THE ROLES OF PROSODY IN CHINESE-ENGLISH READING COMPREHENSION

*Shelley Xiuli Tong[a], Rachel Ka Ying Tsui [a], Nicole Sin Hang Law [a], Leo Shing Chun Fung[a], Ming Ming Chiu[b] & Kate Cain [c]

[a] Human Communication, Development, and Information Sciences, Faculty of Education,

The University of Hong Kong, Hong Kong,

[b]Department of Special Education and Counseling, The Education University of Hong Kong,

[c] Department of Psychology, Lancaster University

Accepted for publication in Learning and Instruction, 23 October 2023

Abstract

Background: Despite being an essential component of children's oral reading fluency, prosodic reading, which involves expressive changes in pitch patterns and pause durations, has not been explored in Cantonese-English bilingual children, whose first language (L1) is tonal, non-alphabetic, and whose second language (L2) is non-tonal, alphabetic.

Aims: This study examined the development of prosodic reading and its within- and cross-language associations with reading comprehension among Cantonese-English bilingual children from second to third grade.

Sample: One hundred and twenty-one 7-to 8-year-old Cantonese-English bilingual children completed initial testing in grade 2, with 52 tested in grade 3.

Methods: Prosodic reading was assessed using one Chinese and one English passage, each comprising six types of syntactic structures: declaratives, clause-final commas, *yes-no* questions, *wh-* questions, complex adjectival phrases, and quotatives. Word-reading efficiency, oral passage-reading fluency, and reading comprehension in Chinese and English were also measured.

Results: Spectrographic analyses revealed that these children were aware of language-independent functions and language-specific manifestations of pitch and pause cues within and across their L1 Chinese and L2 English. *Wh* question pitch contours emerged as the most robust link to reading comprehension across both languages, while a crossover effect occurred from Cantonese pitch to English reading comprehension. Shorter pauses for English declarative quotative sentences and phrase-final commas were concurrently associated with greater English reading comprehension.

Conclusions: These findings are interpreted within a new reading framework, the Prosodic Catalysing Hypothesis (PCH), which proposes that pitch and pause production can bridge prosody and syntax to facilitate reading comprehension.

Keywords: prosodic reading, oral reading fluency, bilingual reading comprehension, *wh* questions, spectrographic analysis, Prosodic Catalysing Hypothesis (PCH)

1. Introduction

When reading stories to children, adults often employ prosodic reading, using animated and expressive tones to act out the characters. Although this is done to entertain the children, something more fundamental is potentially at play. Prosodic reading, which involves expressive changes in pitch patterns and pause durations, has been recognized as an essential component of children's oral reading fluency (e.g., for a review, see Kuhn et al., 2010). Recent studies suggest that Western children who demonstrate stronger prosody while reading develop more proficient reading comprehension skills in English and other alphabetic languages (e.g., Arcand et al., 2014; Miller & Schwanflugel, 2008; Ravid & Mashraki, 2007). However, no study to date has examined the role of prosodic reading in tonal, non-alphabetic languages, such as Chinese, or its role in bilingual literacy, particularly when the children's first language (L1) is tonal, non-alphabetic, such as Cantonese, and their second language (L2) is non-tonal, alphabetic, such as English. Thus, an overarching goal of the present study is to investigate within- and cross-language roles of prosodic reading in reading comprehension among Cantonese-English bilingual children. In doing so, we first characterize these bilingual children's use of pitch and pause in reading syntactically complex sentences, and explore the potential influence of L1 prominent tone features on L2 prosodic acquisition. Next, we examine to what extent various pitch and pause measures concurrently and longitudinally link with L1 and L2 reading comprehension. Our research demonstrates that of the two most salient prosodic features—pitch and pause—*wh* question pitch contours emerge as the prominent, and perhaps even language-independent, correlator of Chinese and English reading comprehension in Hong Kong Cantonese-English bilingual children.

1.1. Language-Independent and Language-Specific Prosody between Cantonese and English

In the realization and function of suprasegmental features (i.e., prosody), Cantonese and English exhibit certain similarities, especially with regard to pitch and pause. Pitch is a perceptual sensation of fundamental frequency (*f*0) that often marks sentence-final positions, commands, and neutral statement questions with falling tones, and sentence non-final phrases, tentative statements, requests, and sympathetic questions with rising tones (Cruttenden, 1997). A global pitch declination tendency, that is an initial rising then falling, appears in declarative sentences (Cruttenden, 1997), whilst yes-no questions evince a continuous rising pitch (Pennington & Ellis, 2000). Pitch or intonation contours signal syntactic units and serve as important universal clues to the chunking of utterances into meaningful syntactic sets (Cutler et al., 1997; Dowhower, 1991; Erekson, 2010). Specifically, intonational contours are universally characterized by falling or rising *f*0 patterns at phrase/clause boundaries and/or terminal markers of sentences (Dowhower, 1991). In addition, larger fall-rise pitch patterns tend to occur at stronger, rather than weaker, syntactic boundaries (Cooper & Sorensen, 1981).

Likewise, pauses, which are brief silent periods between units of speech, coordinate syntactic grouping by marking the boundaries between and within sentences, with intrasentential pauses often shorter than intersentential ones (e.g., Vaissière, 1983). Also, syntactically complex phrases often carry greater information loads, yielding longer pauses before and after them (e.g., Zvonik & Cummins, 2003). Given these similarities, we expect Cantonese-English bilingual children to employ the same strategies when producing and encoding similar pitch- and pause-related structures in their two languages, suggesting a language-independent prosodic pattern.

Despite these similarities, pitch and pause are strikingly different in Cantonese and English. As a complex tonal language, Cantonese employs an extensive pitch range to concurrently: 1) realize six lexical tones[1] (i.e., high level-55, high rising-25, mid-level-33, low falling-21, low rising-23, and low level-22) that distinguish segmentally identical words or morphemes (e.g., /fu55/ *skin*, /fu25/ *tiger*, /fu33/ *trousers*, /fu21/ *symbol*, /fu23/ *woman*, and /fu22/ *father*); and 2) mark intonation (Fox et al., 2008). Consequently, pitch rise occurs at the final syllable of interrogative questions (e.g., yes-no) rather than across the entire utterance (e.g., Gu et al., 2006). Cantonese is characterized as having a weak macro rhythm (June, 2014), as each syllable can be assigned to a different tone and no regular rising and falling tonal patterns appear at the phrase level. For example, the syllable /si/ can convey various meanings depending on the tone. It can mean "poem" with a high-level tone, "history" with a high rising tone, "to try" with a mid-level tone, "time" with a falling tone, "city" with a low rising tone, and "matter" with a low level tone. As Cantonese has no consistent pattern of stress or emphasis on specific syllables to create a rhythmic pattern (ba-NA-na /bəˈnɑːnə/), and the tone of each syllable exerts a significant impact on the meaning of a phrase or a sentence, the overall rhythm of Cantonese is difficult to predict and identify. Furthermore, Cantonese is a *wh*-in-situ language, meaning *wh* words (e.g., /sɵy21/ *who* and /sɐm22mɔ23/ *what*) are not relocated to the beginning of a question, but remain in their original position (Gu et al, 2006; Huang, 1982; Ma et al., 2006). In addition, Cantonese *wh* questions typically end with the final particle /nɛ:55/, which is marked by a rising contour (Yang et al., 2020).

In contrast to Cantonese, English is a stress-timed language whereby stressed syllables in a word serve as "pitch landmarks" for intonation so that the rising *f*0 contour for yes-no questions is signalled by a rising pitch across the entire utterance (e.g., Roach, 2018). English is characterized as having a medium macro rhythm (June, 2014), because English has

[1] The five-level tone transcription system developed by Chao (1930) uses 1 (lowest) to 5 (highest) to describe relative height, shape and duration of pitch contour. High, middle and low indicate pitch ranges while level, rising, and falling denote the slopes of the pitch contours.

stress, and stressed syllables are typically marked as high pitch in declarative statements. English also requires that *wh* words be transposed to the beginning of a *wh* question (e.g., *What* book are you reading?), with the terminal contour of the final syllable not always eliciting an increase in pitch (Chafe, 1988) and commonly marked by a falling pitch contour (Roach, 2018).

Cantonese and English also exhibit distinctive pause structures at the phrasal and sentential levels. In English, pauses are mostly required following basic quotatives (Cooper & Paccia-Cooper, 1980), and always required following phrase-final comma locations (e.g., *“Yes,” says Apple**,** and she starts crying*). However, when commas appear in a series, such as ones using complex adjectival phrases (e.g., *tall, thin, yellow Banana*), they require less pronounced pauses (Chafe, 1988). In contrast, a slight pause often separates complex adjectival phrases sequenced in Cantonese (e.g., /kʰɐn21 lɪk22/ *diligent,* /jit22 sɐm55/ *warm-hearted,* /jʊŋ23 kɐm25/ *courageous*) (Matthews & Yip, 2011).

Given these similarities and differences (see Table 1 for a summary), the first objective of our study is to characterize the acoustic properties of pitch and pause in different Cantonese syntactical structures during prosodic reading. As most Hong Kong children learn L2 English at a relatively young age (i.e., three years old or earlier), they must cope with two prosodically distinct speech codes to establish dual prosodic representations. Research on English prosody acquired in a multilingual context, especially in a Cantonese-dominant ecology, suggests that Hong Kong English exhibits tonal features. Specifically, at the word level, stressed syllables are assigned high tones, whilst unstressed syllables are pronounced with low tones (e.g., Chen & Au, 2004; Wee, 2008). At the phrase level, a pattern of sequential tones occurs (Lim, 2009).

The influence of L1 prominent tone features on L2 prosodic acquisition has been explained by the dynamic interactive processing account, which suggests that processes and representations involved in a bilingual's acquisition of Cantonese tone and English stress are shared and improve with age (Tong et al., 2015). For example, Cantonese-English bilingual children exhibit comparable and correlated L1 tone and L2 stress perception performance, while adults demonstrate a superior L2 stress than L1 tone, indicating an age-related facilitative effect of tone on stress (Tong et al., 2015). Thus, our second objective is to identify the language-independent and language-specific prosodic reading features in L1 Chinese and L2 English as indexed by pitch and pause. Since Cantonese is a pitch-dominant language, we hypothesize that children will demonstrate an early mastery of pitch in their L1 Chinese, but require more time to master pitch patterns in their L2 English. Furthermore, we expect these children to exhibit both language-independent and language-specific pause patterns when reading Chinese and English. As they develop, their pause durations should decrease across both languages.

Table 1
Comparison of Prosodic Features between Cantonese and English

		L1 Cantonese	L2 English
Pitch features	Lexical	Tone: a property of segments	Stress: not a property of segments
	Phrasal	A weak macro rhythm: no regular rising and falling tonal patterns	A medium macro rhythm: a large inventory of pitch accent, with stressed syllables typically marked as high pitch (H*), and relatively frequent deaccenting
	Sentential	Basic declaratives: pitch declination across the utterance	Basic declaratives: pitch declination across the utterance
		Yes-no questions: pitch rise occurs at the final syllable only	Yes-no questions: pitch rise occurs across the entire utterance
		Wh questions: pitch rise marks the terminal contour of the final syllable	*Wh* questions: pitch fall marks the terminal contour of the final syllable
Pause features	Phrasal and Sentential	Shorter intrasentential pauses than intersentential ones	Shorter intrasentential pauses than intersentential ones
		Pauses required for all punctuation, including commas, colons, and full-stops as indicated in complex adjectival phrase commas, phrase-final commas, and declarative quotatives.	Complex adjectival phrase commas: No pronounced pauses needed for a series
			Phrase-final commas: Pauses required for phrase-final locations
			Declarative quotatives: Pauses may or may not be required after declarative quotatives

1.2. Prosodic Reading and Reading Comprehension: Theoretical Foundations and a New Prosodic Catalysing Hypothesis

While many studies have demonstrated prosody's organizational and scaffolding roles in spoken language acquisition and comprehension (e.g., Abboub et al., 2016; Frazier et al., 2006), some suggest that even infants are sensitive to prosodic phrasing (e.g., Nazzi et al., 2000) and recognize its connection with syntactic structures (e.g., Soderstrom et al., 2003). Moreover, psycholinguistic studies of adult sentence processing have provided compelling evidence for the Implicit Prosody Hypothesis (IPH, Fodor, 1998; 2002), which assumes that speech prosody is naturally imposed on sentences during silent reading and that these activated prosodic representations influence a reader's interpretation of the text (for a review, see Breen, 2014). Evidence for this comes from event-related brain potential studies in a range of languages. For example, a closure positive shift (CPS) component, which is evoked by overt prosodic boundaries when listening to sentences, is also induced by commas during German adults' silent reading, indicating implicit prosodic phrasing (Steinhauer & Friederici, 2001). Similarly, Korean readers exhibit CPS during online sight sentence processing after a long subject noun phrase (Hwang & Steinhauer, 2011). Furthermore, Japanese speakers prefer to assign a subvocal prosodic boundary after a long subject noun phrase during silent reading of relative clauses (Hirose, 2003). Finally, German adult readers' eye movements show that lexical stress patterns dominate early syntactical processing when reading syntactically ambiguous sentences (McCurdy et al., 2013). Of note, this effect is influenced less by the context, indicating an effect of implicit prosody on sentence reading.

A series of correlational studies on children's reading acquisition reported that their prosodic competence, defined as the metalinguistic ability to analyse and process prosodic features, was positively associated with their reading comprehension (for a review, see Wade-Woolley et al., 2021). This result aligns with two prominent reading theories: the

Simple View of Reading (SVR), which assumes that reading comprehension is determined by language comprehension and word reading (Hoover & Gough, 1990), and the Reading Systems Framework (RSF), which proposes that a linguistic system comprising phonology, syntax, and morphology contributes to both orthographically precise, phonologically specific, and semantically connected lexicons, and the integration of lexical information and world knowledge to construct meaning from text (Perfetti & Stafura, 2014). Both theories emphasize the critical role of oral language skills by assuming that language comprehension or linguistic systems undergird reading comprehension. However, prosody, as an essential structural element of phonology, has not been explicitly recognized by these two models.

Thus, a recent systematic review by Wade-Woolley et al. (2021) expanded the RSF framework by adding prosodic competence as a component of linguistic phonology, and tone and word stress as components of lexical phonology. According to this revised RFS model, prosody is conceptualized as "a source of linguistic knowledge" that influences: 1) orthographic systems through overt marking (e.g., tone is explicitly marked in Mandarin Pinyin: /mā/ (*mother*), /má/ (hemp), /mǎ/ (*horse*), and /mà/ (*to scold*)) or implicit constraints (e.g., the graphotactic patterns of English word stress placement); 2) word identification processes by specifying phonological lexicons through word stress or lexical tones; and 3) comprehension processes by strengthening syntactic parsing and representation through prosodic reading.

Together, all of these previous theoretical frameworks emphasize that reading comprehension is deeply rooted in spoken language, implying a fundamental role for prosody. However, unlike spoken language comprehension, which occurs naturally and has its prosodic cues overtly marked by acoustic input, reading comprehension is a complex process that necessitates additional cognitive skills, specifically, visual word reading and implicit or silent prosodic representation. The SVR and RSF include word reading, but neglect prosodic representation. Despite the explicit recognition of prosodic reading in Wade-Woolley et al.'s (2021) extended RSF, it neither explains nor predicts how an individual's ability to produce fine-grained acoustic-phonetic features (that is pitch and pauses related to prosodic structures) may be related to their age or level of reading comprehension. Given this conceptual gap, we propose the Prosodic Catalysing Hypothesis (PCH) as presented in Figure 1.

Similar to the models outlined above, the PCH assumes that visual word recognition—comprising interlocking segmental and suprasegmental phonological structures, orthographic components, and semantic constituents—lays the foundation for reading comprehension. However, unlike the other three models, PCH posits that prosodic reading functions as a "catalyst" for optimizing the connection between prosodic cues and syntactic structures, and between word reading and reading comprehension. Pitch and pauses are the psychoacoustic manifestations or outcomes of prosodic structures influenced by both bottom-up factors, such as word-level segmental components and suprasegmental features (e.g., lexical tone in Chinese), and top-down factors, such as different syntactic structures. Optimization occurs at both word and sentence levels.

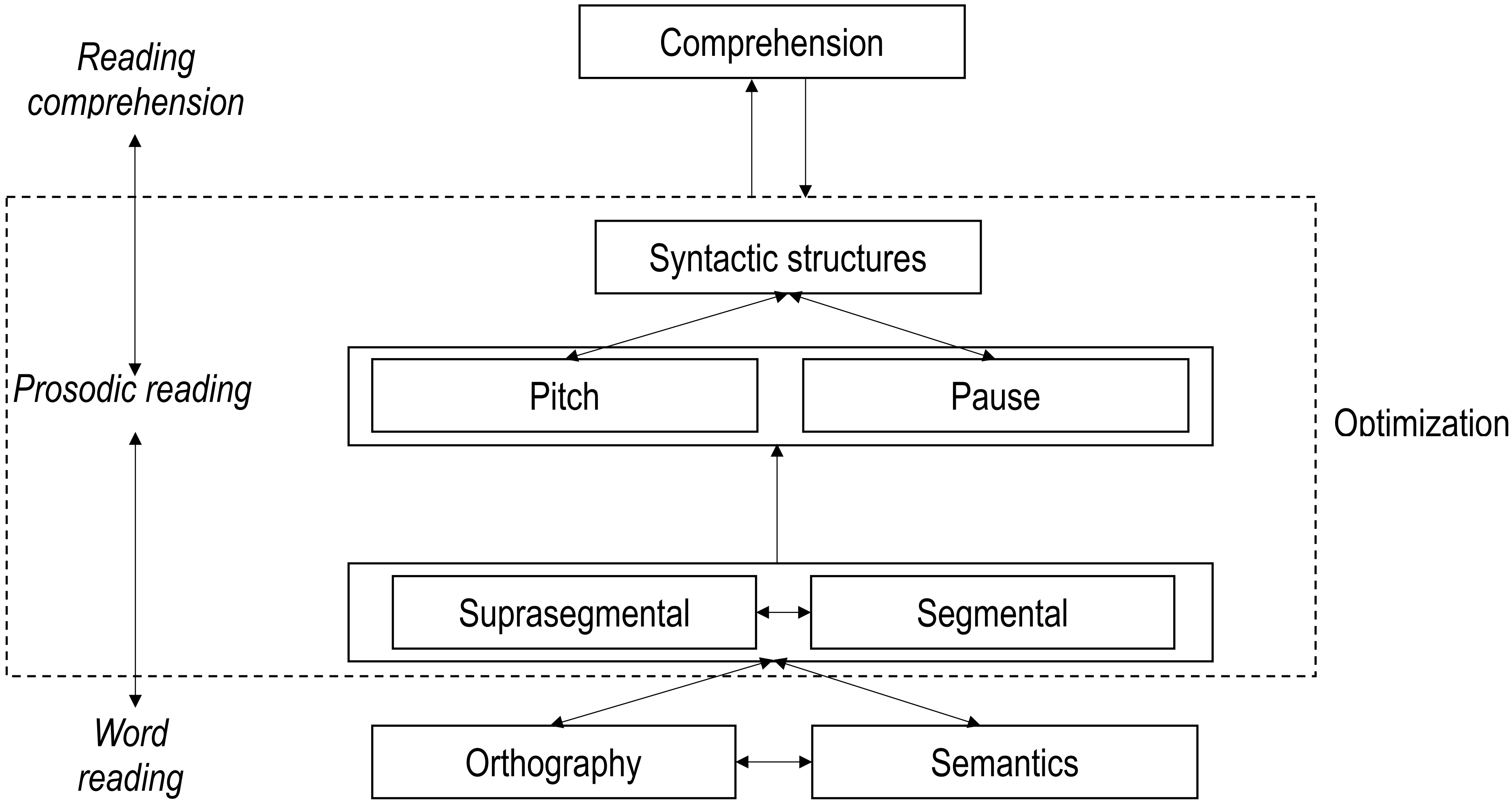
At the word level, suprasegmental features, especially pitch (e.g., lexical tone in Cantonese), and segmentally related information, serve as concurrent indicators of lexical meaning, aiding comprehension (e.g., Cantonese tonal syllable /si55/ "poem" versus syllable /si25/ "history"). At the sentence level, as prosodic reading necessitates assigning of appropriate pitch and pause to different syntactic structures, prior knowledge of prosody-syntax correspondence, established through spoken language, is naturally activated and imposed during the chunking of words into comprehensible phrases and clauses. As such, this may facilitate the prediction and inference of meaning and the construction of meaning-based representations of text. Evidence of the effects of syntax on the detection of prosodic boundaries (e.g., Cole et al., 2010; Buxó-Lugo & Watson, 2016), suggests that an individual's knowledge of syntactic structures guides the construction of prosodic representation, manifested as the appropriate use of pitch and pause patterns to parse sentences. As a result, with prosody signaling syntax, and syntax bootstrapping prosody, a bidirectional (i.e., feedforward and feedback) process optimizes the probabilistic mapping between prosody and syntax, and synchronizes a multi-level representation of words and sentences for generating the optimal interpretation of the text. Furthermore, pitch and pause, as the core perceptual indices of acoustic-phonetic features of prosody, simultaneously contribute to visual word recognition and text comprehension. Such parallel processing of low- and high-level information is supported by the automaticity theory (LaBerge & Samuels, 1974), which suggests that once a child's word reading becomes automatic and effortless, they can reallocate attention to higher level (e.g., syntactic) processing.

Additionally, according to the PCH framework, prosodic reading acts as a necessary stepping stone toward fluent reading comprehension by optimizing word reading and prosody-syntax mapping. As prosodic reading synchronizes written text with speech, children more effectively chunk connected texts into smaller syntactic units (e.g., phrases and clauses) by identifying (i.e., through hearing) appropriate boundaries for assigning emphasis, intonation, and pauses (Rasinski, 1994; Schreiber, 1991). To dramatize expressions that approximate normal speech, children must navigate terminal and internal punctuation and access morphological and syntactic cues in order to place appropriate semantic and syntactic emphasis that establishes a coherent semantic representation of the text (Ravid & Mashraki, 2007). Thus, prosodic reading facilitates the transition from word-by-word reading to reading in meaningful phrases, such as syntactically appropriate chunks (Rasinski, 1994), that ultimately expedites simultaneous constructions and integration of multi-level linguistic representations for text comprehension.

One testable hypothesis of PCH is that an individual's ability to use pitch and pause structures when reading aloud sentences with different syntactic structures may be related to their ability in reading comprehension. Most models of language comprehension emphasize an interactive process involving bottom-up and top-down processing (Gibson et al., 2013). Likewise, PCH recognizes that reading comprehension ability may influence the use of prosody when reading aloud. However, the current study was not designed to test causal or directional hypotheses, instead, it focuses on the individual's ability to use pitch and pause to: 1) read aloud sentences with distinct syntactic structures; and 2) their relation to reading comprehension in both Chinese and English.

Figure 1

The Prosodic Catalyzing Hypothesis (PCH)



1.3. Previous Studies on Prosodic Reading and Reading Comprehension

Our PCH framework is informed by a number of studies that suggest a link between prosodic reading and reading comprehension in monolingual alphabetic and nonalphabetic readers. For example, Miller and Schwanenflugel's (2006) cross-sectional study of English-speaking children found that prosodic reading, measured by pause duration and expression of appropriate intonation through pitch changes, as well as rate and accuracy, accounted for additional substantial variance in reading comprehension. Beyond oral and word reading fluency, pitch changes in declarative sentences and yes-no questions independently contributed to reading comprehension. In a study of Hebrew-speaking fourth graders, Ravid and Mashraki (2007) found that prosody, measured as appropriate pauses and intonation modulation at prosodic sites in the text, contributed significantly to reading comprehension, a relation moderated by morphological skills. Similarly, prosodic reading in French, indexed by inappropriate pauses and attention to punctuation, explained substantial variance in reading comprehension of second graders (Arcand et al., 2014). Moreover, a recent study by Chung and Bidelman (2022) found that oral reading prosody, including pause intrusions, phrase-final comma pause duration, and child-adult pitch match, were significant correlates of Chinese reading accuracy and reading comprehension in Mandarin-speaking third graders.

Finally, a study of Spanish children with poor reading comprehension found that they made more low-frequency word-reading mistakes than better comprehenders, and assigned more inappropriate pauses and pitch ranges in declarative and interrogative sentences (Álvarez-Cañizo et al., 2015). A similar study of Dutch fifth-grade students with poor reading comprehension skills found that they performed significantly worse than their age-matched peers on both text and speech prosody, which was measured using the multidimensional fluency scale to assess expression, phrasing, smoothness, and pace while reading narrative passages or telling stories. Furthermore, poor comprehenders exhibited significant weaknesses compared to their comprehension level-matched controls on a speech rhythm task measuring sensitivity to pitch, intensity and duration of filtered speech. These findings indicate that speech prosody can distinguish between poor comprehenders and typically developing comprehenders (Groven et al., 2018).

The longitudinal studies that have examined the developmental relation between prosodic reading and reading comprehension report mixed results. In support of a relationship, Miller and Schwanenflugel (2008) found that prosodic reading in grades 1 and 2 contributed to reading comprehension in grade 3. In addition, Lai and colleagues (2014) showed that a latent construct of text reading fluency, which included prosodic reading, was linked with reading comprehension across three time points in grade 2 English monolingual readers. However, a study following Dutch children from grades 4 to 6, found grade-dependent connections between text reading prosody and reading comprehension (Veenendaal et al., 2016): specifically, whilst reading comprehension was significantly associated with text reading prosody from grades 4 to 5, text reading prosody accounted for unique variance in reading comprehension in grade 6 only. Finally, after controlling for decoding and word reading efficiency, Fernandes et al. (2018) did not find any evidence for a longitudinal relation between prosodic reading and reading comprehension in Portuguese children between grades 2 and 5.

It is important to note that prosodic reading was assessed using a subjective rating scale, rather than an objective spectrographic analysis, in most of these studies (for an exception see Miller & Schwanenflugel, 2006; 2008). A recent meta-analysis of 35 studies in alphabetic languages, 30 of which used rating scales and 5 of which used spectrographic measures, found a moderate association (r = .51) between prosody and reading comprehension. The association was stronger in studies using rating scales than those using spectrographic analysis (Wolters et al., 2022). However, the limited number of studies that have used spectrographic analysis to measure prosody emphasizes the need for further research to determine if pitch and pause pattern is associated with reading comprehension, especially in non-alphabetic languages.

Furthermore, with most previous studies limited to non-tonal alphabetic monolingual readers, one particular question remains: Does prosodic reading relate to reading comprehension in non-alphabetic, complex tonal-language-speaking children? This question is uniquely important because prosody varies considerably in its acoustic-phonetic manifestations and communicative functions across languages (van Ommen et al., 2020), a fact remarkably evident in the use of pitch between tonal languages (e.g., Cantonese) and non-tonal languages (e.g., English). Specifically, unlike English where pitch is used primarily to signal intonation contours, Cantonese employs pitch at both lexical and intonational levels, resulting in inevitable competition between the lexical tone and intonation of the phrase-final or sentence-final word (Fox et al., 2008). No previous studies have implemented parallel manipulations of syntactic structures in Chinese and English to tap into pitch and pause structures simultaneously. Thus, our PCH framework extends previous studies and contribute to our understanding of the role of prosody in reading comprehension by examining the relation of prosodic reading in Cantonese-speaking children's L1 Chinese and L2 English reading comprehension, with a particular focus on within- and cross-language contributions.

1.4. ***The Present Study***

Despite the established understanding of its significance in reading comprehension among English and other alphabetic monolingual readers, a crucial yet unexplored question remains regarding the role of prosodic reading in bilingual reading comprehension, particularly for children whose first language is tonal Cantonese while their second language is non-tonal English. Thus, this study makes a significant advance in our understanding of the role of prosody in reading comprehension by exploring the role of prosodic reading, specifically pitch and pause, in the reading comprehension abilities of Cantonese-English bilingual children. Additionally, the study investigates how the prominent tone features of the first language (L1) influence the acquisition of prosodic reading in the second language (L2) among Cantonese-English bilingual children. These interrelated research objectives are incorporated into the Prosodic Catalysing Hypothesis (PCH), which posits that prosodic reading functions are a catalyst for optimizing the connection between prosodic cues and syntactic structures, and between word reading and reading comprehension. Our research questions and hypotheses are specified below.

First, in line with the parallel and interactive nature of the PCH, which emphasizes the integration and prediction of prosodic and syntactic representations, we hypothesize that bilingual children will retrieve and employ prototypic prosodic patterns, especially pitch and pause values that correspond with various syntactic structures, to produce language-independent and language-specific prosodic features when reading aloud syntactically complex sentences. Furthermore, we expect that bilingual children's proficiency in utilizing pitch and pause during prosodic reading will be related to their abilities in text reading comprehension. These two core PCH hypotheses are formulated in four research questions: First, what are the language-independent similarities and language-specific differences between Cantonese-English bilingual children's L1 Chinese and L2 English prosodic reading as measured by pitch change and pause duration? Second, as children progress from grade 2 to grade 3, what, if any, developmental changes occur in the use of pitch and pause structures in their L1 and L2 prosodic reading? Third, to what extent does prosodic reading, as indexed by pitch change and pause duration, explain variance in reading comprehension skills beyond word reading efficiency and oral reading fluency? Fourth, does a cross-language relationship exist between prosodic reading and reading comprehension in Cantonese-English bilingual children?

To address these questions, we first conducted a spectrographic analysis of pitch change and pause duration during the reading of six types of syntactically complex sentences (basic declaratives, declarative quotatives, *wh* questions, yes-no questions, complex adjectival phrase commas, and phrase-final commas) in connected Chinese and English texts. We then evaluated both concurrently and longitudinally the within- and cross-language contributions of different pitch and pause variables to reading comprehension, while accounting for word reading efficiency and oral reading fluency. Our results will provide critical evidence to test the basic premises of the PCH model by clarifying how prosodic reading is related in reading comprehension. We do this by examining a population of bilingual children who cope simultaneously with prosodically distinct tonal logographic and non-tonal alphabetic languages.

2. Method

2.1. Participants

Participants were 121 typically developing native Cantonese-speaking second graders (70 females, 51 males; *mean* age = 7 years, 11 months, *SD* = 4.55 months) who started learning English as an L2 around the age of 3 years and studied in Hong Kong mainstream primary schools across three districts (i.e., Hong Kong Island, Kowloon, and the New Territories) where Cantonese is the medium of instruction. Of these participants, 54 returned one year later to complete the same testing. Technological issues rendered the acoustic data of two participants unusable. Thus, 52 participants were included in Grade 3 (32 females, 20 males; mean age = 8 years, 11 months, SD = 3.27 months). For that reason, we analyzed data from the two groups separately: (a) cross-sectional analyses of 121 participants in Grade 2 and (b) cross-sectional (Grade 3) and longitudinal analyses (Grades 2 and 3) of 52 participants. For effect sizes of .40, .30, .20, and .10, the statistical power for a sample of (a) 121 participants is .99, .92, .60, and .19, and (b) 52 participants is .85, .59, .30, and .10, respectively (Kohn & Senyak, 2021). Our University's Human Research Ethics Committee granted ethics approval for this study.

According to parental report, the participants came from various socio-economic backgrounds, but the majority of their parents earned HK$10,000-$40,000 (~US$1,290-$5,161) per month, and their mothers graduated from high school or above. Second graders were chosen because children at this age are expected to develop word reading automaticity and demonstrate varying degrees of fluent prosodic text reading skills (e.g., McBride, 2015; Miller & Schwanenflugel, 2008), allowing for an examination of possible developmental changes. Furthermore, this choice aligns with previous seminal studies of monolingual English-speaking second graders (Miller & Schwanenflugel, 2006; 2008), enabling a comparison of English prosodic reading between bilingual and monolingual children.

In Hong Kong, English L2 reading instruction starts when children enter kindergarten at the age of 3 years. By age 6, these children can typically read and write simple English sentences before entering primary school (Yeung, 2018). Starting from the first grade, both Chinese and English are formally taught as language subjects in mainstream primary schools. While Cantonese is commonly used for daily communication in Hong Kong, children are taught to read and write standard Chinese, known as Modern Standard Written Chinese (Curriculum Development Council, 2017). Although Cantonese and Mandarin are distinct Chinese dialects with variations in vocabulary and grammar, their basic linguistic features are fundamentally similar (Chao, 1968; Tang & Cheung, 2014). However, Cantonese has preserved more ancient elements of the Chinese language (Huang et al., 2022; Matthews, 2014).

Furthermore, as Chinese classes in Hong Kong primary schools mostly use texts directly adopted from Mandarin, standard Mandarin Chinese vocabulary and grammar are taught, but the texts are transcribed for Cantonese pronunciation when reading aloud

(Education Bureau of the Hong Kong Special Administrative Region, 2004). Thus, Cantonese refers to a dialect, and Chinese refers to the Modern standard written form of the language. According to the curriculum of Hong Kong mainstream primary schools (Hong Kong Examinations and Assessment Authority, 2020), children in Grade 3 should develop listening, speaking, reading, and writing skills in Chinese and English. For Chinese language skills, they are expected to understand the meaning of speech and conversations, deliver their messages clearly, use vocabulary and punctuation correctly, and summarise the meaning of texts. For English language skills, they are expected to identify and discriminate sounds, stress, and intonation between common words; comprehend both explicit and implicit meanings from short, simple speech and conversations; provide short answers to simple questions; accurately pronounce simple and familiar words; effectively interact with others in English; and use the basic rules of English grammar to write and understand the meaning of short and simple texts.

2.2. Instruments and measures

By employing a repeated-measures design, participants completed the same acoustic-phonetic assessment of prosodic reading of Chinese and English passages, word reading efficiency, passage reading fluency, and reading comprehension across Grades 2 and 3, and a nonverbal ability control measure in Grade 2 only, all of which are described below.

2.2.1. Prosodic reading: stimuli and procedure

To assess the children's L1 Chinese and L2 English reading prosody in Grades 2 and 3, two structurally similar passages, one in Chinese and one in English, were created in accordance with previous studies (Miller & Schwanenflugel, 2006; 2008). Each passage contained six types of sentence structure that may elicit language-independent and language-specific prosodic reading features: (1) *basic declaratives* that may elicit pitch decline at the end of the sentence in both Chinese and English; (2) *declarative quotatives* that may elicit short pauses following quotations in both Chinese and English; (3) *wh questions* that may elicit pitch rise in Chinese but pitch fall in English; (4) *yes-no questions* that may elicit pitch rise in both Chinese and English; (5) *complex adjectival phrase commas* that may elicit pauses in Chinese but not in English; and (6) *phrase-final commas* that may elicit pauses following phrases in both Chinese and English. Each of the six types occurred three times in the Chinese passage and twice in the English passage.

Because decoding errors prevent the meaningful measurement of reading prosody, words in the passages were carefully selected according to Hong Kong Chinese Lexical Lists for Primary Learning (2007) and Wordlists for the Primary English Language Curriculum (n.d.) (HKSAR Education Bureau, 2009). The Chinese passage consisted of 255 words and 18 sentences, whereas the English passage consisted of 99 words and 12 sentences. Our pilot study of 42 second-grade Cantonese-English bilingual children (27 boys and 15 girls) showed that each passage was easily decodable by most typically developing Hong Kong second graders. Chinese and English passages and comprehension questions, with subscribed numbers indicating the sentence types described above but not shown to participants, are as follows:

Chinese L1 passage:

彩虹覺得自己很美[麗，於] [6]是驕傲起來。
Rainbow feels so [beautiful, that] [6] it makes her very proud.
她笑白雲的衣服太簡單。[1]
She laughs at Cloud’s clothes for being too plain. [1]
彩虹驕傲地問白雲：「你美麗嗎？」[4]
“Are you beautiful?”[4] Rainbow asks Cloud proudly.
白雲說 [2]：「我不美，可是小草[綠油油、軟綿綿、亮晶晶] [5]的衣服把大地打扮得很美麗呢！」
“I’m not beautiful, but Grass’s [green, soft and sparkly] [5] clothes are dressing the ground so beautifully!” says Cloud. [2]
彩虹找到小[草，然] [6]後生氣地問：「你比我美麗嗎？」[4]
Rainbow finds [Grass and] [6] asks angrily, “Do you think you’re more beautiful than me?”[4]

小草說：「你有[七彩、彎彎、長長] [5]的裙子，我只有綠衣服，怎麼能和你比較呢？[3] 你和小花比較吧！」

"Your dress is [rainbow-colored, flowy and long] [5]. I only have green clothes. How can I compare with you? [3] You should compare yourself with Flower!" says Grass.

彩虹又去問小花：「你說我們誰更美麗呢？」[3]

Rainbow then goes and asks Flower, "Who do you think is more beautiful, me or you?"[3]

小花笑着對她說 [2]：「當然是你更美麗！」

Flower smiles and says to her [2], "Of course, you're more beautiful!"

彩虹聽了開開心心地唱起歌。[1]

Rainbow hears this and happily sings a song. [1]

太陽公公看見了，笑着對她說 [2]：「你很美，可是你為大自然做了什麼？[3] [勤力、熱心、勇敢] [5] 的小草和小花都為大自然努力工作！你覺得我說得對嗎？」[4]

Grandpa Sun sees this and says to her while smiling [2], "You are beautiful, but what have you done for nature? [3] [Diligent, enthusiastic and courageous] [5] Grass and Flower both work hard for nature! Would you say I'm correct?"[4]

彩虹沒有回答，她只愛和別人比[較，從][6]來都沒有為別人做過甚麼。

Rainbow does not answer. She only loves to compare herself with [others, but] [6] never does anything for others.

她慚愧得消失於天空中。[1]

She feels ashamed as she disappears from the sky. [1]

Chinese Comprehension Questions (in question-and-answer format):

1. 誰喜歡和別人比較? (正確答案：彩虹)
   Who likes to compare with others? (Correct answer: Rainbow)
2. 小草穿了甚麼衣服? (正確答案：綠衣服)
   What colour of clothes is Grass wearing? (Correct answer: Green clothes)
3. 彩虹和誰比較過? (正確答案：白雲、小草和小花 *必須全部回答)
   Who does Rainbow compare herself with?
   (Correct answer: Cloud, Grass and Flower *answer must include all)
4. 你認為彩虹最後會有甚麼感受? (正確答案：慚愧 / 不開心 / 後悔)
   How do you think Rainbow will feel at the end?
   (Correct answer: ashamed/unhappy/regretful)
5. 你認為彩虹為什麼沒有回答太陽公公? (正確答案：因為彩虹知錯了 / 從來都沒有為別人做過什麼)
   Why do you think Rainbow does not answer Grandpa Sun?
   (Correct answer: Because Rainbow knows that she is wrong/she has never done anything for others.)

English L2 passage:

Apple, Orange, and Banana are talking. [1]

"I am going to be banana ice cream!" says [tall, thin, yellow] [5] Banana.
"That's nice," says [short, fat, round] [5] Orange.
"I am going to be juice!" [2] says Orange.
"How about you?"[3] Banana asks cute, small, red Apple.
Apple does not say a [word, but] [6] she looks at Orange.
"Apple, are you sad?" [4] says Orange.
"Yes," says [Apple, and] [6] she starts crying.
"Why are you crying?" [3] asks Banana.
"Are you sick?" [4] asks Orange.
"No, I do not know what to be," [2] says Apple.
So, Orange tells Apple to also be juice. [1]

English Comprehension Questions (in question-and-answer format):

1. What is Banana going to be? (Correct answer: (Banana) ice cream)
2. What is Orange going to be? (Correct answer: Juice)
3. Why does Apple cry? (Correct answer: She does not know what to be.)
4. What does Orange tell Apple to be? (Correct answer: Juice)
5. How many fruits are there in the story? (Correct answer: three)

The Chinese and English passages were presented on separate laminated A4 pages, following the format of the Gray Oral Reading Test, Fourth Edition (Wiederholt & Bryant, 2001). Children were instructed to read the passages aloud with appropriate expression and accuracy. To demonstrate this, a recording of expressive reading of another short story in the appropriate language was first played. Children's oral reading production was recorded in a sound-attenuated booth or quiet room (with a sampling rate of 44.1 kHz) via an Edirol USB Audio Capture UA-25, a Lenovo ThinkPad E450 Laptop, and an AKG SE300 B microphone pre-amplifier with a CK91 cardioid condenser microphone capsule. We created a digital .wav file for each participant using Praat (Boersma & Weenink, 2015).

To assess their comprehension of the text for the prosodic reading task, the children answered five orally-presented short-answer literal questions by extracting information directly from the text (e.g., the text explicitly states that Rainbow compares herself with Cloud, Grass, and Flower, and the literal question was: *Who does Rainbow compare herself with?*). A threshold of at least three correct answers was used to indicate adequate comprehension for each passage. All participants passed this criterion.

2.2.2. Reading outcomes and control measures

Word reading efficiency. The one-minute Chinese word reading subtest from HKT-SpLD II (Ho et al., 2007) and the sight word efficiency subtest from TOWRE (Torgesen et al., 2012) were used to assess Chinese and English word reading efficiency, respectively. The former comprised 90 two-character Chinese words while the latter comprised 108 single English real words and 66 nonwords. Each child was asked to read as many of the words as accurately and quickly as possible within 60 seconds for Chinese words, and within 45 seconds for English words. Each correctly read word received one point, for a maximum score of 90 and 174 for Chinese and English, respectively. Test-retest reliability coefficients for an approximately 12-month interval were .75 for Chinese and .86 for English.

Passage reading fluency. In accordance with a standardized measure of oral reading (Weiderholt & Bryant, 2001), passage reading fluency was assessed by calculating the number of words that children correctly read aloud per minute for one Chinese and one English passage. Test-retest reliability coefficients for an approximately 12-month interval were .63 for Chinese and .87 for English.

Reading comprehension. The Chinese reading comprehension test consisted of five individual passages presented in order of increasing difficulty ranging from 175 to 366 words arranged in order of increasing difficulty. Three passages (two narratives and one expository) were translated into Chinese from Form A (level 2A and 3A) of the York Assessment of Reading for Comprehension Passage Reading (Snowling et al., 2011) and modified for a Hong Kong context. The other two passages (one narrative and one expository) were chosen from Chinese extra-curricular books not taught to children in Hong Kong and used in previous studies to assess the Chinese reading comprehension of children at the same age (e.g., Deng & Tong, 2021). Children’s comprehension of each passage was evaluated with eight multiple-choice questions, resulting in 40 questions that included literal, inferential, evaluative, and word-definition. Since the eight questions for the first narrative passage were correctly answered by all Grade 2 children, this passage was removed for the testing of these children at Grade 3. Thus, the total score for Chinese reading comprehension was 40 and 32 for Grades 2 and 3, respectively, and the percentage correct was used as an index of reading comprehension for each grade.

English reading comprehension was assessed with the Grade 1 and Grade 2 Comprehension Subtests of the Gates-MacGinitie Reading Test, Fourth Edition (MacGinitie, et al., 2000), which consisted of 80 short segments of stories and non-fiction passages. After each segment, children were asked to select one picture among three choices that best represented the sentence for each question. Each correct response received one point, and the maximum score for English reading comprehension was 80 for both Grades 2 and 3. Test-retest reliability coefficients for an approximately 12-month interval were .80 for Chinese and .89 for English.

Nonverbal intelligence. Nonverbal IQ was measured in Grade 2 with the subtest of matrix reasoning from the Wechsler Abbreviated Scale of Intelligence, 2nd Edition (Wechsler, 2011). Children were presented with an incomplete matrix and asked to select the best option to complete the matrix. Cronbach's alpha was .83.

2.3. Procedure

We recruited our participants across Hong Kong by first sending an invitation letter and information leaflet to randomly selected local mainstream primary schools. Each school distributed a written consent form and language background questionnaire to parents of participating children. Once written consent was obtained, our participants were either tested in a quiet classroom at their school or in the University laboratory. Chinese and English reading comprehension was group tested, while all other tasks were individually administered by well-trained, proficient Cantonese-English bilingual research assistants. Testing took approximately one hour with 15-20 minutes allocated to acoustic recording of prosodic reading and 35-40 minutes to assessment of reading outcomes.

2.4. Acoustic analyses of Chinese and English prosody reading across grades 2 and 3

To characterize the acoustic features of each child's L1 and L2 prosodic reading of six syntactically complex structures, we used Praat (Boersma & Weenink, 2015) to conduct a spectrographic analysis of fundamental frequency ($f0$) and pause durations extracted from each type of syntactic structure in the read-aloud passages (see Figures 2a and 2b).

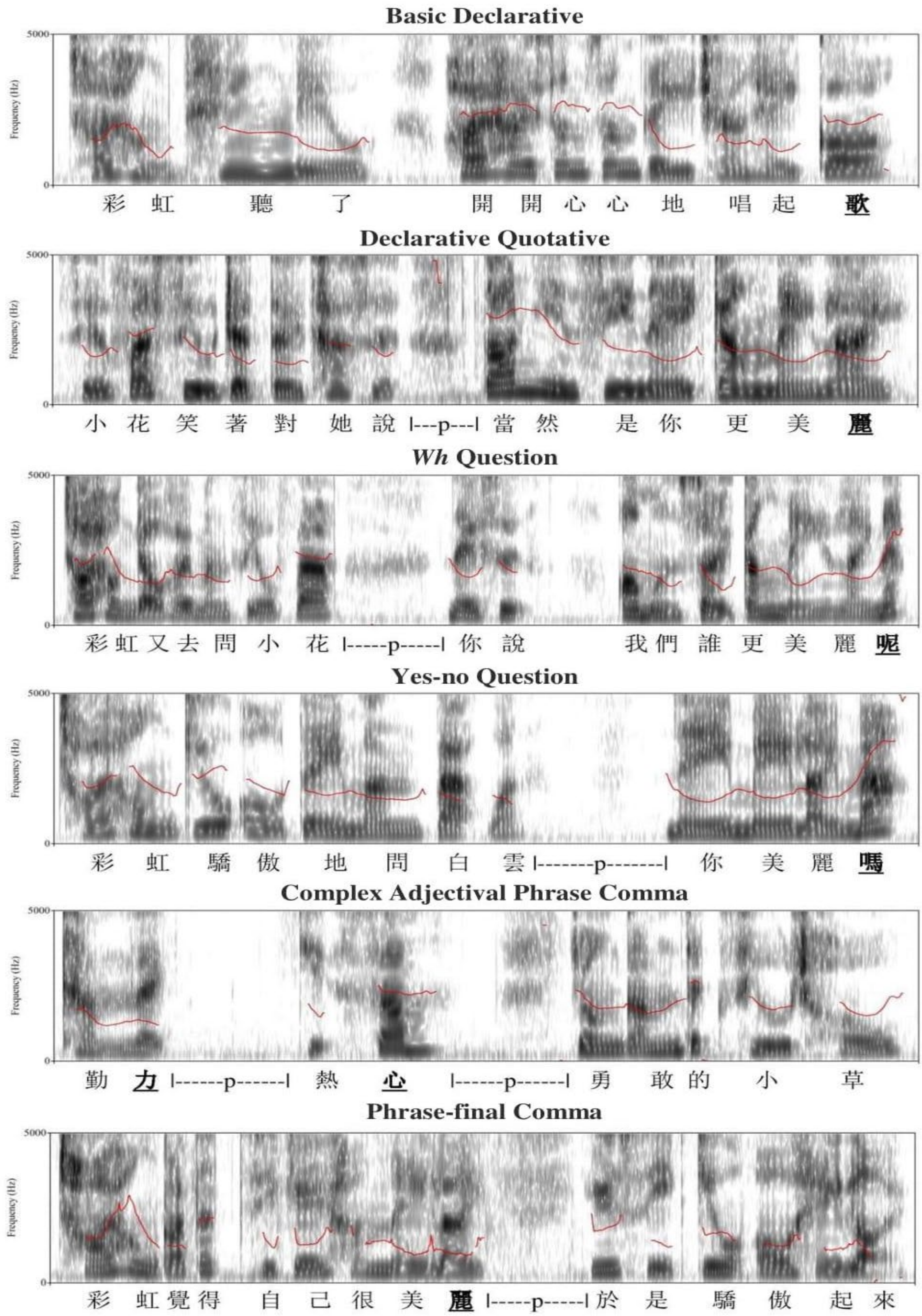


*Figure 2a. An example of the six targeted structures in the Chinese read-aloud passage. (**Words** = pitch pattern. |-p-| = pause.)*

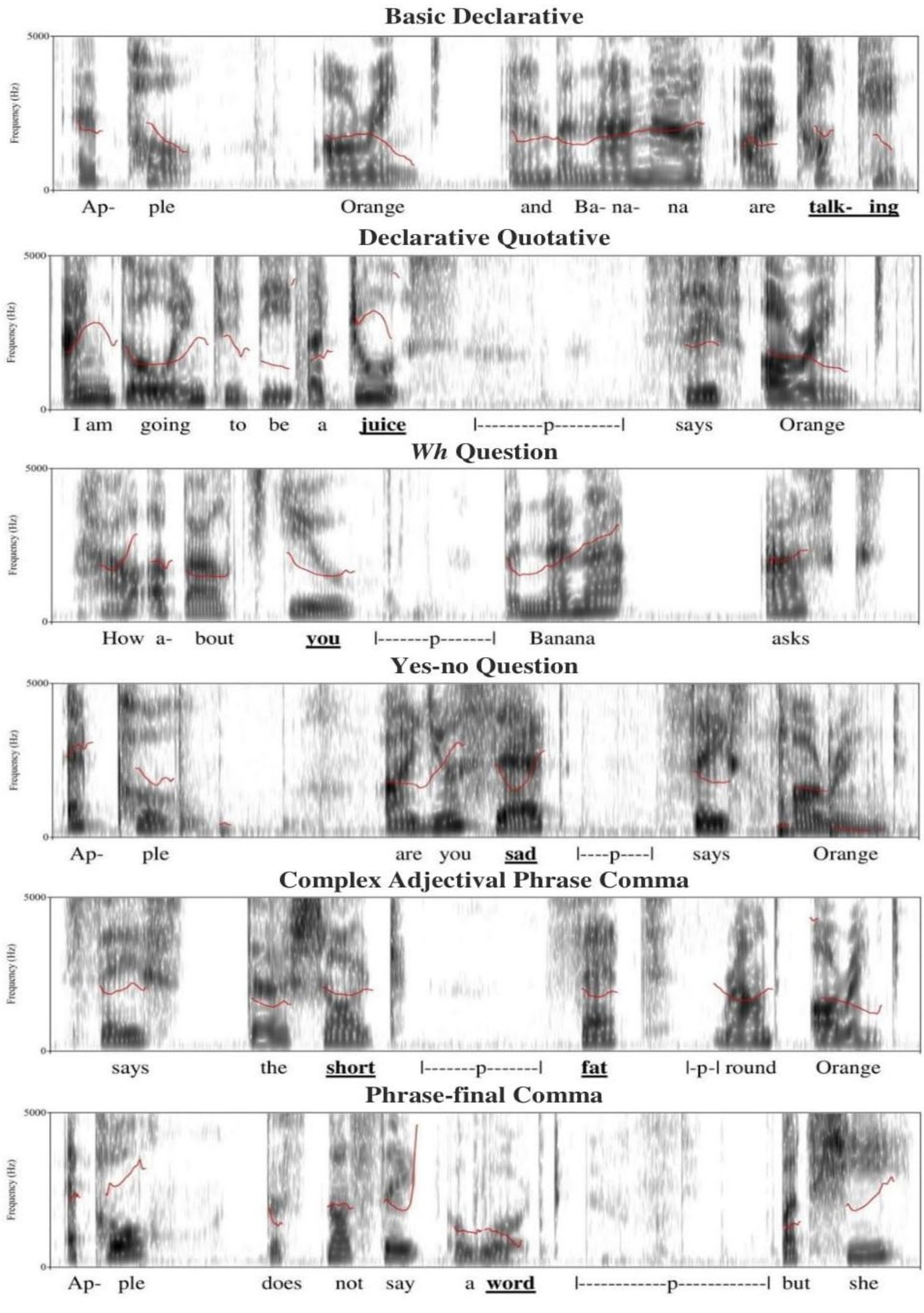


Figure 2b. An example of the six targeted structures in the English read-aloud passage.
***(Words** = pitch pattern. |-p-| = pause.)*

2.3.1. Pitch measurements

Measured in Hertz (Hz), each child's *f0* onset, *f0* offset, maximum *f0*, and minimum *f0* were obtained for each syntactic structure. *f0* values were extracted at the sentence-final word for basic declaratives, yes-no questions, *wh* questions, and basic quotatives, and before the comma for phrase-final commas and complex adjectival phrases. *f*0 onset corresponded to the onset of periodicity in the waveforms of the vocalic segment in the target words, and was extracted at the first stable upward-going zero-crossing of the voicing cycle. *f*0 offset was obtained at the beginning of the last periodic cycle as seen on the waveform, and corresponded to the loss of F2 and/or F1 as seen on the spectrographic display (Khouw & Ciocca, 2007). The direction of pitch contour was determined by *f0* onset and offset. Pitch fell when *f0* onset exceeded *f0* offset, but rose when *f0* offset exceeded *f0* onset. Maximum *f*0 and minimum *f*0 were then obtained, with the slope value of pitch contour determined by maximum *f0* minus minimum *f0*.

These four *f*0 parameters were then converted to *equivalent rectangular bandwidth* (ERB) rate via this formula (Glasberg & Moore, 1990):

$$Freq(ERB) = [21.4 * \log_{10}(0.00437 * Freq(Hz) + 1)]$$

In the equation above, the constants 21.4 and 0.00437 are a direct result of the formula derivation process and ERB, which is a measure used in psychoacoustics to approximate the frequency selectivity in the human auditory system. The frequency expressed in Hz represents a physical scale of the number of periods per time unit, and the ERB scale correlates strongly with the frequency selectivity of our auditory system (Hermes & van Gestel, 1991). The transformation of the fundamental frequency (*f0*) parameters from Hz to ERB-rate therefore yields a psychoacoustic measurement that accurately represents the perceptual value of the physical acoustic data (Glasberg & Moore, 1990). Moreover, this ERB measure allows for a more meaningful and perceptually-realistic comparison across speakers and/or genders (Ladd & Terken, 1995). The constant 21.4 scales the logarithmic function appropriately, while the constant 0.00437 scales the frequency term suitable for logarithmic calculation. Together, these constants ensure that the ERB equation accurately quantifies the rate at which auditory events are resolved by the human auditory system.

2.3.2. Pause duration measurements

In accordance with the Gray Oral Reading Tests (Wiederholt & Bryant, 2001), only pause durations between 100 and 3,000 milliseconds (ms) were included. Pause onset and pause offset were obtained at zero-crossings of waveforms based on visual inspection of waveforms and spectrograms. Pause onset marked the beginning of the last cycle of the preceding syllable (corresponding to the offset of the preceding syllable), whereas pause offset marked the beginning of the first cycle of the following syllable (corresponding to the onset of the following syllable).

Basic declarative pause durations were determined by the intervals between the offset of the final syllable of the declarative sentence and the onset of the first syllable of the following sentence. For phrase-final commas and complex adjectival phrase commas, pause occurred from the end of the word before the comma to the following word. Since quotation tags must come in front of the quotation in Cantonese, but can occur at the end of the quotation in English, pause durations for yes-no questions, *wh* questions, and basic quotatives in Cantonese were extracted between the end of the quotation tag and the beginning of the following quotation; for English, these pause intervals were extracted between the sentence-final word and the beginning of the tag that followed the end of the quotation.

All prosodic production recordings were spectrographically analysed, coded, and double-coded by the second and third authors who have extensive experience in acoustic analysis. To evaluate consistencies between the two coders, approximately 25% of randomly selected, double-coded acoustic data from each phase was submitted to a two-way random-effect model based on a single measurement and absolute agreement of the intraclass correlation coefficient (ICC) for assessing the inter-rater reliability of each pitch and pause variable. All coefficients were between fair and excellent (see Table 2).

2.3. Statistical analysis

As the number of participants in Grade 2 (N = 121) and Grade 3 (N = 52) were different, we separately ran analyses of 121 second graders and 52 third graders. There were no missing data within each grade. All continuous explanatory variables were centered.

2.3.1. Analytical issues and statistical strategies

Our analyses of these data addressed issues involving measurement error, multiple outcomes, and explanatory variables. To reduce measurement error, we used multiple questions for each construct to create a precise index, then determined via *confirmatory factor analysis* (CFA, Joreskog & Sorbom, 2018) whether sets of pitch or pause variables reflected one or more underlying constructs. To assess CFA fit, we used the comparative fit index (CFI), Tucker–Lewis index (TLI), root mean square error approximation (RMSEA), and standardized root mean square residual (SRMR) at two fit thresholds: good (CFI & TLI > .95; RMSEA < .06; SRMR < .08) and moderate (.90 < CFI & TLI < .95; .06 < RMSEA < .10; .08 < SRMR < .10) (Hu & Bentler, 1999).

Multiple outcomes can have correlated residuals, so modeling each one separately can inflate standard errors. To address this issue, we can use *multivariate outcome, multilevel analysis* for observed variables (Hox et al., 2017) or *structural equation model* for latent constructs (Joreskog & Sorbom, 2018). As our EFA and CFA did not yield coherent factors, we used observed outcome variables and a *multivariate outcome, multilevel analysis* (see supplementary Table S4 for a summary of CFA results).

Explanatory variable issues included: 1) categorical explanatory variables (grade, language) modelled with *deviation coding* (a type of *contrast coding*; Schad et al., 2020): 2) indirect mediation, which was determined using an *M-test* (MacKinnon et al., 2004); 3) false positives across many hypotheses, which were reduced using the *two-stage linear step-up procedure* (Benjamini et al., 2006); 4) effect size comparisons, which were evaluated via *Lagrange multiplier tests* (Bertsekas, 2014); and 5) stability of results despite minor differences in the data or analyses (*robustness*, Kennedy, 2008), such as modelling outcomes simultaneously and then separately.

2.3.2. Explanatory model

To examine our first and second questions regarding language-independent similarities and language-specific differences and their developmental changes in prosodic reading, we tested whether pitch or pause (factors or variables) differed across grades (2 vs. 3) or languages (Chinese vs. English) via *multivariate outcome, multilevel analysis* via *MLwin* software (Charlton et al., 2020).

$$\mathbf{Prosody}_{yi} = \beta_y + e_{yi} + \beta_{y1}\mathit{Grade\_3}_{yi} + \beta_{y2}\mathit{English}_{yi} + \beta_{y3}\mathit{Grade\_3{*}English}_{yi} \quad (1)$$

In the **Prosody** vector, outcome y (pitch or pause variables) of student i has grand mean intercept β_y, with unexplained component (*residual*) e_{yi}. The pitch and pause variables are changes and duration of our six sentence structures. In the first model, we entered the deviation coded variables *Grade*_3 (0.5; vs. -0.5 for Grade 2) and *English* (0.5; vs. -0.5 for Cantonese). In the second model, we added their interaction.

Similarly, we tested whether Grade 2 reading comprehension in English or Chinese was related to IQ and concurrent Grade 2 explanatory variables (i.e., word reading efficiency, passage reading fluency, same-language prosody, or cross-language prosody).

$$\begin{aligned}\mathbf{Grade\_2\_Reading\_Comprehension}_{yi} \\ &= \beta_y + e_{yi} + \beta_{y1}\mathit{IQ}_{yi} + \beta_{yu}\mathbf{Same\_language\_reading}_{yi} \\ &+ \beta_{yv}\mathbf{Same\_language\_prosody}_{yi} + \beta_{yw}\mathbf{Cross\_language\_reading}_{yi} \\ &+ \beta_{yz}\mathbf{Cross\_language\_prosody}_{yi} \quad (2)\end{aligned}$$

In the **Grade_2_Reading_Comprehension** vector, outcome *y* (*Grade 2 Chinese reading comprehension* or *Grade 2 English reading comprehension*) of student *i* had grand mean intercept β_y, with residual e_{yi}. We entered IQ, followed by word reading efficiency and passage reading fluency variables in the same language (**Same_language_reading**). (The subscripts u, v, w, and z reflect the different numbers of explanatory variables in each vector.)

To address our third question regarding concurrent within-language contributions of prosodic reading to reading comprehension, we added prosody variables from the same language (**Same_language_prosody**), followed by word reading and passage reading variables in the other language (**Cross_language_reading**). To address our fourth question regarding concurrent cross-language statistical contributions of prosodic reading to reading comprehension, we added prosody variables from the other language (**Cross_language_prosody**). To increase precision and reduce *multicollinearity*, *non-significant* variables were removed without causing *omitted variable bias* (Kennedy, 2008).

Likewise, we tested whether Grade 3 reading comprehension in English or Chinese was related to IQ and concurrent Grade 3 explanatory variables (i.e., word reading efficiency, passage reading fluency, same-language prosody, or cross-language prosody).

$$\begin{aligned}\textbf{Grade\_3\_Reading\_Comprehension}_{yi} &= \beta_y + e_{yi} + \beta_{y1}IQ_{yi} + \beta_{yu}\textbf{Same\_language\_reading}_{yi} \\ &+ \beta_{yv}\textbf{Same\_language\_prosody}_{yi} + \beta_{yw}\textbf{Cross\_language\_reading}_{yi} \\ &+ \beta_{yz}\textbf{Cross\_language\_prosody}_{yi} \qquad (3)\end{aligned}$$

In the **Grade_3_Reading_Comprehension** vector, outcome *y* (*Grade 3 Chinese reading comprehension* or *Grade 3 English reading comprehension*) of student *i* had grand mean intercept β_y, with residual e_{yi}. All other vectors and variables had Grade 3 values, and the procedure was the same as equation (2).

Finally, to address our third and fourth questions regarding **longitudinal** within- and cross-language statistical contributions of prosody to reading comprehension, we tested whether Grade 3 reading comprehension in English or Chinese was related to IQ and Grade 2 explanatory variables (i.e., word reading efficiency, passage reading fluency, same language prosody, or cross language prosody).

$$\begin{aligned}\mathbf{Longitudinal\_Reading\_Comprehension}_{yi} \\ = \beta_{y} + e_{yi} + \beta_{y1} IQ_{yi} + \beta_{yu}\mathbf{Same\_language\_reading}_{yi} \\ + \beta_{yv}\mathbf{Same\_language\_prosody}_{yi} + \beta_{yw}\mathbf{Cross\_language\_reading}_{yi} \\ + \beta_{yz}\mathbf{Cross\_language\_prosody}_{yi} \qquad (4)\end{aligned}$$

In the **Longitudinal_Reading_Comprehension** vector*,* outcome y (*Grade 3 Chinese reading comprehension* or *Grade 3 English reading comprehension*) of student i had grand mean intercept β_y, with residual e_{yi}. All other vectors and variables had Grade 2 values, and the procedure was the same as equation (2).

We computed the Schwarz Bayesian information criterion (BIC) for each set of variables to find the optimal model with the lowest BIC (Grasa, 1989). Information criteria measure whether a model strikes a good balance between goodness of fit and a parsimonious specification. Unlike other information criteria, BIC provides consistent estimates regardless of the sample size (Grasa, 1989).

3. Results

Table 2 shows the means, standard deviations of reading outcome measures, and pitch and pause features produced by children in Grades 2 and 3 as they read aloud six syntactic types of sentences in connected Chinese and English passages (the other descriptive statistics, including maximum, minimum and range, are provided in supplementary Table S1). Correlations of all pitch and pause variables across languages are provided in supplementary Tables S2 and S3.

Although Miller and Schwanenflugel's (2006) study of English monolingual children's prosodic reading indicated a two-factor solution, with pitch and pause emerging as distinct factors, our CFAs did not support a single factor for either pitches or pauses (see summary of CFA results in supplementary Table S4). Furthermore, follow-up exploratory factor analyses (EFAs) showed low ratios of adjacent eigenvalues (less than 2) for pitches and for pauses, indicating an absence of latent constructs (Tabacknick & Fidell, 2012) (see EFA results in supplementary Table S5).

Table 2

Descriptive Statistics and Inter-Rater Reliability (r_{xx}) for Prosodic Reading of Six Syntactic Structures and Reading Skills in L1 Chinese and L2 English for Children in Grades 2 (G2) and 3 (G3)

Measure	Grade 2 Cantonese (n = 121)			Grade 2 English (n = 121)			Grade 3 Cantonese (n = 52)			Grade 3 English (n = 52)		
	M	*SD*	r_{xx}	*M*	*SD*	r_{xx}	*M*	*SD*	r_{xx}	*M*	*SD*	r_{xx}
Basic declarative												
Pitch (ERB-rate)	-0.79	0.94	.83	-2.54	1.18	.85	-0.72	0.70	.89	-2.00	1.44	.96
Pause (ms)	968	400	.79	851	427	.98	871	315	1.0	681	286	.99
Declarative quotative												
Pitch (ERB-rate)	-0.85	0.87	.85	-1.85	1.24	.78	-0.68	-0.99	.87	-1.61	1.38	.76
Pause (ms)	599	263	.99	588	397	.98	571	173	.99	422	272	.99
Wh question												
Pitch (ERB-rate)	1.03	0.69	.79	-1.47	2.06	.92	1.04	0.62	.71	-1.64	1.73	.84
Pause (ms)	743	284	.99	574	349	1.0	663	199	1.0	409	229	.98
Yes-no question												
Pitch (ERB-rate)	0.92	0.68	.93	0.69	1.45	.86	0.81	0.63	.79	.53	1.29	.96
Pause (ms)	623	251	.86	764	538	.93	594	209	.97	615	315	1.0
Complex adjectival phrase Comma												
Pitch (ERB-rate)	-0.31	0.55	.78	-0.42	0.86	.99	-0.41	0.55	.83	-0.27	0.64	.96
Pause (ms)	437	243	.88	470	369	.99	365	145	.89	275	258	.98
Phrase-final comma												
Pitch (ERB-rate)	-0.54	0.79	.75	-1.38	1.54	.93	-0.66	0.72	.79	0.39	2.04	.75
Pause (ms)	537	284	.99	826	366	1.0	461	181	.99	624	297	.99
Reading skills												
Word reading efficiency	63.45	17.60		41.26	15.13		76.10	16.52		56.31	14.41	
Reading fluency	124.93	26.75		70.40	21.76		141.50	22.06		86.56	14.70	
Reading comprehension	64.26	16.33		55.41	18.99		72.06	14.13		76.13	17.18	

3.1. Research questions 1 and 2: Language-independent similarities and language-specific differences in L1 Chinese and L2 English prosodic reading development

Figures 3 and 4 illustrate the cross-language similarities and differences in the six pitch and six pause measures between L1 Chinese and L2 English across Grades 2 and 3. Table 4 shows the results of multivariate outcome, multilevel analysi*s* tested for differences in each pitch and pause measure across both grades.

3.1.1. Pitch

Basic declarative pitch falls were often smaller for children in Grade 2 than in Grade 3 (b = 0.306; see Table 3, panel A, top, left). Meanwhile, basic declarative pitch falls were often greater in English than Cantonese (b = -1.609). Together, they accounted for over 36% of the variance in basic declarative pitch (see Table 3, panel A, bottom, left). Similarly, declarative quotative pitch falls were often greater in English than Cantonese (b = -0.982), accounting for 17% of the variance. *Wh* question pitch fall was greater in English than in Cantonese (b = -2.560), accounting for over 43% of the variance. Figure 3 illustrates the results showing frequent pitch falls in English, but pitch rises in Cantonese.

Analyses of three other pitch measures revealed no robust solution. Yes-no question pitch rises were greater in Cantonese than in English (b = -0.245), which aligned with the lowest pitch rise in Grade 3 English compared to all other grade and language combinations. However, the large standard deviation in Grade 3 English indicates caution when interpreting this result. This model accounted for less than 2% of the variance. The analysis of complex adjectival phrase comma pitch revealed that language and grade were not significant. Phrase-final comma pitch falls were often smaller for children in Grade 2 than in Grade 3 (b = 0.826) with larger differences at the interaction with language (b = .943). These explanatory variables accounted for nearly 17% of the variance. However, the large standard deviation in Grade 3 English (SD = 2.04) far exceeds its mean (M = 0.39), which suggests caution when interpreting this result.

3.1.2. Pause

Basic declarative pause was shorter in Grade 3 than in Grade 2 (b = -0.134) and shorter in English than in Cantonese (b = -0.139; see Table 3, panel B, top, left, and Figure 4), accounting for nearly 6% of the variance (see Table 3, panel B, bottom, left). Declarative quotative pause was also shorter in Grade 3 than in Grade 2 (b = -0.096) and in Cantonese than in English among Grade 2 students, (b = -0.081; see model 2), accounting for nearly 4% of the variance.

Wh question pause was shorter in Grade 3 than in Grade 2 (b = -0.122) and shorter in English than in Cantonese (b = -0.195), accounting for over 13% of the variance. Yes-no question pause was shorter in Grade 3 than in Grade 2 (b = -0.092) and longer in English than in Cantonese (b = 0.101), accounting for nearly 3% of the variance. Complex adjectival phrase comma pause was shorter in Grade 3 than Grade 2 (b = -0.133), accounting for over 4% of the variance. Phrase-final pause was shorter in Grade 3 than in Grade 2 (b = -0.139) and longer in English than in Cantonese (b = 0.251), accounting for nearly 3% of the variance.

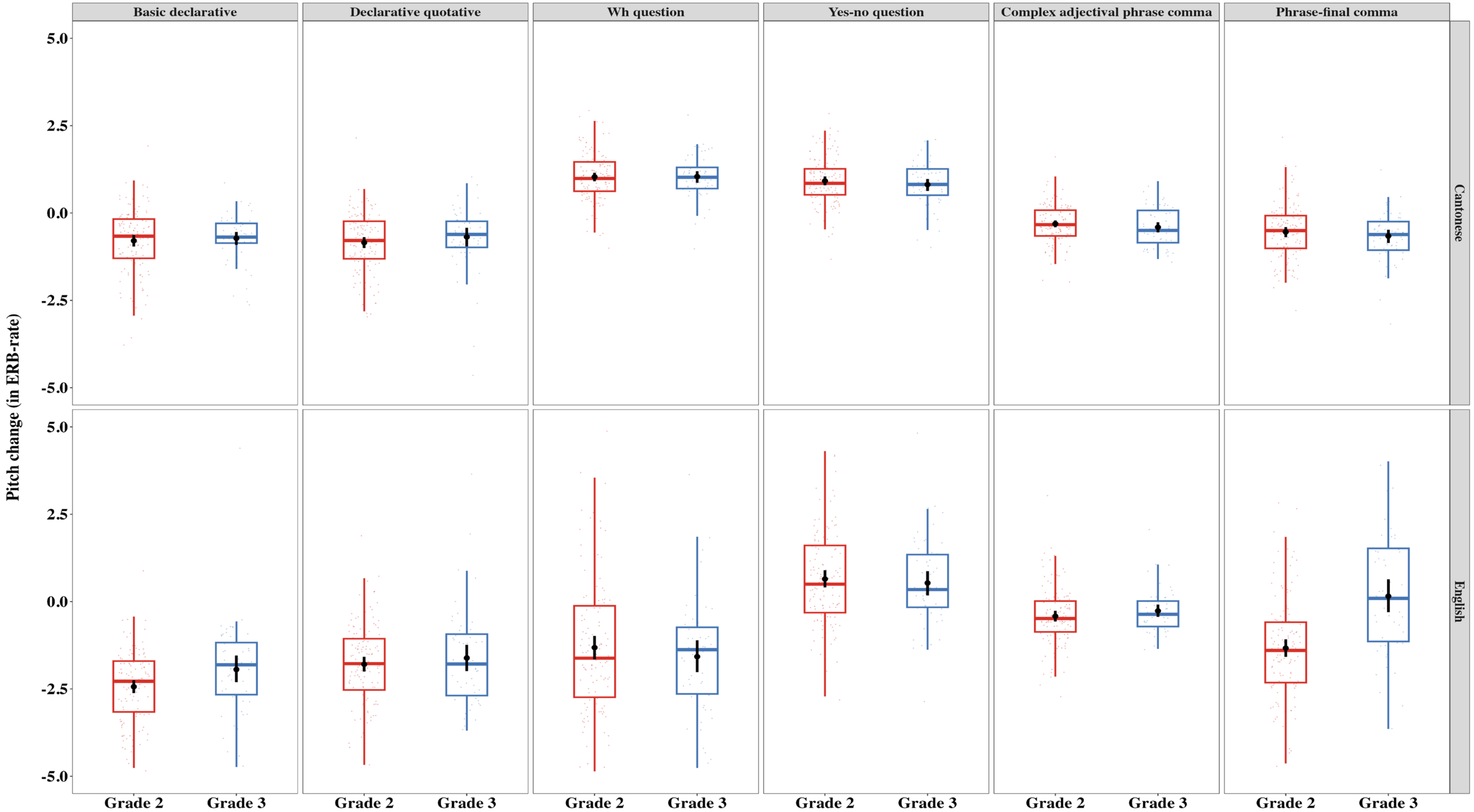


Figure 3. Box plots representing average pitch values (in ERB-rate) for six Cantonese and English prosodic structures produced by Cantonese-English bilingual children in grade 2 and grade 3. Error bars represent 95% confidence intervals. Positive values represent pitch rise, whereas negative values represent pitch fall

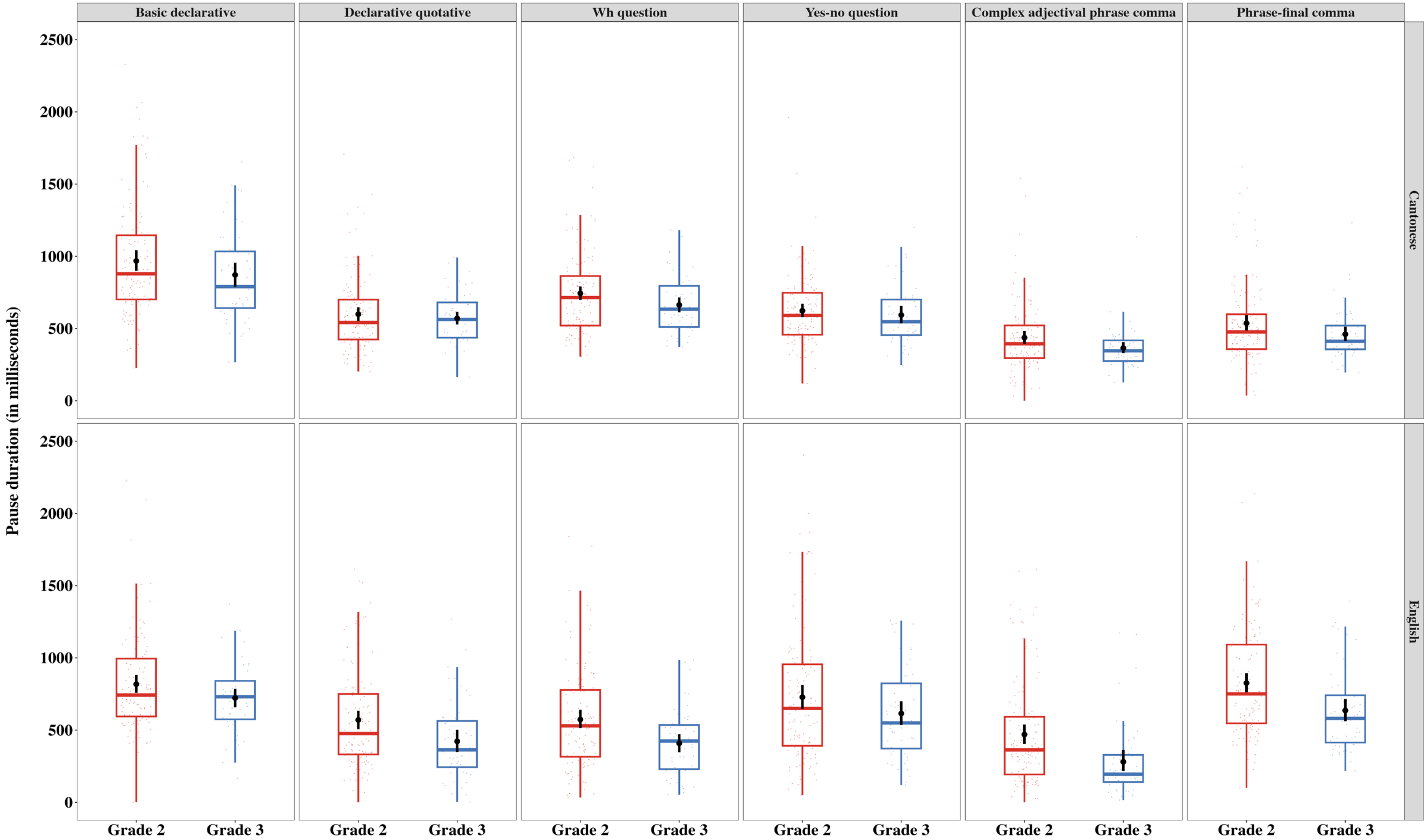


Figure 4. Box plots representing average pause values (in milliseconds) for six Cantonese and English prosodic structures produced by Cantonese-English bilingual children in grade 2 and grade 3. Dots represent the mean. Error bars represent 95% confidence intervals.

Table 3

Results of Multivariate Outcome Multilevel Analysis for Six Pitches and Six Pauses with Unstandardized Regression Coefficients and (Standard Errors)

Panel A	Pitch											
	Basic declarative		Declarative quotative		*Wh* question		Yes-no question		Complex adjectival phrase comma		Phrase-final comma	
Explanatory variable	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Grade 3	0.306*	0.306*	0.201	0.201	-0.081	-0.081	-0.130	-0.130	0.025	0.027	0.820***	0.826***
	(0.128)	(0.127)	(0.130)	(0.130)	(0.172)	(0.172)	(0.129)	(0.129)	(0.081)	(0.081)	(0.163)	(0.155)
English	-1.609***	-1.515***	-0.982***	-0.967 ***	-2.560***	-2.596***	-0.245*	-0.255*	-0.036	0.015	-0.277	0.103
	(0.117)	(0.127)	(0.119)	(0.130)	(0.158)	(0.172)	(0.118)	(0.127)	(0.074)	(0.081)	(0.149)	(0.155)
Grade 3 × English		0.234		0.039		-0.089		-0.026		0.125		0.943***
		(0.127)		(0.130)		(0.172)		(0.129)		(0.081)		(0.155)
Variance at each level												
School	0%		5%		0%		1%		0%		3%	
Student	100%		95%		100%		99%		100%		97%	
Explained variance at each level												
School	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.553	0.561
Student	0.361	0.367	0.179	0.180	0.435	0.435	0.015	0.015	0.001	0.008	0.063	0.156
Total variance explained	0.361	0.367	0.170	0.171	0.435	0.435	0.015	0.015	0.001	0.008	0.078	0.168
BIC	3.087	3.094	3.096	3.113	3.679	3.695	3.091	3.108	2.167	2.177	3.564	3.476

Note: Equation (1) $\textbf{Prosody}_{yi} = \beta_y + e_{yi} + \beta_{y1}\mathit{Grade}_{yi} + \beta_{y2}\mathit{English}_{yi} + \beta_{y3}\mathit{Grade{*}English}_{yi}$
Each model includes a constant term. All *random effects* are not significant. BIC = Bayesian Information Criterion
**p* < .05, ***p* < .01, ****p* < .001

Panel B	Pause

Explanatory variable	Basic declarative		Declarative quotative		*Wh* question		Yes-no question		Complex adjectival phrase comma		Phrase-final comma	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Grade 3	-0.134**	-0.134**	-0.096**	-0.096**	-0.122***	-0.122***	-0.092*	-0.092*	-0.133***	-0.133***	-0.139***	-0.139***
	(0.045)	(0.045)	(0.036)	(0.036)	(0.034)	(0.034)	(0.045)	(0.045)	(0.034)	(0.034)	(0.036)	(0.036)
English	-0.139***	-0.154***	-0.055	-0.081*	-0.195***	-0.212***	0.101*	0.079	-0.004	-0.028	0.251***	0.226***
	(0.041)	(0.045)	(0.033)	(0.036)	(0.031)	(0.034)	(0.041)	(0.045)	(0.031)	(0.034)	(0.033)	(0.036)
Grade 3 × English		-0.037		-0.067		-0.043		-0.057		-0.061		-0.063
		(0.045)		(0.036)		(0.034)		(0.045)		(0.034)		(0.036)
Variance at each level												
School	0%		0%		3%		5%		12%		7%	
Student	100%		100%		97%		95%		88%		93%	
Variance at each level												
School	0	0	0	0	0.088	0.089	0.000	0.000	0.211	0.382	0.157	0.291
Student	0.056	0.058	0.028	0.039	0.132	0.136	0.030	0.035	0.021	0.043	0.179	0.180
Total variance explained	0.056	0.058	0.028	0.039	0.131	0.133	0.029	0.034	0.044	0.053	0.176	0.183
BIC	0.995	1.010	4.989	5.000	0.442	0.454	0.974	0.986	0.354	0.360	6.363	6.379

Note: Equation (1) $\mathbf{Prosody}_{yi} = \beta_y + e_{yi} + \beta_{y1} Grade_{yi} + \beta_{y2} English_{yi} + \beta_{y3} Grade*English_{yi}$

Each model includes a constant term. All *random effects* are not significant. BIC = Bayesian Information Criterion

**p* < .05, ***p* < .01, ****p* < .001

3.2. Research Question 3: To what extent does prosodic reading explain variance in reading comprehension skills beyond word reading efficiency and oral reading fluency?

3.2.1. Concurrent Grade 2 and Grade 3 reading comprehension

Table 4 shows the results of multivariate outcome, multilevel analysis of the concurrent relation between prosodic reading and reading comprehension for L1 Chinese and L2 English.

L1 Chinese. Grade 2 children with higher word reading efficiency (b = 0.185) or passage reading fluency (b = 0.009) exhibited better Chinese reading comprehension. However, no Grade 2 prosodic reading variables showed a significant link to Grade 2 reading comprehension. This model accounted for nearly 35% of the variance.

Grade 3 children with higher word reading efficiency (b = 0.295) exhibited greater Chinese reading comprehension. No Grade 3 prosodic variables were linked to Grade 3 reading comprehension. This model accounted for over 35% of the variance.

L2 English. Grade 2 children with higher word reading efficiency (b = 0.569; model 2) exhibited better English reading comprehension (see Table 4, middle, English reading comprehension, Grade 2, middle). However, passage fluency and all English prosody variables in Grade 2 were not significantly linked to Grade 2 English reading comprehension. Grade 2 children with greater Chinese reading comprehension showed better English reading comprehension (b = 0.129). Grade 2 children with greater Cantonese pitch yes-no question rises exhibited better English reading comprehension (b = 0.167). These explanatory variables accounted for nearly 59% of the variance.

As with Grade 2, Grade 3 children with higher word reading efficiency (b = 0.347) exhibited better English reading comprehension and, in addition, passage reading fluency (b = 0.027) was also associated with better English reading comprehension(see Table 4, middle, English reading comprehension, Grade 3, right). In contrast to the Grade 2 analyses, several

Grade 3 prosodic variables were associated with Grade 3 English reading comprehension: A larger pitch fall in English *wh* questions (b = - 0.153), shorter declarative quotative English pauses (b = - 1.439), and shorter phrase-final English pauses (b = -1.181) were associated with better English reading comprehension. Furthermore, Grade 3 quotative English pause mediated the link between English word reading efficiency and English reading comprehension by 20% (z = 2.624; p = .009); specifically, Grade 3 children with higher English word reading efficiency exhibited shorter quotative English pauses and more proficient English reading comprehension.

Table 4

Results of multivariate outcome, multilevel analysis for Concurrent Predictions of Prosody to Chinese Reading Comprehension and to English Reading Comprehension

	Chinese Reading Comprehension	
Explanatory variable	Grade 2	Grade 3
Nonverbal IQ	0.285 ***	0.305 **
	(0.077)	(0.107)
Chinese word reading efficiency	0.185 *	0.295 *
	(0.090)	(0.125)
Chinese passage reading fluency	0.009 **	0.007
	(0.003)	(0.005)
Variance at each level		
School	19%	32%
Student	81%	68%
Explained variance at each level		
School	0.718	0.546
Student	0.264	0.262
Total variance explained	0.348	0.352
BIC	2.604	2.715
	English reading comprehension	
	Grade 2	Grade 3
Nonverbal IQ	0.062	0.164 *
	(0.064)	(0.064)
English word reading efficiency	0.569 ***	0.347 **
	(0.105)	(0.112)
English passage reading fluency	0.002	0.027 ***
	(0.004)	(0.007)
Chinese reading comprehension	0.129 *	
	(0.065)	
English pitch *Wh* question		-0.153 ***
		(0.038)
Log (English Pause Declarative quotative)		1.439 ***
		(0.403)
Log (English Pause Phrase-final comma)		-1.181 **
		(0.408)
Cantonese pitch Yes-no question	0.167 *	
	(0.082)	
Variance at each level		
School	34%	35%
Student	66%	65%
Explained variance at each level		

School	0.904	1.000
Student	0.426	0.733
Total variance explained	0.586	0.828
BIC	2.137	1.848

Note: n = 121 for equation (2), and n = 52 for equation (3)

$$\mathbf{Grade\_2\_Reading\_Comprehension}_{yi} = \beta_y + e_{yi} + \beta_{y1} IQ_{yi} + \beta_{yu}\mathbf{Same\_language\_reading}_{yi} + \beta_{yv}\mathbf{Same\_language\_prosody}_{yi} + \beta_{yw}\mathbf{Cross\_language\_reading}_{yi} + \beta_{yz}\mathbf{Cross\_language\_prosody}_{yi} \quad (2)$$

$$\mathbf{Grade\_3\_Reading\_Comprehension}_{yi} = \beta_y + e_{yi} + \beta_{y1} IQ_{yi} + \beta_{yu}\mathbf{Same\_language\_reading}_{yi} + \beta_{yv}\mathbf{Same\_language\_prosody}_{yi} + \beta_{yw}\mathbf{Cross\_language\_reading}_{yi} + \beta_{yz}\mathbf{Cross\_language\_prosody}_{yi} \quad (3)$$

Each model includes a constant term. All variables that are not shown are not significant. BIC = Bayesian Information Criterion *p < .05, **p < .01, ***p < .001

3.2.2. Longitudinal predictors from Grade 2 prosodic variables to Grade 3 reading comprehension

L1 Chinese. Higher Grade 2 passage fluency (b = 0.012) and larger Grade 2 *wh* question pitch rise (b = 0.263) were associated with better Grade 3 Chinese reading comprehension (see Table 5, top). These explanatory variables accounted for nearly 47% of the variance.

L2 English. Children with higher Grade 2 word reading efficiency (b = 0.641) exhibited better Grade 3 English reading comprehension (see Table 5, middle). Higher Grade 2 Chinese reading comprehension (b = 0.176) and larger Chinese pitch phrase-final comma fall (b = -0.294) predicted better Grade 3 English reading comprehension. This model accounted for over 77% of the variance.

3.3. Research Question 4: Does a cross-language relationship exist between prosodic reading and reading comprehension in Cantonese-English bilingual children?

Our concurrent analyses of L2 English prosody to L1 Chinese reading comprehension in Grades 2 and 3 indicated no significant crossover effects across grades (Table 4, top left). The final models of the multivariate outcome multilevel analyses accounted for 35% of the variances in both Grade 2 and Grade 3 Chinese reading comprehension. In contrast, our concurrent analyses of L1 Cantonese prosody to L2 English reading comprehension indicated that Grade 2 children with better Chinese reading comprehension exhibited better Grade 2 English reading comprehension. More importantly, children with larger pitch rise for Chinese yes-no questions exhibited better English reading comprehension. The final model accounted for nearly 59% of the variance in Grade 2 English reading comprehension. However, for Grade 3 children, neither Chinese reading nor Cantonese prosody was significantly linked to English reading comprehension, with the final model accounting for nearly 83% of the variance.

The results of the longitudinal analyses (see Table 5) also demonstrated that children with better Grade 2 Chinese reading comprehension exhibited better Grade 3 English reading comprehension. Furthermore, the larger Grade 2 Cantonese phrase-final pitch fall was associated with better Grade 3 English reading comprehension. The final model accounted for over 77% of the variance.

Table 5

Results of Multivariate Outcome Multilevel Analysis for Within- and Across-Language Longitudinal Predictions from Grade 2 to Grade 3 for Chinese and English Reading Comprehension (n = 52)

Explanatory variable	Grade 3 Chinese reading comprehension	
Nonverbal IQ	0.206	*
	(0.104)	
Grade 2 Chinese word reading efficiency	0.166	
	(0.126)	
Grade 2 Chinese passage reading fluency	0.012	*
	(0.005)	
Grade 2 Cantonese Pitch *Wh* question	0.263	*
	(0.120)	
Variance at each level		
School (32%)	0.649	
Student (68%)	0.383	
Total variance explained	0.467	
BIC	2.608	
	Grade 3 English reading comprehension	
Nonverbal IQ	0.045	
	(0.085)	
Grade 2 English word reading efficiency	0.641	***
	(0.123)	
Grade 2 English passage reading fluency	0.002	
	(0.006)	
Grade 2 Chinese reading comprehension	0.176	*
	(0.081)	
Grade 2 Chinese pitch Phrase-final comma	-0.294	**
	(0.098)	
Variance at each level		
School (35%)	1.000	
Student (65%)	0.651	
Total variance explained	0.775	
BIC	2.041	

Note. Equation (4)

$$\begin{aligned}\mathbf{Longitudinal\_Reading\_Comprehension}_{yi} &= \beta_y + e_{yi} + \beta_{y1} IQ_{yi} + \beta_{yu}\mathbf{Same\_language\_reading}_{yi} \\ &+ \beta_{yv}\mathbf{Same\_language\_prosody}_{yi} + \beta_{yw}\mathbf{Cross\_language\_reading}_{yi} \\ &+ \beta_{yz}\mathbf{Cross\_language\_prosody}_{yi} \end{aligned} \tag{4}$$

The longitudinal prediction was conducted using nonverbal IQ, reading fluency, and Cantonese prosodic reading variables measured in grade 2 to predict grade 3 English reading comprehension. Each model includes a constant term. All variables that are not shown are not significant. BIC = Bayesian Information Criterion * $p < .05$. ** $p < .01$. *** $p < .001$.

4. Discussion

This study explored the role of prosodic reading in bilingual reading comprehension, specifically for children with tonal Cantonese as their first language and non-tonal English as their second language. Spectrographic analysis of Hong Kong Cantonese-English bilingual children's expressive reading of six types of parallel L1 and L2 syntactically-complex sentences in connected text revealed common cross-linguistic and language-specific patterns for the same syntactic structure. Furthermore, the pitch patterns used by these bilingual children while reading aloud, particularly for *wh* questions, were found to be associated with their reading comprehension across languages. Moreover, their use of short pauses in certain English structures were significant correlators of their greater English reading comprehension. Finally, the pitch patterns used in their L1 Cantonese were linked with L2 English reading comprehension. For example, using higher pitch for yes-no questions and lower pitch for phrase-final commas in Cantonese was linked to better English reading comprehension skills. These findings highlight the importance of prosody in bilingual reading comprehension, support the PCH framework, and extends previous research on prosodic reading in monolingual children (e.g., Chung & Bidelman, 2022; Miller & Schwanenflugel, 2006) by focusing on Cantonese-English bilingual children.

In relation to pitch, this study has several notable findings. First, we found that bilingual children are sensitive to pitch variations in both L1 Chinese and L2 English, displaying a pitch fall in declarative sentences and a pitch rise in yes-no questions, as well as increased pitch movements before major syntactic boundaries. More importantly, our findings demonstrate for the first time that these bilingual children showed language-specific differences in pitch structures when producing different question types. Specifically, English yes-no questions had a large pitch rise, while English *wh* questions did not maintain an upswing pitch, aligning with patterns in monolingual English children (Miller &

Schwanenflugel, 2008). In Cantonese, both question types were uttered with a rise in pitch, consistent with previous research (e.g., Ma et al., 2006). This suggests that these bilingual children recognize language and question type differences in pitch early in reading development. Additionally, greater pitch declination in English basic declarative, declarative quotative, and *wh* questions compared to Cantonese reflects language-specific constraints on prosodic reading. This may explain why our participants exhibited stronger English prosodic reading ability, as Cantonese has limited application of pitch at the utterance level (Fox et al., 2008; Ladd, 1996). Additionally, our findings suggest the development of prosodic reading from Grade 2 to 3, as indicated by smaller pitch fall for basic declaratives in Grade 3 and shorter pauses across various structures compared to Grade 2. These results suggest that bilingual children demonstrate an increased ability to mark pitch declination, as well as a heightened level of reading fluency.

There were also several notable findings in relation to pause duration. First we found that our participants exhibited proficiency in recognizing and utilizing pauses as boundary markers within and between sentences in both languages but, critically, exhibited language-specific differences in pause duration. They produced longer pauses after yes-no questions, declarative quotative, and phrase-final comma in English compared to Cantonese, but shorter pauses after *wh* questions and basic declarative. These longer pauses in English contrast with the shorter pauses typically observed in skilled English readers after basic declarative sentence, basic quotative, yes-no questions and phrase-final comma (Miller & Schwanenflugel, 2006). This suggests that our participants, with their stronger Chinese and weaker-English profiles, may be less proficient in utilizing pause structures in L2 English than in L1 Cantonese. However, the shorter pause observed in *wh* questions and basic declarative may indicate that our participants were influenced by English-specific pause structures.

How do these bilingual children's utilization of pitch and pause patterns when reading aloud syntactically complex sentences relate to their ability to comprehend the text, both within and across languages? Our within-language results showed that pitch, and especially *wh* question pitch contour, and certain pause structures were significantly associated with reading comprehension. These findings cannot be explained by current theories of reading comprehension, but are identified by the PCH framework, which characterizes prosodic reading as a process involving parallel word decoding (i.e., converting written text into spoken elements) and linguistic encoding (i.e., assigning grammatical roles to spoken elements).

To achieve appropriate prosody in sentence reading, words must be assigned proper syntactic roles, while fluent reading necessitates suitable phrasing and chunking of its syntactic structure (Chafe, 1998; Schreiber, 1991). As *wh* questions are syntactically complex, many children, including typically developing ones, struggle with producing and comprehending them across languages (e.g., van der Lely et al., 2011; Schulz & Roeper, 2011). Thus, our participants' appropriate use of rising (for Cantonese) and falling (for English) contours to read aloud *wh* questions indicates their tacit understanding of the language-specific manifestations of syntactical structures of *wh* questions in these two languages, aiding comprehension. This explanation is supported by the *wh* question's unique status regarding its syntactic complexity and pragmatic function.

In Cantonese, although *wh*-words are in situ in syntax, they do not necessarily map onto their sentence structure position, as their placement is influenced by selection restrictions of certain verbs (Huang, 1982; Shi, 1994). Depending on the verb, some complement clauses containing a wh-word can only be a statement (e.g., 李四問我[邊個郁咗嗰塊芝士]。Lei5sei3 man6 ngo5 [bin1go3 juk1 zo2 go2 faai3 zi1si2]. *Leisei asked me* *who moved that piece of cheese.*) or a question (e.g., 李四相信[邊個郁咗嗰塊芝士]? Lei5sei3 soeng1seon3 [bin1go3 juk1 zo2 go2 faai3 zi1si2]? *Who does Leisei believe moved that piece of cheese*?), while others (e.g., 李四知道[邊個郁咗嗰塊芝士]? Lei5sei3 zi1dou3 [bin1go3 juk1 zo2 go2 faai3 zi1si2].). Others can be either a question (i.e., *Who does Leisei know moved that piece of cheese?*) or a statement (i.e., *Leisei knows who moved that piece of cheese.*) contingent on the context. As wh-words in Cantonese do not overly mark questions, assigning rising pitch contour to a *wh* question may reflect children's covert understanding of *wh*-in-situ syntax, enabling them to make predictions during text comprehension.

Similarly, in English, *wh* questions are syntactically complex due to subject-object asymmetry. The two types of *wh* questions, that is wh-subject questions (wh-words are subjects: e.g,, *who* is watching the boy?) and wh-object questions (e.g., who-words are objects: e.g., *who* is the boy watching?), exhibit different acquisition order; wh-object questions are more difficult to acquire because of grammatical movement and lower input frequency (Fahn, 2003; Rowland et al, 2003; O'Grady, 1997). Furthermore, a corpus study of 200 spontaneous *wh* questions in American English showed that falling pitch contour occurs more frequently as most *wh* questions serve a linguistic function, such as requesting elaborating details, opening a subtopic, and directing information flow (Hedberg et al., 2010).

In our study, one of the two *wh* questions in the English passages of our prosodic reading task directed information flow (e.g., *while Banana and Orange are talking about what they want to be, Banana asks Apple, "How about you?"*), while the other requested elaborating details (e.g., *when Orange asks Apple whether she is sad, and Apple responds, "Yes" and starts crying, Orange follows up with, "Why are you crying?"*). Children who successfully produced a falling pitch contour for these two questions demonstrated their ability to encode the linguistic functions of the *wh* questions during prosodic reading, which, in turn, indicates their understanding of the text's meaning.

In addition, shorter pauses in English, specifically in declarative quotatives and phrase-final commas were linked to greater Grade 3 English reading comprehension. Grade 3 children with higher English word reading efficiency also exhibited shorter quotative English pauses and stronger English reading comprehension. These results differ from previous studies of native English-speaking children (e.g., Miller & Schwanenflugel, 2006), which found no association between pause duration and English reading comprehension. This discrepancy may be due to the fact our participants are Cantonese-dominant Chinese-English bilingual children whose English skills had yet to reach native proficiency level. Therefore, the shorter pauses suggest improved English fluency, allowing them to focus more on understanding the meaning of the text.

Nevertheless, it should be noted that, except for *wh* pitch contour and some pause structures, most prosodic reading variables did not directly associate with either Chinese or English reading comprehension. These findings can be explained by PCH, which suggests that the role of prosodic reading in reading comprehension can be manifested in either direct and explicit or indirect and implicit ways. Some prosodic measures, like *wh* pitch contours, are directly linked with reading comprehension in both Chinese and English. However, other prosodic variables may influence reading comprehension through their relationship with other linguistic and cognitive constructs, such as syntactic structure and word reading. This explanation aligns with a recent meta-analysis showing that prosody contributed to reading comprehension through metalinguistic awareness and word reading (e.g., Tong et al., 2023). Thus, the lack of direct association between certain prosodic variables and reading comprehension emphasizes that the contribution of prosodic reading comprehension can be manifested through other variables that are directly related to reading comprehension, such as syntactic structure and word reading. Further research is needed to test this hypothesis.

Regarding cross-language associations, we found that L1 yes-no question pitch rise and phrase-final comma pitch fall were associated with L2 English reading comprehension, even after considering key predictors such as L1 and L2 word reading efficiency and text reading fluency. However, no L2 English prosodic variable was linked to L1 Chinese reading comprehension. The language interdependence hypothesis and the threshold hypothesis (Cummins, 2000) suggest that differences in proficiency levels between L1 and L2 proficiency determines the direction of cross-language association. However, these theories cannot explain why certain L1 prosodic cues are transferable to L2 English reading comprehension. In contrast, PCH frames comprehension as a data-prediction and optimization problem whereby comprehenders use implicit prosody or covert imitation to predict words, syntactic structure, and meanings. Meanwhile, prosodic reading provides feedback to reinforce this process by facilitating the transfer of the prosodic cues or parameters that best match the prediction.

Since pitch rise in yes-no questions and pitch fall in phrase-final commas are common in both Cantonese and English, our data can be taken to show that Cantonese-English children can rely on their existing L1 pitch cues to understand the syntactic structures in English. By contrast, excessive pitch variations in English prosodic reading cannot be fully applied to L1 Cantonese, where pitch variation is more limited. As a result, some L1 prosodic variables may contribute to English reading comprehension, while their L2 prosodic counterparts are not associated with Chinese reading comprehension. Although existing reading theories cannot adequately explain our findings outlined above, the PCH provides insights into the roles of pitch and pause cues in reading comprehension. Unlike previous models (e.g., Hoover & Gough, 1990; Perfetti & Stafura, 2014), PCH acknowledges the link between language production and text reading comprehension. It views prosodic reading as covert imitation generated by the production system, facilitating the connections between prosody, syntax, and words. This reinforces the dynamic production-comprehension feedforward and feedback loop, optimizing linguistic and cognitive resources for comprehension. Further empirical research is needed to test this hypothesis. Furthermore, our results show that, while cross-language facilitation occurred in pitch cues, the direction of such facilitation was from L1 Chinese to L2 English reading comprehension, but not vice versa. This highlights the interdependence of prosodic cues between L1 and L2, and the impacts of oral language proficiency on bilingual reading comprehension.

Quantifying reading fluency remains an unresolved issue in the field (Kuhn et al., 2010). Our spectrographic approach objectively assesses oral reading fluency, providing detailed analysis of pitch and pause in various syntactic structures. This can complement the traditional rating-scale-based classroom assessment, leading to a precise characterization of individual oral reading profiles.

Finally, although our study cannot establish causal links, the findings have implications for educational practice and instructional strategies to support bilingual children's reading comprehension. Prosodic reading, including pitch and pause patterns, is important for both L1 Chinese and L2 English reading comprehension. Including multiple syntactic structures in assessment can more precisely characterize children's oral reading fluency. Incorporating prosodic cues such as pitch and pause might be an effective way to enhance reading comprehension in both L1 Chinese and L2 English. However, reading comprehension is complex and relies on multiple factors; so prosodic reading should be considered simultaneously with all other factors (e.g., word reading efficiency, passage reading fluency, and oral language comprehension).

While our findings have significant implications, there are limitations that suggest areas for further investigation. First, our sample size was limited by the labour-intensive spectrographic analysis of an individual children's oral reading of two passages. The advantage of our approach was the use of an objective measure. However, we recommend that future research should seek to reproduce these findings, and also the utility of our approach, by testing a larger sample of bilingual children with multiple measures of prosodic reading, including both spectrographic analysis and a rating scale. Second, we did not directly assess participants' prior syntactic knowledge. Given the evidence showing a link between syntactic knowledge and reading comprehension (Deacon & Kieffer, 2017), future studies should examine how an individual's prior syntactic knowledge may influence the contribution of prosodic reading to reading comprehension. Third, our study primarily focused on expressive prosody and did not explore perceptual acuity in recognizing acoustic-phonetic cues associated with prosodic structure, known as prosodic sensitivity. Increasing evidence has shown the connection between prosodic sensitivity, which is further influenced by oral language skills such as vocabulary (Chung et al., 2021), or syntactic awareness (e.g., Tong et al., 2022). An extension of this study would be to simultaneously evaluate bilingual children's prosodic reading and prosodic sensitivity, and other oral language skills, in order to determine the potential impacts of these language skills on the link between prosodic reading and reading comprehension.

Additionally, our study does not speak to directionality. In particular, our findings do not allow us to differentiate between two possibilities: Whether prosody aids in better reading comprehension abilities for bilingual children or if proficient bilingual readers are more skilled at utilizing prosody. Other work has found that greater reading comprehension can yield better prosodic reading (e.g., Hudson et al., 2005; Nuria et al., 2017; Ravid & Mashraki, 2007; Torgesen & Hudson, 2006). Hence, the prosody-reading comprehension link might be bi-directional, such that prosody may help chunk texts into syntactically and semantically related units that facilitate processing and comprehension. Conversely, a proficient understanding of text may bootstrap children's precise assignment of boundary or phrase breaks according to major syntactic and semantic elements, thereby enhancing prosodic reading. Preliminary exploration of the reverse direction (comprehension affects pitch or pause) show some positive results (see supplementary Table S6) that future studies can explore in detail. Thus, future research can distinguish these possible relations in order to further elucidate the prosody-reading comprehension link in bilingual readers.

Finally, although our findings are best explained by PCH framework, which frames reading comprehension as a data-prediction and optimization problem using implicit prosody, this aspect has been neglected in previous theories. To validate this model, further testing using experimental and interventional designs is necessary. One plausible approach is to provide prosodic reading training to young children and directly measure the impacts of PCH on their reading comprehension. Moreover, despite demonstrating a significant cross-language association between L1 Cantonese prosodic reading and L2 English reading comprehension, our participants were Chinese-dominant bilingual children whose L2 English skills were not as proficient as their L1 Chinese. This imbalance may affect our results. Therefore, further work that includes a more balanced Chinese-English bilingual cohort is necessary for a complete understanding of the crossover effect of prosodic reading.

5. Main conclusions

Our findings of spectrographic analyses of Cantonese-English bilingual children's L1 and L2 prosodic reading in grades 2 and 3 demonstrated that: 1) these children are aware of language-independent functions and language-specific manifestations of pitch and pause cues within and across their L1 Chinese and L2 English; and 2) their use of these pitch and pause cues depends on language-specific syntactic structures. Furthermore, pitch, and especially *wh* question pitch contour, and some pause structures were significantly associated with reading comprehension in either L1 Chinese or L2 English. The cross-language association occurred from L1 Cantonese yes-no pitch rise and phrase-final pitch fall to L2 English reading comprehension. These results are best explained by the prosodic catalysing hypothesis (PCH), which characterizes the roles of pitch marking and pause duration in reading comprehension. However, given the correlational and explorative nature of our work, further experimental and intervention research is necessary to test PCH and clarify the nature and also direction of the prosody-reading comprehension relation.

References

Abboub, N., Nazzi, T., & Gervain, J. (2016). Prosodic grouping at birth. *Brain and Language*, *162*, 46-59. https://doi.org/10.1016/j.bandl.2016.08.002

Álvarez-Cañizo, M., Suárez-Coalla, P., & Cuetos, F. (2015). The role of reading fluency in children’s text comprehension. *Frontiers in Psychology*, *6*, Article 1810. https://doi.org/10.3389/fpsyg.2015.01810

Arcand, M.-S., Dion, E., Lemire-Théberge, L., Guay, M.-H., Barrette, A., Gagnon, V., Caron, P.-O., & Fuchs, D. (2014). Segmenting texts into meaningful word groups: Beginning readers’ prosody and comprehension. *Scientific Studies of Reading, 18*(3), 208–223. https://psycnet.apa.org/doi/10.1080/10888438.2013.864658

Bandura, A. (1977). *Social learning theory.* Prentice-Hall.

Bandura, A., & Harris, M. B. (1966). Modification of syntactic style. *Journal of Experimental Child Psychology, 4*(4), 341–352. https://doi.org/10.1016/0022-0965(66)90036-1

Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, *93*(3), 491–507.

Bertsekas, D. P. (2014). *Constrained optimization and Lagrange multiplier methods*. Academic Press.

Boersma, P., & Weenink, D. (2015). *Praat: Doing phonetics by computer (Version 5.4.22)* [computer software]. http://www.praat.org/

Breen, M. (2014). Empirical investigations of the role of implicit prosody in sentence processing: Implicit prosody. *Language and Linguistics Compass*, *8*(2), 37–50. https://doi.org/10.1111/lnc3.12061

Buxó-Lugo, A., & Watson, D. G. (2016). Evidence for the influence of syntax on prosodic parsing. *Journal of Memory and Language*, *90*, 1–13.

https://doi.org/10.1016/j.jml.2016.03.001

Cain, K. (2007). Syntactic awareness and reading ability: Is there any evidence for a special relationship? *Applied Psycholinguistics*, *28*(4), 679–694. https://doi.org/10.1017/S0142716407070361

Cain, K., & Oakhill, J. (2007). Reading Comprehension Difficulties: Correlates, Causes, and Consequences. In K. Cain & J. Oakhill (Eds.), *Children's comprehension problems in oral and written language: A cognitive perspective* (pp. 41–75). The Guilford Press.

Chafe, W. (1988). Punctuation and the prosody of written language. *Written Communication, 5*(4), 395–426. https://doi.org/10.1177/0741088388005004001

Chao, Y. R. (1930). A system of tone letters. *Le Maitre Phonetique, 45*, 24-27. Chen, C.C., & C-P., Au (2004). Tone assignment in secondlanguage prosodic learning". Paper presented at Speech Prosody 2004. Nara, Japan. http://sprosig.org/sp2004/PDF/Chen-Au.pdf Cole, J., Mo, Y., & Baek, S. (2010). The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes*, *25*(7), 1141-1177. https://doi.org/10.1080/01690960903525507

Charlton, C., Rasbash, J., Browne, W.J., Healy, M. and Cameron, B. (2020) *MLwiN Version 3.05*. Centre for Multilevel Modelling, University of Bristol.

Chung, W.-L., Bidelman, G. M. (2022). Acoustic features of oral reading prosody and the relation with reading fluency and reading comprehension in Taiwanese children. Journal of Speech, Language, and Hearing Research, 65, 334-343. doi.org/10.1044/2021_JSLHR-21-00252

Chung, W.-L., Jarmulowicz, L., & Bidelman, G. M. (2021). Cross-linguistic contributions of acoustic cues and prosodic awareness to first and second language vocabulary knowledge. Journal of Research in Reading, 44(2), 434-452. doi:10.1111/1467-9817.12349

Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech.* Harvard University Press.

Cooper, W. E., & Sorensen, J. M. (1981). *Fundamental frequency in sentence production*. Springer New York. https://doi.org/10.1007/978-1-4613-8093-1

Courtright, J. A., & Courtright, I. C. (1976). Imitative modeling as a theoretical base for instructing language-disordered children. *Journal of Speech & Hearing Research, 19*(4), 655–663. https://doi.org/10.1044/jshr.1904.655

Cruttenden, A. (1997). *Intonation (2nd ed.).* Cambridge University Press.

Cummins, J. (2000). The threshold and interdependence hypotheses revisited. In *Language, power and pedagogy* (pp. 173–200). Multilingual Matters. https://doi.org/10.21832/9781853596773-008

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40*(2), 141–201. https://doi.org/10.1177/002383099704000203

Deacon, S. H., & Kieffer, M. (2018). Understanding how syntactic awareness contributes to reading comprehension: Evidence from mediation and longitudinal models. *Journal of Educational Psychology, 110*(1), 72-86. https://doi.org/10.1037/edu0000198

Deng, Q., & Tong, X. (2021). Understanding text reading comprehension in Chinese students who are d/deaf and hard of hearing: The roles of segmental phonological awareness and suprasegmental lexical tone awareness. *American Annals of the Deaf*, *166*(4), 462–477. https://doi.org/10.1353/aad.2021.0039

Dowhower, S. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory into Practice*, *30*(3), 165–175. https://doi.org/10.1080/00405849109543497

Education Bureau of the Hong Kong Special Administrative Region. (2004). *Primary one admission (POA) system* (Education Bureau Circular No. 9/2004).

Erekson, J. A. (2010). Prosody and interpretation. *Reading Horizons: A Journal of Literacy* *and Language Arts, 50* (2). Retrieved from https://scholarworks.wmich.edu/reading_horizons/vol50/iss2/3

Fahn R-L.(2003). Chinese-speaking children's production of Wh-questions. *Concentric: Studies in English Literature and Linguistics, 29 (2)*, 82-117.

Fernandes, S., Querido, L., Verhaeghe, A., & Araújo, L. (2018). What is the relationship between reading prosody and reading comprehension in European Portuguese? Evidence from grades 2 to 5. *Journal of Research in Reading*, *41*(S1), S102–S129. https://doi.org/10.1111/1467-9817.12248

Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research, 27*, 285-319.

Fodor, J.D. (2002). Psycholinguistics cannot escape prosody. *Proceedings of Speech Prosody 2008*, Aix-en-Provence, France.

Fox, A., Luke, K. K., & Nancarrow, O. (2008). Aspects in intonation in Cantonese. *Journal of Chinese Linguistics, 36*(2), 321-366. https://www.jstor.org/stable/23756112

Frazier, L., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences, 10 (6)*, 244-249. http://dx.doi.org/10.1016/j.tics.2006.04.002

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences*, *110*(20), 8051–8056. https://doi.org/10.1073/pnas.1216438110

Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, *47*(1–2), 103–138. https://doi.org/10.1016/0378-5955(90)90170-T

Grasa, A. A. (1989). *Econometric model selection: A new approach*. Dordrecht: Kluwer.

Groen, M. A., Veenendaal, N. J., & Verhoeven, L. (2018). The role of prosody in reading comprehension: evidence from poor comprehenders. *Journal of Research in Reading*, 42(1), 37-57. doi.org/10.1111/1467-9817.12133

Gu, W., Hirose, K., & Fujisaki, H. (2006). Modeling the effects of emphasis and question on fundamental frequency contours of Cantonese utterance. *IEEE Transactions on Audio, Speech, and Language Processing, 14*(4), 1155–1170. https://doi.org/10.1109/TASL.2006.876132

Gutman, A., Dautriche, I., Crabbé, B., & Christophe, A. (2015). Bootstrapping the syntactic bootstrapper: Probabilistic labeling of prosodic phrases. *Language Acquisition, 22*(3), 285–309. https://doi.org/10.1080/10489223.2014.971956

Hermes, D. J., & van Gestel, J. C. (1991). The frequency scale of speech intonation. *The Journal of the Acoustical Society of America*, *90*(1), 97−102. https://doi/10.1121/1.402397

Hirose, Y. (2003). Recycling prosodic boundaries. *Journal of Psycholinguistic Research*, *32*(2), 167–195. https://doi.org/10.1023/A:1022448308035

Hong Kong Educational Bureau, Curriculum Development Council (2007). 香港小學學習字詞表 [Hong Kong Chinese lexical lists for primary learning]. Retrieved April 26, 2021, from http://www.edbchinese.hk/lexlist_ch/

HKSAR Education Bureau. (2009). Wordlists for the primary English language curriculum [Computer software]. http://cd1.edb.hkedcity.net/cd/eng/vocab09/appendix_1_content.htm/

Ho, C. S., Chan, D., Chung, K., Tsang, S. M., Lee, S. H., & Cheng, R. Y. (2007). *The Hong Kong test of specific learning difficulties in reading and writing for primary school students-[HKT-P (II)].* Hong Kong Specific Learning Disabilities Research Team.

Hong Kong Examinations and Assessment Authority. (2020). *Territory-wide system* *assessment 2020 (primary schools) Part 3 English language quick guide*. https://www.bca.hkeaa.edu.hk/web/en/2020QuickGuidePri/QG_P_Part3.pdf

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2*(2), 127–160. https://doi/10.1007/BF00401799

Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis*. Routledge.

Huang, C. (1982). Move WH in a language without WH movement. *The Linguistics Review*, *1*(4), 369–416. https://doi.org/10.1515/tlir.1982.1.4.369

Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). Reading fluency assessment and instruction: What, why, and how? *The Reading Teacher*, *58*(8), 702–714. https://doi.org/10.1598/RT.58.8.1

Hwang, H., & Steinhauer, K. (2011). Phrase length matters: The interplay between implicit prosody and syntax in Korean "garden path" sentences. *Journal of Cognitive Neuroscience*, *23*(11), 3555–3575. https://doi.org/10.1162/jocn_a_00001

Joreskog, K., & Sorbom, D. (2018). *LISREL 10.1*. Scientific Software International.

June, S.-A. (2014). Prosodic typology: By prominence type, word prosody, and macro-rhythm. In S.-A. June (Ed.). *Prosodic Typology II: The phonology of Intonation and Phrasing* (pp.520-539). Oxford University Press.

Kennedy, P. (2008). *Guide to econometrics*. Wiley-Blackwell.

Khouw, E., & Ciocca, V. (2007). Perceptual correlates of Cantonese tones. *Journal of Phonetics*, *35*(1), 104–117. https://doi.org/10.1016/j.wocn.2005.10.003

Kohn M. A., & Senyak J. (2021). *Sample Size Calculators*. UCSF CTSI. [computer software]. https://www.sample-size.net/

Kuhn, M. R., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly, 45*(2), 230–251. https://doi/10.1598/RRQ.45.2.4

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293–323. https://doi/10.1016/0010-0285(74)90015-2

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

Ladd, D. R., & Terken, J. M. B. (1995). Modelling intra- and inter-speaker pitch range variations. In K. Elenius, & P. Branderud (Eds.), *Proceedings of the XIIIth International Congress of Phonetic Sciences, ICPhS, Stockholm, Sweden, August 13-19, 1995* (Vol. 2, pp. 386-389)

Lai, S. A., Benjamin, R. G., Schwanenflugel, P. J., & Kuhn, M. R. (2014). The longitudinal relationship between reading fluency and reading comprehension skills in second-grade children. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 30*(2), 116–138. https://doi.org/10.1080/10573569.2013.789785

Li, L., Valcke, M., Badan, L, & Anderl, C.(2022). Video self-modeling (VSM) as a strategy to instruct CFL students' sentence-level stress. *Sustainability, 14*, 15509. https://doi.org/10.3390/su142315509

Lim, L. (2009). Revisiting English prosody: (Some) New Englishes as tone languages? In Lim, Lisa & Nikolas Gisborne, eds. *The Typology of Asian Englishes* Special Issue, *English World-Wide* 30(2): 218-239.

Ma, J. K-Y., Ciocca, V., & Whitehill, T. L. (2006). Quantitative analysis of intonation patterns in statements and questions in Cantonese. *Proceedings of the Third Intonational Conference on Speech Prosody*, 277–280.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie reading tests* (4th ed.). Riverside.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate*

Behavioral Research, 39, 99–128. https://doi.org/10.1207/s15327906mbr3901_4

Matthews, S., & Yip, V. (2011). *Cantonese: A comprehensive grammar*. Routledge

McBride, C. (2015). Is Chinese special? Four aspects of Chinese literacy acquisition that might distinguish learning Chinese from learning alphabetic orthographies. *Educational Psychology Review, 28(3)*, 523-549. https://doi/10.1007/s10648-015-9318-2

McCurdy, K., Kentner, G., & Vasishth, S. (2013). Implicit prosody and contextual bias in silent reading. *Journal of Eye Movement Research*, *6*(2). https://doi.org/10.16910/jemr.6.2.4

Miller, J., & Schwanenflugel, P. J. (2006). Prosody of syntactically complex sentences in the oral reading of young children. *Journal of Educational Psychology, 98*(4), 839–853. https://doi/10.1037/0022-0663.98.4.839

Miller, J., & Schwanenflugel, P. J. (2008). A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. *Reading Research Quarterly, 43*(4), 336–354. https://doi.org/10.1598%2FRRQ.43.4.2

Montgomerie, R., Little, S. G., & Akin-Little, A. (2014). Video self-modeling as an intervention for oral reading fluency. *New Zealand Journal of Psychology, 43*(1), 18–27.

Nazzi, T., Kemler-Nelson, D. G., Jusczyk P. W., Jusczyk, A. M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy, 1*, 123-147. https://doi/10.1207/S15327078IN0101_11

O'Grady, W. (1997). *Syntactic Development: The Acquisition of English.* Chicago: The University of Chicago Press.

Pennington, M. C., & Ellis, N. C. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *The Modern Language Journal, 84*(3), 372–389. https://www.jstor.org/stable/330567

Perfetti, C. & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*, 22-37. http://dx.doi.org/10.1080/10888438.2013.827687

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105–110. https://doi.org/10.1016/j.tics.2006.12.002

Rasinski, T. V. (1994). Developing syntactic sensitivity in reading through phrase-cued texts. *Intervention in School and Clinic, 29*(3), 165–168. https://doi.org/10.1177/105345129402900307

Ravid, D., & Mashraki, Y. E. (2007). Prosodic reading, reading comprehension and morphological skills in Hebrew-speaking fourth graders. *Journal of Research in Reading*, *30*(2), 140–156. https://doi/10.1111/j.1467-9817.2007.00340.x

Roach, P. (2018). *English phonetics and phonology: A practical course*. Cambridge University Press.

Rowland, C., Pine, J., Lieven, E., & Theakston, A. (2003). Determinants of acquisition order in wh-questions: Re-evaluating the role of caregiver speech. *Journal of Developmental Linguistics, 4*, 5-48.

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, 110, 104038.

Schreiber, P. A. (1991). Understanding prosody's role in reading acquisition. *Theory into Practice, 30*(3), 158–164. https://doi.org/10.1080/00405849109543496

Schulz, P., & Roeper, T. (2011). Acquisition of exhaustivity in wh-questions: A semantic dimension of SLI?. *Lingua*, *121*(3), 383-407. https://doi.org/10.1016/j.lingua.2010.10.005

Schwanenflugel, P. J., Hamilton, A. M., Wisenbaker, J. M., Kuhn, M. R., & Stahl, S. A. (2004). Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers. *Journal of Educational Psychology, 96*(1), 119–129. https://doi.org/10.1037/0022-0663.96.1.119

Shi, D. (1994). The nature of Chinese Wh-questions. *Natural Language and Linguistic Theory, 12(2)*, 301-333.

Snowling, M. J., Stothard, S. E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., Nation, K., & Hulme, C. (2011). *York assessment of reading for comprehension: Passage reading* (2nd ed.). GL Assessment.

Soderstrom, M., Seidl, A., Kemler, N., D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language, 49(2)*, 249-267. https://doi.org/10.1016/S0749-596X%2803%2900024-X

Steinhauer, K., & Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research, 30*(3), 267–295. https://doi.org/10.1023/A:1010443001646

Tong, X., Deacon, S. H., & Cain, K. (2014). Morphological and syntactic awareness in poor comprehenders: Another piece of the puzzle. *Journal of Learning Disabilities*, *47*(1), 22-33. https://doi.org/10.1177/0022219413509971

Tong, X., Deng, Q., & Tong, S.X. (2023). Speech prosody and reading comprehension in Chinese-English bilingual children: The mediating role of syntactic awareness. Journal of Speech, Language, and Hearing Research, 66. doi.org/10.1044/2023_JSLHR-22-00026

Tong, X., Lee, S. M.-K., Lee, M. M.-L., & Burnham, D. (2015). A tale of two features: Perception of Cantonese lexical tone and English lexical stress in Cantonese-English bilinguals. *PLoS ONE*, *10*(11), e0142896. https://doi.org/10.1371/journal.pone.0142896

Tong, X., Kwan, J. L. Y., Xiuli Tong, S., & Deacon, S. H. (2021). How Chinese–English bilingual fourth graders draw on syntactic awareness in reading comprehension: Within-and cross-language effects. *Reading Research Quarterly*. https://doi.org/10.1002/rrq.400

Torgesen, J. K., & Hudson, R. F. (2006). Reading fluency: Critical issues for struggling readers. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 130–158). International Reading Association.

Torgesen, J. K., Wagner, R., & Rashotte, C. (2012). *Test of word reading efficiency:(TOWRE-2)*. Pearson Clinical Assessment.

Vaissière, J. (1983). Language-independent prosodic features. In *Prosody: Models and measurements* (pp. 53-66). Springer.

van der Lely, H. K. J., Jones, M., & Marshall, C. R. (2011). Who did Buzz see someone? Grammaticality judgement of wh-questions in typically developing children and children with Grammatical-SLI. *Lingua*, *121*(3), 408–422. https://doi.org/10.1016/j.lingua.2010.10.007

van Ommen S, Boll-Avetisyan N, Larraza S, Wellmann C, Bijeljac-Babic R, Höhle B, Nazzi T. (2020). Language-specific prosodic acquisition: A comparison of phrase boundary perception by French- and German-learning infants. *Journal of Memory and Language, 112* (104108). https://doi/10.1016/j.jml.2020.104108

Veenendaal, N. J., Groen, M. A., & Verhoeven, L. (2016). Bidirectional relations between text reading prosody and reading comprehension in the upper primary school grades:

A longitudinal perspective. *Scientific Studies of Reading*, *20*(3), 189–202. https://doi.org/10.1080%2F10888438.2015.1128939

Wade-Woolley, L., Wood, C., Chan, J., & Weidman, S. (2021). Prosodic competence as the missing component of reading processes across languages: Theory, evidence and future research. *Scientific Studies of Reading*. https://doi.org/10.1080/10888438.2021.1995390

Wechsler, D. (2011). *WASI-II: Wechsler abbreviated scale of intelligence* (2nd ed.). PsychCorp.

Wee, L. H. (2008). Phonological patterns in the Englishes of Singapore and Hong Kong. World Englishes, 27(3-4), 480-501. https://doi.org/10.1111/j.1467-971X.2008.00580.x.

Wiederholt, J. L., & Bryant, B. R. (2001). Gray Oral Reading Tests–Fourth Edition.

Wolters, A. P., Kim, Y-S. G., & Szura, J. W. (2020). Is reading prosody related to reading comprehension? A meta-analysis. Scientific Studies of Reading, 26(1), 1-20. doi-org.eproxy.lib.hku.hk/10.1080/10888438.2020.1850733

Yang, Y., Gryllia, S., & Cheng, L. L. S. (2020). Wh-question or wh-declarative? Prosody makes the difference. *Speech Communication, 118*, 21–32. https://doi.org/10.1016/j.specom.2020.02.002

Yeung, S. (2018). Second language learners who are at-risk for reading disabilities: a growth mixture model study. *Research in Developmental Disabilities, 78*, 35-43.

Zvonik, E. & Cummins, F. (2003). The effect of surrounding phrase lengths on pause duration. In *European Conference on Speech Communication and Technology (EUROSPEECH-2003).* 777–780