Mandarin tone perception compared to phonological awareness after controlling for age and nonverbal IQ

Model	Step	Chinese character		Chinese character	
		Final β	R^2 change	Final β	R^2 change
1 & 2	1. Age	−.257	.001	−.226	.001
	2. Nonverbal IQ	.256	.099*	.266	.099*
1	3. Disyllabic tone perception	.340**	.113**	.321*	.113**
	4. Mandarin phonological awareness	.096	.007	.161	.024
2	3. Mandarin phonological awareness		.013		.043
	4. Disyllabic tone perception		.107**		.094*
	Total R^2		.220**		.238**

Mandarin phonological awareness refers to rhyme awareness when the dependent variable is Chinese character in the left column and to final phoneme awareness when the dependent variable is Chinese character in the right column

*** $p < .001$; ** $p < .01$; * $p < .05$

Table 5 Hierarchical regressions showing the variance in English L2 word reading accounted for by English stress perception compared to phonological awareness after controlling for age and nonverbal IQ

Model	Step	English real word		English nonword	
		Final β	R^2 change	Final β	R^2 change
1 & 2	1. Age	.000	.015	.006	.012
	2. Nonverbal IQ	.201	.081*	.172	.062*
1	3. English stress perception	.122	.025	.088	.016
	4. English rhyme awareness	.489***	.230***	.472***	.215***
2	3. English rhyme awareness		.242***		.223***
	4. English stress perception		.014		.007
	Total R^2		.352***		.305***

*** $p < .001$; **$p < 0.1$; * $p \leq .05$

Table 6 shows the contributions of English stress production and rhyme awareness to English reading. These results are notably different from the stress perception data. When English stress production was entered before rhyme awareness, it explained a small, but significant amount of variance after age and nonverbal IQ in both English real word reading (8.7%) and nonword decoding (6.3%). However, when English rhyme awareness was entered before stress production, it eliminated any independent contribution of stress production. Indeed, the only important predictor for both English real word and nonword reading was English rhyme awareness (β = .45 and β = .44 for real and nonword reading respectively).

Total contributions of prosody and PA to Mandarin L1 and English L2 word reading

In Table 4, about 12–14% of the unique variance in Mandarin L1 word reading could be explained by linguistic prosody awareness (i.e., disyllabic tone perception)

Table 6 Hierarchical regressions showing the unique variance in English L2 word reading accounted for by English stress production compared to phonological awareness after controlling for age and nonverbal IQ

Model	Step	English real word		English nonword	
		Final β	R^2 change	Final β	R^2 change
1 & 2	1. Age	.030	.015	.028	.012
	2. Nonverbal IQ	.201	.081*	.171	.062*
1	3. English stress production	.186	.087*	.144	.063*
	4. English rhyme awareness	.452***	.186***	.443***	.179***
2	3. English rhyme awareness		.242***		.223***
	4. English stress production		.032		.019
	Total R^2		.370***		.317***

*** $p < .001$; ** $p < .01$; * $p \leq .05$

and phonological awareness after controlling for age and nonverbal IQ. Turning to English L2 word reading, in Tables 5 and 6, about 26–27% of the variance in English real word reading and about 23–24% of the variance in English nonword decoding was explained by linguistic prosody awareness (English stress perception/production) and phonological awareness after controlling for age and nonverbal IQ.

Discussion

While there is a strong assumption that auditory processing, linguistic prosody awareness, phonological awareness, and reading are related (Zhang & McBride-Chang, 2010), little empirical evidence has established relationships between these variables. In studies with English monolingual children, auditory processing and linguistic prosody awareness were found to make substantial contributions to English L1 word reading (Goswami et al., 2010; Whalley & Hansen, 2006). However, to date, relatively little research has been conducted on the contributions of auditory processing and linguistic prosody awareness to word reading in a tone language with logographic writing system (e.g., Mandarin has four tones and printed words in Chinese characters) and in English as a second language. In the current study, we examined the links among auditory processing, language-specific linguistic prosody awareness, phonological awareness, and Mandarin L1 and English L2 word reading in Taiwanese children. In the sections that follow, we revisit the three specific aims. For each aim, Mandarin measures are discussed first and English measures second. Finally, limitations and future directions are presented.

Auditory cues specific to Mandarin L1 and English L2 word reading

Pitch contour discrimination predicted Mandarin L1 word reading, and rise time discrimination predicted English L2 real word reading (after controlling age and nonverbal IQ). These findings support the hypothesis that some auditory cues might be more specific to one language relative to others (Antoniou et al., 2015). In this case, detecting pitch fluctuations (at the contour but not interval level) plays a more important role in Mandarin L1 word reading relative to detecting the percept of each individual beat (i.e., rise time); whereas the reverse is true in English L2 word reading. Pitch contour discrimination, however, did not predict unique variance in English L2 word reading, supporting the proposition that listeners are more adept at processing pitch variations specific to their native prosodic system (Bidelman et al., 2011; Krishnan et al., 2010). Collectively, Mandarin-speaking children may rely on pitch contour processing for Chinese character recognition, but amplitude envelope onset for English L2 real word reading.

In accordance with Goswami et al.'s (2010, 2011) studies with English monolingual children, rise time discrimination made substantial contributions to English L2 word reading. These results suggest that rise time discrimination is important to English word reading regardless of L1 or L2. Contrary to Wang et al. (2012), our study did not find that rise time discrimination was a unique predictor of

Mandarin L1 word reading in typically developing children. The inconsistency between the two studies might be due to differences between the current study and Wang, Huss, Hämäläinen, and Goswami's study. First, substantial individual differences were found in Wang et al. (2012) study with Mandarin-speaking children with and without dyslexia. Wang et al. (2012) reported that Mandarin-speaking children with dyslexia had difficulty on a rise time task when compared to a control group. It may be that sensitivity to rise time is attenuated as a result of dyslexia, regardless of L1 (Goswami et al., 2011). Hence, the relationship between rise time discrimination and Mandarin L1 word reading might be strengthened by substantial individual differences in a group of children with and without dyslexia.

Second, children's rise time discrimination thresholds were calculated in different ways across different studies. In light of thresholds of sinusoidal amplitude modulation estimated in children with and without dyslexia (Lorenzi, Dumont, & Füllgrabe, 2000), children's *thresholds*, in the current study, were computed by adaptively measuring the smallest difference in rise time stimuli that children could reliably discriminate. In contrast, previous studies have calculated the number of trials children needed to reliably discriminate the smallest difference in rise time stimuli (larger numbers represent less sensitivity) which is insufficient to measure thresholds and provides a much coarser estimate of rise-time sensitivity (Goswami et al., 2010, 2013; Wang et al., 2012; Zhang & McBride-Chang, 2014). Hence, the relationship between rise time discrimination and Mandarin L1 word reading might be diminished by the smaller individual differences in this measure when assessed at millisecond resolution (present study).

Relative contributions of prosody and PA to word reading differed in L1 and L2

Direct comparisons between languages allowed us to assess the relative contributions of linguistic prosody awareness and phonological awareness to Mandarin L1 and English L2 word reading. Linguistic prosody awareness made a more substantial contribution to Mandarin L1 word reading than did phonological awareness; whereas the reverse was observed in English L2. The different contributions of linguistic prosody awareness and phonological awareness to Mandarin and English are discussed in terms of different phonological structures between languages, and interference between L1 and L2.

Mandarin disyllabic tone perception, but not rhyme awareness, predicted the unique variance in Mandarin L1 word reading after controlling for age and nonverbal IQ. These results may be explained by the phonological structure in Mandarin and its sound-character mapping. That is, children need to use lexical tones to distinguish partial homophones (e.g., ma^1 'mother', ma^2 'numbness, ma^3 'horse', and ma^4 'scold'), map these sounds to their corresponding Chinese characters, and then sound out Chinese characters. This may explain why linguistic prosody awareness plays a more important role in Mandarin L1 word reading than does phonological awareness. This finding is mirrored in work with English monolingual children in which English stress perception and production made substantial contributions to English L1 word reading, independent of phonological

awareness (Goswami et al., 2010; Holliman et al., 2008; Jarmulowicz et al., 2007; Whalley & Hansen, 2006). Taken together, linguistic prosody awareness seems to be a more important predictor of first language word reading, Mandarin (L1) in the current study and English (L1) in previous work, than phonological awareness.

Interestingly, Mandarin rhyme awareness did not explain significant variance in Mandarin L1 word reading above and beyond age and nonverbal IQ. One possible explanation is that explicit instruction in Mandarin phonetic symbols helps children master the Mandarin phonetic inventory in a short time, and consequently diminishes the contribution of phonological awareness to Chinese character recognition. Mandarin has a simple syllable structure, a small number of syllable types with no consonant clusters, and only about 1300 syllables with tonal distinctions used by Mandarin speakers (Duanmu, 2000). In first grade, Taiwanese children receive 10 week intensive instruction in Mandarin phonetic symbols for onsets (i.e., initial consonants in a syllable), rhymes (i.e., syllable vowels, diphthongs or vowels followed consonants), and four tones, which in turn may attenuate individual variability in phonological awareness.

Contrary to Mandarin L1 word reading, English rhyme awareness made a more substantial contribution to English real word reading and nonword decoding than did English stress perception and production. Our findings do not support the results of previous studies with English monolingual children that English stress perception/production, independent of phonological awareness, predicted significant additional variance in L1 word reading (Goswami et al., 2010; Holliman et al., 2008; Jarmulowicz et al., 2007; Whalley & Hansen, 2006). The inconsistency in performance between native English speakers and Mandarin-speaking children might be the result of cross-language differences and interference. First, the English syllable is more complex, with a wider range of permissible syllable types and more complex coda sequences. This characteristic of English (but not Mandarin) may be particularly salient for Mandarin-speaking children, especially in an alphabetic language in which sounds are represented by letters. Second, Mandarin L1 speakers may not be as familiar with the English stress system as native English speakers. This may be due either to limited exposure, or practice, or both.

It is worth noting that English L2 real word reading was predicted by English stress production, but not by English stress perception. It may be that the perception task was not sensitive enough with these school-aged children; however, the descriptive data show that accuracy on these two tasks was similar (66 and 70% for perception and production respectively). Another possibility is that the process of word production is what is important. As we have investigated in work with adults (Chung & Jarmulowicz, 2017), syllable assembly may be important, particularly in relation to reading. Furthermore, because English stress production task requires readers to assemble syllables and stress them accurately, it may be a more important predictor of English L2 word reading in Mandarin speakers in comparison with English stress perception task.

With regard to English nonword decoding, English stress perception predicted no significant additional variance after controlling for age and nonverbal IQ. The results corroborate previous studies that English stress perception was not a significant predictor of nonword decoding in either English monolingual children

(Whalley & Hansen, 2006) or adult Mandarin speakers (Chung & Jarmulowicz, 2017). However, the findings contradict Goswami et al. (2010) study with English-speaking children with and without dyslexia. The inconsistency between studies might be attributable to substantial individual differences in English stress perception and nonword decoding in Goswami et al. (2010) study population with dyslexia. Nevertheless, the contribution of English stress perception to nonword decoding in the current study was negligible.

Total contributions of prosody and PA to Mandarin L1 and English L2 word reading

We also found that language-specific linguistic prosody awareness explained significant variance in Mandarin L1 and English L2 word reading. This suggests that linguistic prosody awareness is an important contribution to reading development across languages and orthographies. Additionally, linguistic prosody awareness and phonological awareness together accounted for more variance in English L2 word reading than in Mandarin L1 word reading. These findings suggest that English has a closer letter-sound relationship than does Mandarin.

It is noteworthy that linguistic prosody awareness and phonological awareness made more substantial contributions to English L1 word reading (e.g., 39–40%) in Goswami et al. (2010) study relative to Mandarin L1 (12–14%) and English L2 (e.g., 23–27%) word reading in the current study. There are three possible explanations for these results. First, as mentioned above, substantial individual differences may have been observed in children of the Goswami et al. (2010) given the inclusion of dyslexic children, an effect that may not be as evident in typically developing children. Hence, the addition of an atypical population was likely a source of variability in Goswami's study. Second, segmental and prosodic information may not be as important to a logographic language like Mandarin, because its printed words (i.e., Chinese characters) cannot be decoded into sounds directly (Hu & Catts, 1998; Siok & Fletcher, 2001). Hence, in the current study, the fourth-grade children who mastered Mandarin syllable structure, perhaps through intensive instruction in Mandarin phonetic symbols in first grade, might rely on morphological knowledge (Hu, 2013) and orthographic skills (Li, Shu, McBride-Chang, Liu, & Peng, 2012) to map sounds to characters for Mandarin L1 word reading. Third, the fourth graders in the current study had only 3–5 years of experience learning to read English in a Mandarin-dominant context. They may still be developing the abilities to use English stress and phonemes to sound out English L2 real words and nonwords.

Future directions

The current study implemented a correlational design to examine links among auditory processing, linguistic prosody awareness, phonological awareness, and word reading. Mandarin-speaking children's awareness of prosodic patterns was measured by non-speech (i.e., pure tones varying in rise time and pitch) and speech

stimuli. As native Mandarin speakers, the children recruit pitch contour as an auditory cue to distinguish Mandarin tone patterns, and awareness of Mandarin tone patterns is more important for Chinese character recognition (e.g., homophones differing in tones) than is phonological awareness. As young beginning English learners, the children employ rise time to distinguish syllable boundaries that carry relative stress information in English, and awareness of stressed and unstressed syllables is less important for English word and nonword reading than is phonological awareness.

Nevertheless, the development of the three abilities (i.e., auditory processing, linguistic prosody awareness, and phonological awareness) is dynamic and interrelated from birth to school age (Zhang & McBride-Chang, 2010). Hence, longitudinal studies should be conducted to examine the relative contributions of the three abilities at preschool age to reading abilities at school age. Indeed, the relative contributions of linguistic prosody awareness and phonological awareness to Mandarin L1 and English L2 word reading might be more clearly delineated when both abilities are measured before children receive formal instruction in Mandarin phonetic symbols.

Additionally, a few questions remain regarding English real and nonword performance. It is unclear whether the English learners in the current study actually recognized the real words they read, or if they used the same decoding approach to sound out both real and nonwords, or if they used both a decoding and sight word strategy. This may also account for differences between previous findings with monolingual children and our findings with Mandarin-speaking English learners. Thus, future work might explore the contributions of phonological and prosodic awareness to English reading with a broader cross-section of children with more experience with English reading or with a longitudinal design that following children whose proficiency improved over time.

In light of previous studies suggesting continuous fluctuations in intensity are critical for accurate perception of English stress patterns (Goswami & Leong, 2013; Chung & Bidelman, 2016), future studies are needed to clarify the degree to which Mandarin speakers rely on intensity as a prosodic cue to processing English stress. Recently, a neurophysiological study revealed that adult Mandarin speakers showed poorer auditory cortical tracking of intensity variations signaling English stress patterns relative to adult native English speakers (Chung & Bidelman, 2016). Hence, future research might include Goswami et al. (2010) DEEdee task in which pitch variations in DEEdee sequences are eliminated. In this scenario, Mandarin speakers would need to exploit intensity variations as a cue to discriminate English stress patterns.

A possible limitation in the current study might have been the English reading measures. Because the two *TOWRE-II* reading subtests were timed tasks, the variability in English L2 real word reading and nonword decoding might have been reduced. However, phonological awareness and linguistic prosody awareness still predicted more variance in English L2 real word reading and nonword decoding than that in Mandarin L1 word reading (an untimed task). Hence, timed or untimed reading tasks did not contaminate comparisons between Mandarin L1 and English L2 word reading.

In summary, results of the current study indicate that pitch contour discrimination contributes to Mandarin L1 word reading and rise time discrimination predicts English L2 word reading. Moreover, linguistic prosody awareness is more important to Mandarin L1 word reading than phonological awareness, whereas the reverse is observed in English L2 word reading. Therefore, the current study highlights the roles of auditory processing, linguistic prosody awareness, and phonological awareness in word reading and how those roles differ in Mandarin as an L1 and English as an L2.

References

Antoniou, M., To, C. K. S., & Wong, P. C. M. (2015). Auditory cues that drive language development are language specific: Evidence from Cantonese. *Applied Psycholinguistics, 36,* 1493–1507. doi:10.1017/S0142716414000514.

Arciuli, J., & Cupples, L. (2006). The processing of lexical stress during visual word recognition: Typicality effects and orthographic correlates. *Quarterly Journal of Experimental Psychology, 59,* 920–948. doi:10.1080/02724980443000782.

Bidelman, G. M., & Chung, W.-L. (2015). Tone-language speakers show hemispheric specialization and differential cortical processing of contour and interval cues for pitch. *Neuroscience, 305,* 384–392. doi:10.1016/j.neuroscience.2015.08.010.

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience, 23,* 425–434. doi:10.1162/jocn.2009.21362.

Bowey, J. A., Cain, M. T., & Ryan, S. M. (1992). A reading-level design study of phonological skills underlying fourth-grade children's word reading difficulties. *Child Development, 63,* 999–1011.

Chan, I.-C., Hu, C.-F., & Wan, I.-P. (2005). Learning to read and spell: The relative role of phonemic awareness and onset-rime awareness. *Taiwan Journal of Linguistics, 3*(1), 65–100.

Chen, H., & Chen, R. (2006). *Raven's standard progressive matrices*. Taipei: Chinese Behavioral Science Corporation.

Chiat, S. (1983). Why Mikey's right and my key's wrong: The significance of stress and word boundaries in a child's output system. *Cognition, 14*(3), 275–300. doi:10.1016/0010-0277(83)90007-0.

Choi, J.-Y., Hasegawa-Johnson, M., & Cole, J. (2005). Finding intonational boundaries using acoustic cues related to the voice source. *The Journal of the Acoustical Society of America, 118,* 2579–2587. doi:10.1121/1.2010288.

Chung, W.-L., & Bidelman, G. M. (2016). Cortical encoding and neurophysiological tracking of intensity and pitch cues signaling English stress patterns in native and nonnative speakers. *Brain and Language, 155–156,* 49–57. doi:10.1016/j.bandl.2016.04.004.

Chung, W.-L., & Hu, C.-F. (2007). Morphological awareness and learning to read Chinese. *Reading and Writing: An Interdisciplinary Journal, 20,* 441–461. doi:10.1007/s11145-006-9037-7.

Chung, W.-L., & Jarmulowicz, L. (2017). Stress judgment and production in English derivation, and word reading in adult Mandarin-speaking English learners. *Journal of Psycholinguistic Research.* doi:10.1007/s10936-017-9475-1.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284–290. doi:10.1037/1040-3590.6.4.284.

Common Core State Standard Initiative. (2015). *Reading: Foundational skills.* Retrieved from http://www.corestandards.org/ELA-Literacy/RF/K.

Corriveau, K., Pasquini, E., & Goswami, U. (2007). Basic auditory processing skills and specific language impairment: A new look at an old hypothesis. *Journal of Speech, Language, and Hearing Research, 50,* 647–666. doi:10.1044/1092-4388(2007/046).

Cutler, A. (1996). Prosody and the word boundary problem. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 87–99). Mahwah, NJ: Lawrence Erlbaum Associates.

Dowling, W. J. (1982). Melodic information processing and its development. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 413–429). New York: Academic Press.

Duanmu, S. (2000). *The phonology of Standard Chinese*. New York, NY: Oxford University Press.

Echols, C. H. (1996). A role for stress in early speech segmentation. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 151–170). Mahwah, NJ: Lawrence Erlbaum Associates.

Foxton, J. M., Talcott, J. B., Witton, C., Brace, H., McIntyre, F., & Griffiths, T. D. (2003). Reading skills are related to global, but not local, acoustic pattern perception. *Nature Neuroscience, 6,* 343–344. doi:10.1038/nn1035.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech, 1*(2), 126–152.

Goswami, U., Gerson, D., & Astruc, L. (2010). Amplitude envelope perception, phonology and prosodic sensitivity in children with developmental dyslexia. *Reading and Writing: An Interdisciplinary Journal, 23,* 995–1019.

Goswami, U., & Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives. *Laboratory Phonology, 4,* 67–92.

Goswami, U., Mead, N., Fosker, T., Huss, M., Barnes, L., & Leong, V. (2013). Impaired perception of syllable stress in children with dyslexia: A longitudinal study. *Journal of Memory and Language, 69*(1), 1–17. doi:10.1016/j.jml.2013.03.001.

Goswami, U., Wang, H.-L. S., Cruz, A., Fosker, T., Mead, N., & Huss, M. (2011). Language-universal sensory deficits in developmental dyslexia: English, Spanish, and Chinese. *Journal of Cognitive Neuroscience, 23,* 325–337. doi:10.1162/jocn.2010.21453.

Greenberg, S. (1999). Speaking in shorthand—A syllable-centric perspective for understanding pronunciation variation. *Speech Communication, 29*(2–4), 159–176. doi:10.1016/S0167-6393(99)00050-3.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8,* 23–34.

Holliman, A. J., Wood, C., & Sheehy, K. (2008). Sensitivity to speech rhythm explains individual differences in reading ability independently of phonological awareness. *British Journal of Developmental Psychology, 26,* 357–367. doi:10.1348/026151007X241623.

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

Hu, C.-F. (2013). Predictors of reading in children with Chinese as a first language: A developmental and cross-linguistic perspective. *Reading and Writing: An Interdisciplinary Journal, 26,* 163–187. doi:10.1007/s11145-012-9360-0.

Hu, C.-F., & Catts, H. W. (1998). The role of phonological processing in early reading ability: What we can learn from Chinese. *Scientific Studies of Reading, 2,* 55–79. doi:10.1207/s1532799xssr0201_3.

Hu, C.-F., & Schuele, C. M. (2005). Learning nonnative names: The effect of poor native phonological awareness. *Applied Psycholinguistics, 26,* 343–362. doi:10.1017/S0142716405050204.

Huang, H. S. (2004). *Graded Chinese character recognition test*. Taipein: Psychological Publishing.

Jarmulowicz, L., Taran, V. L., & Hay, S. E. (2007). Third graders' metalinguistic skills, reading skills, and stress production in derived English words. *Journal of Speech, Language, and Hearing Research, 50,* 1593–1605. doi:10.1044/1092-4388(2007/107).

Kehoe, M., Stoel-Gammon, C., & Buder, E. H. (1995). Acoustic correlates of stress in young children's speech. *Journal of Speech and Hearing Research, 38,* 338–350.

Kitzen, K. R. (2001). *Prosodic sensitivity, morphological ability, and reading ability in young adults with and without childhood histories of reading difficulty*. Available from ProQuest Dissertations and Theses database (UMI No. 3005743).

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *The Journal of the Acoustical Society of America, 118,* 1038–1054.

Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics, 23*(1), 81–95. doi:10.1016/j.jneuroling.2009.09.001.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America, 49,* 467–477. doi:10.1121/1.1912375.

Li, H., Shu, H., McBride-Chang, C., Liu, H., & Peng, H. (2012). Chinese children's character recognition: Visuo-orthographic, phonological processing and morphological skills. *Journal of Research in Reading, 35,* 287–307. doi:10.1111/j.1467-9817.2010.01460.x.

Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar*. Berkeley, CA: University of California Press.

Liu, Y.-H., & Hu, C.-F. (2010). The role of Chinese EFL learners' sensitivity to English lexical stress patterns in grammatical category assignments. *English Teaching and Learning, 34*(4), 1–32.

Lorenzi, C., Dumont, A., & Füllgrabe, C. (2000). Use of temporal envelope cues by children with developmental dyslexia. *Journal of Speech, Language, and Hearing Research, 43,* 1367–1379.

McBride-Chang, C., Tong, X., Shu, H., Wong, A. M.-Y., Leung, K., & Tardif, T. (2008). Syllable, phoneme, and tone: Psycholinguistic units in early Chinese and English word recognition. *Scientific Studies of Reading, 12,* 171–194. doi:10.1080/10888430801917290.

Mehta, G., & Cutler, A. (1988). Detection of target phonemes in spontaneous and read speech. *Language and Speech, 31,* 135–156.

Ministry of Education. (2002). *Report on frequent characters and words in elementary school children.* Taipei: Ministry of Education.

Morton, J., & Jassem, W. (1965). Acoustic correlates of stress. *Language and Speech, 8,* 159–181. doi:10.1177/002383096500800303.

Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1998). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology, 71*(1), 3–27. doi:10.1006/jecp.1998.2453.

Ou, S.-C. (2010). Taiwanese EFL learners' perception of English word stress. *Concentric: Studies in Linguistics, 36*(1), 1–23.

Patel, A. D., Peretz, I., Tramo, M., & Labreque, R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain and Language, 61,* 123–144. doi:10.1006/brln.1997.1862.

Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech, 46,* 115–154.

Siok, W. T., & Fletcher, P. (2001). The role of phonological awareness and visual-orthographic skills in Chinese reading acquisition. *Developmental Psychology, 37,* 886–899.

So, D., & Siegel, L. S. (1997). Learning to read Chinese: Semantic, syntactic, phonological and working memory skills in normally achieving and poor Chinese readers. *Reading and Writing: An Interdisciplinary Journal, 9,* 1–21. doi:10.1023/A:1007963513853.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *Test of word reading efficiency-II.* Austin, TX: Pro-Ed.

Vihman, M., & Croft, W. (2007). Phonological development: Toward a "radical" templatic phonology. *Linguistics, 45,* 683–725. doi:10.1515/LING.2007.021.

Wang, H.-L. S., Huss, M., Hämäläinen, J. A., & Goswami, U. (2012). Basic auditory processing and developmental dyslexia in Chinese. *Reading and Writing: An Interdisciplinary Journal, 25,* 509–536. doi:10.1007/s11145-010-9284-5.

Whalley, K., & Hansen, J. (2006). The role of prosodic sensitivity in children's reading development. *Journal of Research in Reading, 29,* 288–303. doi:10.1111/j.1467-9817.2006.00309.x.

Wood, C., & Terrell, C. (1998). Poor readers' ability to detect speech rhythm and perceive rapid speech. *British Journal of Developmental Psychology, 16,* 397–413. doi:10.1111/j.2044-835X.1998.tb00760.x.

Yu, V. Y., & Andruski, J. E. (2010). A cross-language study of perception of lexical stress in English. *Journal of Psycholinguistic Research, 39,* 323–344. doi:10.1007/s10936-009-9142-2.

Zhang, J., & McBride-Chang, C. (2010). Auditory sensitivity, speech perception, and reading development and impairment. *Educational Psychology Review, 22,* 323–338. doi:10.1007/s10648-010-9137-4.

Zhang, J., & McBride-Chang, C. (2014). Auditory sensitivity, speech perception, L1 Chinese, and L2 English reading abilities in Hong Kong Chinese children. *Developmental Psychology, 50,* 1001–1013. doi:10.1037/a0035086.

Zhang, Y., Nissen, S. L., & Francis, A. L. (2008). Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *The Journal of the Acoustical Society of America, 123,* 4498–4513. doi:10.1121/1.2902165.