Article

How visual word decoding and context-driven auditory semantic integration contribute to reading: A test of additive vs. multiplicative models

Yu Li[1], Hongbing Xing[2], Linjun Zhang[3*], Hua Shu[4] and Yang Zhang[5*]

[1] Division of Science and Technology, BNU-HKBU United International College, Zhuhai, Guangdong, China
[2] Institute on Education Policy and Evaluation of International Students, Beijing Language and Culture University, Beijing, China
[3] Beijing Advanced Innovation Center for Language Resources and College of Advanced Chinese Training, Beijing Language and Culture University, Beijing, China
[4] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China
[5] Department of Speech-Language-Hearing Sciences and Center for Neurobehavioral Development, University of Minnesota, Minneapolis, MN, USA

[*] Correspondence: Linjun Zhang zhanglinjun75@gmail.com ; Yang Zhang zhang470@umn.edu

Abstract: Theories of reading comprehension emphasize decoding and listening comprehension as two essential components. The current study aimed to investigate how Chinese character decoding and context-driven auditory semantic integration contribute to reading comprehension in Chinese middle school students. Seventy-five middle school students were tested. Context-driven auditory semantic integration was assessed with speech-in-noise tests in which the fundamental frequency (F_0) contours of spoken sentences were either kept natural or acoustically flattened with the latter requiring a higher degree of contextual information. Statistical modelling with hierarchical regression was conducted to examine the contributions of Chinese character decoding and context-driven auditory semantic integration to reading comprehension. Performance on Chinese character decoding and auditory semantic integration scores with the flattened (but not natural) F_0 sentences significantly predicted reading comprehension. Furthermore, the contributions of these two factors to reading comprehension were better fitted with an additive model instead of a multiplicative model. These findings indicate that reading comprehension in middle schoolers is associated with not only character decoding but also the listening ability to make better use of the sentential context for semantic integration in a severely degraded speech-in-noise condition. The results add to our better understanding of the multi-faceted reading comprehension in children. Future research could further address the age-dependent development and maturation of reading skills by examining and controlling other important cognitive variables, and apply neuroimaging techniques such as functional magmatic resonance imaging to reveal the neural substrates for the contribution of auditory semantic integration and the observed additive model to reading comprehension.

Keywords: reading comprehension; speech-in-noise recognition; nature F_0 contours; flattened F_0 contours; Chinese character decoding

1. Introduction

Reading comprehension involves the construction of literal and inferred meanings from linguistic input in print, which requires cognitive operations at lower-level processes of visual feature extraction and word decoding as well as higher-level processes of lexical access, syntactic analysis, and semantic integration. While different levels of cognitive and linguistic processing might have unique contributions to reading comprehension (Gough & Tunmer, 1986; Perfetti, Landi, & Oakhill, 2005; Vellutino, Tunmer, Jaccard, & Chen,

2007; McNamara & Magliano, 2009; Perfetti & Stafura, 2014), these processes can be simplified to two separate components, i.e., decoding (recognizing printed words) and linguistic comprehension (taking lexical information and deriving sentence and discourse interpretations with listening comprehension most often measured as a substitute) (Gough & Tunmer, 1986; Hoover & Gough, 1990). Specifically, the Simple View of Reading (SVR) postulates that once isolated printed words are recognized, the remainder processes in reading comprehension are analogous to those involved in listening comprehension. At the neural level, functional magnetic resonance imaging studies have revealed that in addition to modality-specific brain activation, the two modalities of language comprehension activate common brain regions of the temporal and inferior frontal cortices associated with semantic processing (Buchweitz et al., 2009; Regev et al., 2013; Deniz et al., 2019), thus providing neural evidence for shared cognitive processes underlying reading and listening comprehension. Of these cognitive processes, semantic integration is an essential bridging element by which listeners integrate lexical-semantic information with pragmatic and sociolinguistic knowledge to reach an interpretation that reflects semantic representations (Cutler & Clifton, 1999). Given the critical role of auditory semantic integration over the context of discourses in listening comprehension, the abilities to recognize printed words and make use of top-down contextual semantic information could serve as good predictors for reading comprehension.

During speech recognition and understanding, top-down semantic integration is particularly important in adverse conditions in which acoustic-phonetic analysis alone is insufficient for the identification of particular words because of suboptimal listening backgrounds and/or poor speech signal quality. Specifically, the presence of various types of interference such as broadband noise and one-/multi-talker babbles deteriorates speech intelligibility dramatically, but listeners are able to use semantic context to offset the detrimental effects to a great extent (Dubno et al., 2000; Scott, Rosen, Wickham, & Wise, 2004; Golestani, Rosen, & Scott, 2009; Calandruccio, Dhar, & Bradlow, 2010). Similarly, semantic context is also used by listeners to aid speech recognition and comprehension when the speech signal itself is degraded (e.g., Patel, Xu, & Wang, 2010; Wang, Shu, Zhang, Liu, & Zhang, 2013). Furthermore, when degraded speech signal is presented against interference, listeners benefit even more from semantic context. For example, the intelligibility difference between acoustically degraded sentences and semantically unrelated words is much greater when they are presented in suboptimal listening backgrounds than in quiet, indicating that listeners rely more on the top-down semantic context to aid speech recognition and comprehension in adverse conditions (Binns & Culling, 2007; Wang et al., 2013; Jiang, Li, Shu, Zhang, & Zhang, 2017).

The development of auditory semantic integration ability to help speech recognition and comprehension occurs during early age and continues to improve in the entire childhood. While children are able to use semantic context to assist speech recognition in quiet as early as 2 years old (Fernald, 2011), the ability to use semantic context in adverse listening backgrounds continues to improve into adolescence. Previous studies on speech intelligibility of high- and low-predictability sentences in children at 5-17 years of age consistently showed that contextual semantic cues assisted children of all ages in identifying words masked by babbles with older children performing better than younger children in the high context conditions (Elliot, 1979; Fallon, Trehub, & Schneider, 2000; Stelmachowicz, Hoover, Lewis, Kortekaas, & Pittman, 2000). For example, elementary school children aged 11 and 13 years performed significantly poorer than middle school children aged 15 and 17 years on speech recognition in noise and this difference occurred primarily for high-predictability sentences presented at the 0 dB signal-to-noise (SNR) ratio (Elliot, 1979). In addition, there are developmental changes in children's ability to use semantic context to decode degraded signals in quiet and interfering backgrounds (Cole & Perfetti, 1980; Craig, Kim, Rhyner, & Chirillo, 1993). For example, children aged 5-7 years required as much acoustic-phonetic information with high-predictability as with low-predictability contexts to identify target words presented in quiet. By contrast, children aged 8-10 years required shorter portions of the target words (i.e., less acoustic-phonetic information) for

recognition in high-predictability than in low-predictability sentences (Craig et al., 1993). Our previous study (Zhou et al., 2017) showed that elementary school children aged 10 years could not make better use of semantic context in recognizing speech with flattened fundamental frequency (F_0) contours compared to speech with natural F_0 contours. By contrast, middle school children aged 14 years benefited more from semantic context when natural F_0 contours were altered regardless of whether the spoken sentences were presented in quiet or suboptimal listening backgrounds. These findings suggest that semantic context can be better used to facilitate listening comprehension and speech recognition by middle school students compared with younger children.

Reading is a cultivated process that maps written symbols to phonological and semantic representations, and learning to read universally involves solid and stable interactions between written and spoken languages through a large amount of training in schools or other learning experiences. Chinese has a logographic writing system, which is markedly different from alphabetic languages. In alphabetic languages, graphemes that correspond to phonemes of spoken languages are used as visual symbols and word decoding follows grapheme-to-phoneme conversion rules. In written Chinese, however, characters are used as basic writing units with no parts in a character corresponding to phonemes. It is never the case in Chinese writing system that a phonetic component maps onto a subsyllabic phonological representation in the way that a letter maps onto a part of a word's phonological form in an alphabetic system (Perfetti, Landi, & Oakhill, 2005). Furthermore, the layout of graphemes in an alphabetic language is linear, whereas a Chinese character is composed of intricate strokes packed into a square configuration. Full literacy in Chinese requires knowledge of between 3000 and 4000 different characters and it relies to a large degree on rote memory, which is in sharp contrast to that of alphabetic languages which use a relatively limited number of symbols (typically 22-30 or so letters of the alphabet) to produce all words in the languages (Li, Shu, McBride-Chang, Liu, & Peng, 2012). Therefore, learning to read can present a much greater challenge for Chinese-speaking children than for children who speak alphabetic languages, and there is evidence that the challenge persists into the late stage of reading development. It takes children longer time to achieve decoding accuracy when learning to read in a deep orthography compared with a shallow orthography (Seymour, 2005). Moreover, studies on alphabetic languages have revealed that the relation between decoding and reading comprehension decreases sharply after age 9-10 years, indicating a developmental transition (Florit & Cain, 2011; Garcia & Cain, 2014). Considering the opaqueness and visual complexity of Chinese orthography, the transition may be much delayed, probably occurs after elementary schooling. Indeed, previous studies of Chinese reading have suggested that decoding is still a strong predictor of reading comprehension in middle school children aged 13-14 years (Chung, Ho, Chan, Tsang, & Lee, 2011, 2013; Chung, Lo, Ho, Xiao, & Chan, 2014). The correlation coefficients (*.55 - .70*) reported in these studies are higher than that found in English-speaking children with similar age (*.41* in Garcia & Cain, 2014). However, during the development of Chinese reading ability, it remains unclear how the processes of character decoding and contextual semantic integration for speech recognition may independently or interactively contribute to reading comprehension. This critical issue is particularly pertinent to the middle school students who have already got through the early stage of learning to read with a sizable vocabulary (Chall, 1983) but are still facing the challenge of reading development in Chinese.

Based on these extant studies on the development of auditory semantic integration ability and character decoding, it can be predicted that the two abilities, i.e., auditory semantic integration and character decoding, can serve as good predictors of reading comprehension in Chinese middle school children. However, to our knowledge, no previous studies have investigated this issue systematically. This model-testing study aimed to fill this gap by examining the relative contributions of the two subskills to reading comprehension. The Chinese language is particularly interesting and suitable for our purpose due to its special orthographic system that takes considerable amounts of time to learn and a phonological system that deploys pitch contour variations at each syllable (known

as the lexical tones) for different words. Specifically, to assess the role of contextual semantic integration in speech recognition, we introduced acoustic manipulations in two speech-in-noise conditions, one with natural F_0 contours kept in the target sentences presented against interfering background speech and the other with flattened F_0 contour that disrupted the critical cue for Chinese lexical tones for proper word recognition. This speech-in-noise test protocol with Mandarin Chinese materials had been previously adopted in a number of studies on elementary school students, middle school students, and adults including the elderly population (Wang et al., 2013; Jiang et al., 2017; Zhou et al, 2017), and the results demonstrated that greater auditory semantic integration at the sentence level is required to recognize the words in the F_0-degraded condition. Furthermore, two statistical models, i.e., the multiplicative model (product of the two subskills) and additive model (sum of the two subskills), were tested to clarify how the subskills would predict the variance in reading comprehension. The multiplicative model predicts that if an individual performs poorly on one subskill, reading comprehension would also be poor no matter how he or she performs on the other subskill. The additive model predicts that reading comprehension is adequately explained by the linear combination of the two subskills. However, given the mounting evidence from studies of reading comprehension predicted by decoding and linguistic comprehension that the additive model fitted data well and could provide a better fit than the multiplicative model of the two subskills in alphabetic languages (Chen & Vellutino, 1997; Savage, 2006; Savage & Wolforth, 2007; Georgiou, Das, & Hayward, 2009) and Chinese (Yeung, Ho, Chan, & Chung, 2016), we hypothesized that the additive model might provide better predictive modelling results for our data. Findings of this study will add to our understanding of the cognitive model of reading to further delineate the developmental trajectory of reading skills, which can provide the foundation to examine the brain mechanisms underlying the acquisition of normal reading skills in children as well as the disruptive neural correlates accompanying developmental dyslexia (McCandliss et al., 2003; Zhang et al., 2012; Finn et al., 2014; Protopapas & Parrila, 2018; Stein, 2018; Lohvansuu et al., 2021).

2.   Methods and Materials

Participants

Seventy-five middle school students (37 females; mean age, 14.2 years; age range, 13.25-15.44 years) were recruited from a middle school located in Beijing to participate in this study. Each had normal or corrected-to-normal vision, and had no known hearing impairments or history of neurological disorders. Participation was voluntary with consent obtained from their legal guardians. This study was approved by the Institute Review Board of Beijing Normal University in compliance with the Declaration of Helsinki for the protection of human subjects.

Test Protocols and Measures

Three independent tests were carried out, measuring Chinese character decoding, semantic integration during speech recognition against interfering background speech, and reading comprehension, respectively. The children were individually tested in quiet rooms with ambient noise level below 45 dB(A). Details of these test procedures were described below.

Chinese character decoding test

The test that was used to estimate children's decoding ability consisted of 150 characters. The first 40 items were taken from a Chinese character decoding test for kindergarten children (Shu, Peng & McBride-Chang, 2008). One hundred of the remaining 110 characters were taken from Chinese language textbooks for elementary school children with 20 items from each grade level from Grades 2 to 6 (Shu, Chen, Anderson, Wu, & Xuan, 2003). The last 10 characters had not been introduced in textbooks for both kindergarten and elementary school children. During the test, children were asked to read the characters aloud one by one and the testing was terminated after reading 15 consecutive items

incorrectly. Each character was worth one mark. The performance score was the number of correctly pronounced characters. This test has been widely used in our previous studies of Chinese child reading and has been demonstrated excellent reliability and validity (e.g., Zhang et al., 2012; Xue, Shu, Li, Li, & Tian, 2013; Xia, Hoeft, Zhang, & Shu, 2016).

Auditory semantic integration during speech recognition against interference

Performance on speech recognition against interference was used to indicate auditory semantic integration ability. Stimuli in this speech recognition test were taken from a corpus used in our previous study (Zhou et al., 2017) and only a brief description is provided here. In this speech-in-noise test, two types of sentences, i.e., sentences with natural and flattened F_0 contours were used as targets. Sentences with natural F_0 contours were read by a male native Chinese speaker and the monotonous sentences with flat F_0 contours were created by flattening the natural F_0 contours at each sentence's mean F_0. Masker stimuli were consonant-misplaced sentences, which were constructed based on the normal sentences. Specifically, the initial consonant of every syllable in each normal sentence was replaced with another consonant as long as the phonotactic rules of Chinese were not violated. Consonant-misplaced sentences were unintelligible at both word and sentence levels, thus having minimal effect of informational masking. A female native Mandarin speaker read the masker sentences, which was to enable separation of the target message from the interfering speech (Figure 1). The target and masker sentences were first edited to be at the sound pressure levels of 75 dB and 70 dB respectively and then combined to form the speech against interference stimuli with SNR level set at +5 dB (refer to the supplementary materials for the sample stimuli).

The speech stimuli were delivered via a pair of Edifier R18 loudspeakers with the sound level set at 65 dB SPL calibrated at the approximate center position of the listener's head. The participants were instructed to listen to the target sentences in male voice carefully and verbally repeat what they heard, as commonly administered in a typical speech intelligibility test. Before the actual test phase, all children participated in a brief practice session representing samples of the experimental conditions. This test protocol adopted a self-paced paradigm, and the participants were encouraged to take a guess when they were not sure which words they had heard. Each child was presented with a total of 28 sentences, half with natural F_0 contours and the other half with flat F_0 contours. The speech materials were prepared to ensure that each participant did not listen to the same sentence with natural or flat F_0 contours and the order of F_0 patterns was counterbalanced across the participants. The responses were recorded and scored by the first author of this paper and checked by an independent auditor blind to the experiment. A strict score standard was adopted. Specifically, only words with consonants, vowels and lexical tones all correctly identified were considered correct answers. The performance score was the proportion of the number of correctly reported words, which was then converted to rationalized arcsine transform units (RAU) (Studebaker, 1985) for the final statistical analyses. The use of RAU transformation helps to reduce ceiling-level saturation and restore homoscedasticity required of normal distribution for parametric statistical tests as the distribution of the scores, i.e., proportion data, was not strictly Gaussian.

Reading comprehension test

The test, developed by following the procedure in Landerl, Fussenegger, Moll, & Willburger (2009), consists of 100 sentences or short passages with gradual increase in length. In the task, each child was given 3 minutes to read as many items as possible and indicate whether they were semantically correct or not. Three examples were given to ensure understanding of the task before the actual test. The performance score was the number of correctly judged items. Excellent reliability and validity have been demonstrated in our previous work on children with dyslexia and typically developing children (Zhang et al., 2012, Xia et al., 2017).

Data analysis

Partial correlation analyses after controlling for age and sex were first carried out to examine the relationship between different measures. Hierarchical regression analyses were then conducted to extract the relative contributions of character decoding and auditory semantic integration to reading comprehension with different orders of the performance on the tests entered the analyses, and the relative contributions of the additive model and multiplicative model. The additive model was defined as the total variance of character decoding and speech recognition no matter what the order entered the hierarchical regression analyses was used, and the multiplicative model as the variance of the multiplication of the two factors (see Johnston & Kirby, 2006 for the same approach).
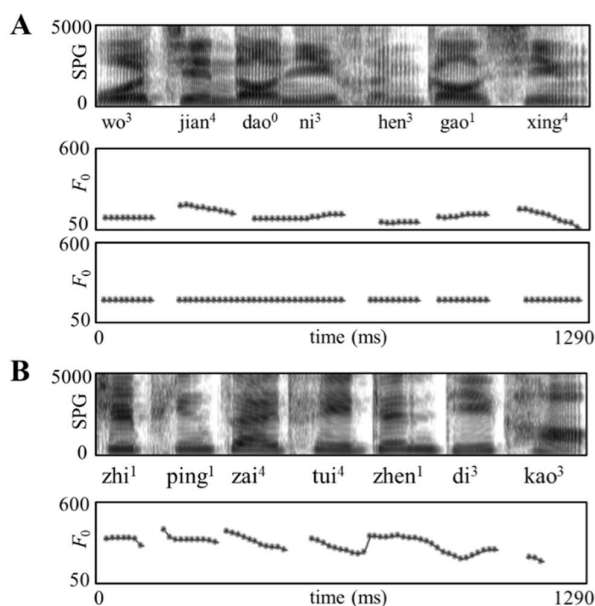

Figure 1. Acoustic features of sample speech stimuli. (A) Broadband spectrograms (SPG: 0 to 5 kHz) and fundamental frequency (F0: 50 to 600 Hz) contours are displayed for a sentence with natural F0 contours and its pitch-flattened counterpart; (B) An interfering sentence with natural F0 contours. Note: The samples are transcribed in Pinyin (the official Romanization system for Standard Chinese) and the figures in the upper right corners signify lexical tones.

3. Results

Performance scores on the Chinese character decoding, speech recognition against interference, and reading comprehension are summarized in Table 1. A paired-samples t-test analysis revealed that recognition of speech with normal F_0 contours was significantly better than that of speech with flattened F_0 contours (t(74) = 9.69, $p < 0.001$), indicating that the latter was more difficult for the listeners.

Table 1. Means, ranges, and standard deviations of each measure.

Measures	M	Range	SD
Chinese character decoding	137	118-148	7.07
Recognition of speech with normal F_0 contours	.93	.31-1	.12
Recognition of speech with flattened F_0 contours	.80	.23-1	.14
Reading comprehension	74.51	46-91	9.57

Table 2 depicts the partial correlations between different measures after controlling for age and sex. The results showed that both Chinese character decoding and reading comprehension were positively correlated with recognition of speech with flattened F_0 contours, but not significantly correlated with recognition of speech with normal F_0 contours. Accuracy scores on the two speech recognition measures were positively correlated.

Table 2. Correlations between variables after controlling for age and sex.

	1	2	3	4
1. Chinese character decoding	-			
2. Recognition of speech with normal F_0 contours [a]	.126	-		
3. Recognition of speech with flattened F_0 contours [a]	.362**	.547***	-	
4. Reading comprehension	.384***	.182	.39**	-

Note. a, transformed data. * $p < .05$. ** $p < .01$. *** $p < .001$.

Hierarchical regression analyses further revealed that Chinese character decoding significantly accounted for the most variance in reading comprehension when it was entered before recognition of speech with normal F_0 contours (13.1%), and the contribution decreased slightly (11.7%) when speech recognition was entered before character decoding. Recognition of speech with normal F_0 contours did not significantly contribute to the variance in reading comprehension no matter it was entered before or after Chinese character decoding (Table 3A, B). However, both Chinese character decoding and recognition of speech with flattened F_0 contours significantly or marginally significantly accounted for the variance in reading comprehension with either of them contributing more when first entered (Table 4A, B).

Table 3. Hierarchical regression analyses predicting reading comprehension from Chinese character decoding and recognition of speech with normal F_0 contours.

Step	Variables	Reading comprehension	
		R^2	ΔR^2
1	Age	.112	.112*
	Sex		
A			
2	Chinese character decoding	.243	.131***
3	Recognition of speech with normal F_0 contours [a]	.252	.009
B			
2	Recognition of speech with normal F_0 contours [a]	.135	.023
3	Chinese character decoding	.252	.117***
C			
2	Chinese character decoding × Recognition of speech with normal F_0 contours [a]	.219	.106**

Note. a, transformed data.   * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 4. Hierarchical regression analyses predicting reading comprehension from Chinese character decoding and recognition of speech with flattened F_0 contours.

Step	Variables	Reading comprehension	
		R^2	ΔR^2
1	Age	.112	.112*
	Sex		
A			
2	Chinese character decoding	.243	.131***
3	Recognition of speech with flattened F_0 contours [a]	.275	$.032^{\Psi}$
B			
2	Recognition of speech with flattened F_0 contours [a]	.200	.088**
3	Chinese character decoding	.275	.074**
C			
2	Chinese character decoding × Recognition of speech with flattened F_0 contours [a]	.261	.148***

Note. a, transformed data. $^{\Psi} p < .09$. * $p < .05$. ** $p < .01$. *** $p < .001$.

The hierarchical regression analyses also revealed that the additive model contributed to 14.0% of the variance in reading comprehension for recognition of speech with normal F_0 contours and the percentage increased to 16.3% for recognition of speech with flattened F_0 contours. By comparison, the multiplicative model contributed to 10.6% of the variance in reading comprehension for recognition of speech with normal F_0 contours and the percentage increased to 14.8% for recognition of speech with flattened F_0 contours

(Table 3C and Table 4C). These results indicate that the additive model accounted for more variance than the multiplicative model.

4. Discussion

Reading comprehension encompasses various components and the past decades have witnessed great progress on the understanding of its complexity. Based on the Simple View of Reading (SVR) framework, the present study examined the contribution of Chinese character decoding and auditory semantic integration to reading comprehension among Chinese middle school students who have largely passed the early stage of learning to read during elementary schooling and are at the stage of reading to learn new knowledge (Chall, 1983). The results showed that Chinese character decoding significantly contributed to reading comprehension irrespective of the extent of auditory semantic integration, indicating that children who know a greater amount of characters have better reading comprehension. Recognition of speech with natural F_0 contours did not significantly contribute to reading comprehension; however, a significant contribution was observed when natural F_0 contours were flattened, indicating that children who can make better use of sentence-level contextual semantic integration in a more adverse condition have better reading ability. Furthermore, based on the predictive power results, the contributions of the two subskills to reading comprehension fitted better with the additive model than the multiplicative model. By examining the contribution of auditory semantic integration to reading, the current study highlights the complexity of reading comprehension.

Because of the lack of clear grapheme-to-phoneme mapping and complex layout of characters in Chinese logographic writing system, the amount of recognized characters plays a critical role in early reading development. For example, Joshi, Tao, Aaron, and Quiroz (2012) found that the performance on Chinese character decoding explained 22% of the variance in reading comprehension among Grade 2 (about 8 years old) children and 32% in Grade 4 (about 10 years old) children. Together with their findings, results of the current study confirmed that Chinese character decoding contributed significantly to reading comprehension for middle school students, which suggests that the role of Chinese character decoding persists throughout the reading development process before maturity. Because different materials were adopted for the measurement of Chinese character decoding and reading comprehension in the current study and Joshi et al.'s study, it is impossible to make direct comparisons between the contributions of character decoding to reading comprehension and obtain the developmental trend from elementary school to middle school in children. Chung et al. (2011, 2013, 2014) recently found high correlations (rs > .50) between Chinese character decoding and reading comprehension in middle school children. These findings indicate that it is still hard to establish a developmental trend in Chinese, at least not in elementary and middle school children. Studies in alphabetic languages, however, indicate an overall trend in which the importance of word decoding in reading comprehension decreases from elementary school to middle school (García & Cain, 2014). Further investigations are needed to explore the developmental trend in Chinese logographic writing system. It is noteworthy that Chinese character decoding ability can be measured with both accuracy and fluency. In the current study, only accuracy scores were obtained. There has been some evidence that Chinese character decoding fluency contributes to reading comprehension in elementary school children (Joshi et al., 2012; Ho et al., 2017). Because middle school children are at the intermediate stage from premature readers to mature readers (Chall, 1983), future longitudinal studies need to examine which measure is a stronger predictor of reading comprehension.

As auditory semantic integration in the sentential context is essential to listening comprehension (Cutler & Clifton, 1999), we predicted that it could serve as a good predictor for reading comprehension. In the present study, the extent of semantic integration needed for recognizing the target spoken sentences was manipulated by presenting two types of stimuli. Specifically, in suboptimal listening conditions, greater auditory semantic integration is required for successful recognition of speech with flattened F_0 contours

compared with speech with nature F_0 contours (Wang et al., 2013; Jiang, et al., 2017). Interestingly, our results showed that recognition of speech with flattened F_0 contours contributed to reading comprehension, while the contribution decreased dramatically when the natural F_0 contours in the target spoken sentences were intact. These results indicate that the more auditory semantic integration is involved in a speech recognition task, the greater the listening task contributes to reading comprehension. There are developmental changes for children to use semantic context to aid speech recognition in adverse conditions. For example, our previous study showed that middle school children could make better use of semantic context in recognizing speech with flattened F_0 contours compared to speech with natural F_0 contours, but elementary school children could not, although they could utilize semantic context during recognizing speech with natural F_0 contours (Zhou et al., 2017). How the ability to make use of semantic information to aid recognition of speech with natural F_0 contours contributes to reading comprehension among elementary school children needs further investigation.

Regarding the statistical modelling issue of how word decoding and linguistic comprehension combine to explain the variance in reading comprehension, previous studies have found that the additive model performed as well as or better than the multiplicative model in explaining the variance in reading comprehension (Chen et al., 1997; Savage, 2006; Savage & Wolforth, 2007; Georgiou et al., 2009). However, this issue was still rarely examined in Chinese. A recent study on Chinese reading comprehension among elementary school children confirmed the superiority of the additive model over the multiplicative model in explaining the contributions of the subskills to reading (Yeung et al., 2016). Although our findings are not meant to be directly compared with those of the previous studies because different subskills were measured, all these studies collectively provide evidence in support of the additive model from a general perspective. Interestingly, the respective contributions of character decoding and recognition of speech with flattened F_0 contours to reading comprehension depended on which one was first entered. When speech recognition was first entered, the relative contribution of character decoding was still significant but much reduced, and vice versa. These results echo the multiplicative model showing higher contribution of the two components than either of them. However, similar patterns were not observed for the recognition of speech with natural F_0 contours. It is very likely that Chinese character decoding and recognition of speech with flattened F_0 contours shared some cognitive processes in explaining the variance in reading comprehension. What these common cognitive processes are, however, remains to be clarified in future studies.

Although the present study has specifically examined the relative contributions of Chinese character decoding and auditory semantic integration in adverse conditions to reading comprehension in middle school students, it is limited in several aspects. Firstly, our study was narrowly focused on testing the multiplicative vs. additive models based on the theoretical framework of SVR. We measured only Chinese character decoding and auditory semantic integration abilities but did not measure other reading-related linguistic subskills (e.g., phonological awareness and vocabulary) and cognitive competence (e.g., non-verbal IQ and executive function) that also contribute to reading comprehension. The relative contributions of these interrelated linguistic and cognitive subskills to reading comprehension in Chinese middle school students have been examined in Chinese primary school students (e.g., Shu, McBride-Chang, Wu, & Liu, 2006; Chik et al., 2012; Yeung, Ho, Chan, & Chung, 2016) and Chinese middle school children (Chung, Ho, Chan, Tsang, & Lee, 2011, 2013; Chung, Lo, Ho, Xiao, & Chan, 2014). These findings have extended the original version of SVR on which we framed the current experimental design, adding to the better understanding of the complexity of reading comprehension. Further investigations are needed to understand how auditory semantic integration and character decoding contribute to reading comprehension after controlling for these linguistic or cognitive subskills. Secondly, reading comprehension can be measured at various levels such as sentence, paragraph, and passage. Long paragraph and passage comprehension require greater semantic integration than sentence or short passage

comprehension used in the present study, which might increase the contribution of auditory semantic integration during speech-in-noise recognition. This issue needs further investigation.

The last two decades have witnessed fast development in neuroimaging techniques such as event-related potentials and functional magnetic resonance imaging and the application of these techniques in revealing neural correlates of cognitive processing. Previous research has shown that the left posterior fusiform gyrus is much involved in visual word form processing after the basic visual processes (McCandliss, Cohen, and Dehaene, 2003; Dehaene & Cohen, 2011) and that both listening and reading comprehension is supported by widely distributed brain regions in the temporal and inferior frontal cortices (Buchweitz et al., 2009; Regev et al., 2013; Deniz et al., 2019). How these regions work in concert to support efficiently the reading processes from visual word form analysis to semantic integration is still a question to be addressed (Protopapas & Parrila, 2018; Stein, 2018) . Neuroimaging techniques can be useful for revealing and characterizing the functional and structural connections between the fusiform gyrus and distributed temporal and frontal regions in typically developing children (Li et al., 2017), thus providing neural evidence for the additive model of reading. Future neuroimaging investigations may also help to reveal how functional and structural disruption of these connections affect reading behaviors in children with dyslexia who are widely thought to have word reading deficits (Lyon, Shaywitz, and Shaywitz, 2003; Finn et al., 2014).

5. Conclusion

In conclusion, our findings revealed that in Chinese middle school students, Chinese character decoding and auditory semantic integration during speech-in-noise recognition contributed to reading comprehension in an additive manner. The current study emphasizes that in addition to decoding and linguistic comprehension measures that are traditionally thought to be reliable predictors of reading comprehension, children's performance on speech recognition under some adverse conditions can also serve as a predictor of reading comprehension because a high degree of semantic integration is involved in these conditions. These new findings from middle school students enrich our understanding of the complexity and decomposition of reading comprehension from a developmental perspective, which support the additive model for conducting further cognitive brain research with potential clinical implications for intervention with dyslexia.

References

Binns, C., & Culling, J. F. (2007). The role of fundamental frequency contours in the perception of speech against interfering speech. *The Journal of the Acoustical Society of America, 122*(3), 1765-1776. doi: 10.1121/1.2751394

Buchweitz, A., Mason, R. A., Tomitch, L., & Just, M. A. (2009). Brain activation for reading and listening comprehension: An fMRI study of modality effects and individual differences in language comprehension. *Psychology & neuroscience, 2*(2), 111-123.

Calandruccio, L., Dhar, S., & Bradlow, A. R. (2010). Speech-on-speech masking with variable access to the linguistic content of the masker speech. *The Journal of the Acoustical Society of America*, *128*(2), 860-869. doi: 10.1121/1.3458857

Chall, J. S. (1983). *Stages of Reading Development*. New York: McGraw-Hill Book Company.

Chen, R. S., & Vellutino, F. R. (1997). Prediction of reading ability: A cross-validation study of the simple view of reading. *Journal of Literacy Research*, *29*(1), 1-24. doi: 10.1080/10862969709547947

Chik, P.P. -M., Ho, C.S. -H., Yeung, P. -S., Chan, D.W. -O., Chung, K.K. -H., Luan, H., ... Lau, W.S. -Y. (2012). Syntactic skills in sentence reading comprehension among Chinese elementary school children. *Reading and Writing, 2*, 679–699. doi: 10.1007/s11145-010-9293-4

Chung, K. K., Ho, C. S., Chan, D. W., Tsang, S. M., & Lee, S. H. (2011). Cognitive skills and literacy performance of Chinese adolescents with and without dyslexia. *Reading and Writing, 24*(7), 835-859. doi: 10.1007/s11145-010-9227-1Chung, K. K., Ho, C. S., Chan, D. W., Tsang, S. M., & Lee, S. H. (2013). Contributions of syntactic awareness to reading in Chinese-speaking adolescent readers with and without dyslexia. *Dyslexia, 19*(1), 11-36. doi: 10.1002/dys.1448

Chung, K. K., Lo, J. C., Ho, C. S., Xiao, X., & Chan, D. W. (2014). Syntactic and discourse skills in Chinese adolescent readers with dyslexia: a profiling study. *Annuals of Dyslexia, 64*(3), 222-247. doi: 10.1007/s11881-014-0095-2

Craig, C. H., Kim, B. W., Rhyner, P. M. P., & Chirillo, T. K. B. (1993). Effects of word predictability, child development, and aging on time-gated speech recognition performance. *Journal of Speech, Language, and Hearing Research*, *36*(4), 832-841. doi: 10.1044/jshr.3604.832

Cutler, A., & Clifton, C. J. (1999). Comprehending spoken language: A blueprint of the listener. In C. Brown & P. Hagoort (Eds.), *The Neurocognition of Language* (pp. 123–166). Oxford: Oxford University Press.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in cognitive sciences, 15*(6), 254-262.

Deniz, F., Nunez-Elizalde, A. O., Huth, A. G., & Gallant, J. L. (2019). The representation of semantic information across human cerebral cortex during listening versus reading is invariant to stimulus modality. *Journal of Neuroscience, 39*(39), 7722-7736.

Dubno, J. R., Ahlstrom, J. B., & Horwitz, A. R. (2000). Use of context by young and aged adults with normal hearing. *The Journal of the Acoustical Society of America*, *107*(1), 538-546. doi: 10.1121/1.428322

Elliott, L. L. (1979). Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *The Journal of the Acoustical Society of America*, *66*(3), 651-653. doi: 10.1121/1.383691

Fallon, M., Trehub, S. E., & Schneider, B. A. (2000). Children's perception of speech in multitalker babble. *The Journal of the Acoustical Society of America*, *108*(6), 3023-3029. doi: 10.1121/1.1323233

Fernald, A. (2001). "Making use of semantic context in early language understanding," paper presented at the Society for Research in Child Development, Minneapolis, MN.

Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review, 23*(4), 553-576. doi: 10.1007/s10648-011-9175-6

Finn, E. S., Shen, X., Holahan, J. M., Scheinost, D., Lacadie, C., Papademetris, X., ... & Constable, R. T. (2014). Disruption of functional networks in dyslexia: a whole-brain, data-driven analysis of connectivity. *Biological psychiatry, 76*(5), 397-404.

García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, *84*(1), 74-111. doi: 10.3102/0034654313499616

Georgiou, G. K., Das, J. P., & Hayward, D. (2009). Revisiting the "simple view of reading" in a group of children with poor reading comprehension. *Journal of Learning Disabilities*, *42*(1), 76-84. doi: 10.1177/0022219408326210

Golestani, N., Rosen, S., & Scott, S. K. (2009). Native-language benefit for understanding speech-in-noise: The contribution of semantics. *Bilingualism (Cambridge, England)*, *12*(3), 385. doi: 10.1017/S1366728909990150

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and special education*, *7*(1), 6-10. doi: 10.1177/074193258600700104

Ho, C. S. H., Zheng, M., McBride, C., Hsu, L. S. J., Waye, M. M., & Kwok, J. C. Y. (2017). Examining an extended simple view of reading in Chinese: The role of naming efficiency for reading comprehension. *Contemporary Educational Psychology*, *51*, 293-302. doi: 10.1016/j.cedpsych.2017.08.009

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

Jiang, W., Li, Y., Shu, H., Zhang, L., & Zhang, Y. (2017). Use of semantic context and F 0 contours by older listeners during Mandarin speech recognition in quiet and single-talker interference conditions. *The Journal of the Acoustical Society of America*, *141*(4), EL338-EL344. doi: 10.1121/1.4979565

Joshi, R. M., Tao, S., Aaron, P. G., & Quiroz, B. (2012). Cognitive component of componential model of reading applied to different orthographies. *Journal of Learning Disabilities*, *45*(5), 480-486. doi: 10.1177/0022219411432690

Johnston, T. C., & Kirby, J. R. (2006). The contribution of naming speed to the simple view of reading. *Reading and Writing*, *19*(4), 339-361. doi: 10.1007/s11145-005-4644-2

Landerl, K., Fussenegger, B., Moll, K., & Willburger, E. (2009). Dyslexia and dyscalculia: Two learning disorders with different cognitive profiles. *Journal of Experimental Child Psychology, 103*(3), 309-324. doi: 10.1016/j.jecp.2009.03.006

Li, H., Shu, H., McBride-Chang, C., Liu, H., & Peng, H. (2012). Chinese children's character recognition: Visuo-orthographic, phonological processing and morphological skills. *Journal of Research in Reading*, *35*(3), 287-307. doi: 10.1111/j.1467-9817.2010.01460.x

Li Y, Zhang L, Xia Z, Yang J, Shu H, Li P. The Relationship between Intrinsic Couplings of the Visual Word Form Area with Spoken Language Network and Reading Ability in Children and Adults. Front Hum Neurosci. 2017 Jun 23;11:327. doi: 10.3389/fnhum.2017.00327

Lohvansuu, K., Torppa, M., Ahonen, T., Eklund, K., Hämäläinen, J. A., Leppänen, P. H. T., & Lyytinen, H. (2021). Unveiling the Mysteries of Dyslexia—Lessons Learned from the Prospective Jyväskylä Longitudinal Study of Dyslexia. *Brain Sciences, 11*(4), 427. doi:10.3390/brainsci11040427

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of dyslexia, 53*(1), 1-14.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in cognitive sciences, 7*(7), 293-299.

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. *Psychology of learning and motivation, 51*, 297-384. doi: 10.1016/S0079-7421(09)51009-2

Regev, M., Honey, C. J., Simony, E., & Hasson, U. (2013). Selective and invariant neural responses to spoken and written narratives. *Journal of Neuroscience, 33*(40), 15978-15988.

Patel, A. D., Xu, Y., & Wang, B. (2010). The role of F0 variation in the intelligibility of Mandarin sentences. In *Speech Prosody 2010-Fifth International Conference* (Chicago, IL).

Perfetti, C. A., Landi, N., & Oakhill, J. V. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: a handbook*. Oxford: Blackwell Publishers

Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The lexical constituency model: some implications of research on Chinese for general theories of reading. *Psychological review, 112*(1), 43. doi: 10.1037/0033-295X.112.1.43

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific studies of Reading, 18*(1), 22-37. doi: 10.1080/10888438.2013.827687

Protopapas, A., & Parrila, R. (2018). Is Dyslexia a Brain Disorder? Brain Sciences, 8(4), 61. doi:10.3390/brainsci8040061

Savage, R. (2006). Reading comprehension is not always the product of nonsense word decoding and linguistic comprehension: Evidence from teenagers who are extremely poor readers. *Scientific Studies of Reading*, *10*(2), 143-164. doi: 10.1207/s1532799xssr1002_2

Savage, R., & Wolforth, J. (2007). An additive simple view of reading describes the performance of good and poor readers in higher education. *Exceptionality Education Canada*, *17*(2), 243-268.

Scott, S. K., Rosen, S., Wickham, L., & Wise, R. J. (2004). A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception. *The Journal of the Acoustical Society of America*, *115*(2), 813-821. doi: 10.1121/1.1639336

Seymour, P. H. K. (2005). Early reading development in European orthographies. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 296-315). Oxford, England: Blackwell. doi: 10.1002/9780470757642.ch16

Shu, H., McBride-Chang, C., Wu, S., & Liu, H. (2006). Understanding Chinese developmental dyslexia: Morphological awareness as a core cognitive construct. *Journal of Educational Psychology, 98*, 122–133. doi: 10.1037/0022-0663.98.1.122

Shu, H., Peng, H., & McBride-Chang, C. (2008). Phonological awareness in young Chinese children. *Developmental Science, 11*(1), 171–181. doi: 10.1111/j.1467-7687.2007.00654.x

Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development*, *74*(1), 27-47. doi: 10.1111/1467-8624.00519

Stein, J. (2018). What is Developmental Dyslexia? Brain Sciences, 8(2), 26. doi:10.3390/brainsci8020026

Stelmachowicz, P. G., Hoover, B. M., Lewis, D. E., Kortekaas, R. W., & Pittman, A. L. (2000). The relation between stimulus context, speech audibility, and perception for normal-hearing and hearing-impaired children. *Journal of Speech, Language, and Hearing Research, 43*(4), 902-914. doi: 10.1044/jslhr.4304.902

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *Journal of Speech, Language, and Hearing Research, 28*(3), 455-462. doi: 10.1044/jshr.2803.455

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skill model of reading development. *Scientific Studies of Reading, 11*, 3-32. doi: 10.1080/10888430709336632

Wang, J., Shu, H., Zhang, L., Liu, Z., & Zhang, Y. (2013). The roles of fundamental frequency contours and sentence context in Mandarin Chinese speech intelligibility. *The Journal of the Acoustical Society of America*, *134*(1), EL91-EL97. doi: 10.1121/1.4811159

Xia, Z., Hoeft, F., Zhang, L., & Shu, H. (2016). Neuroanatomical anomalies of dyslexia: disambiguating the effects of disorder, performance, and maturation. *Neuropsychologia, 81*, 68-78. doi: 10.1016/j.neuropsychologia.2015.12.003

Xia, Z., Zhang, L., Hoeft, F., Gu, B., Gong, G., & Shu, H. (2017). Neural correlates of oral word reading, silent reading comprehension, and cognitive subcomponents. *International Journal of Behavioral Development, 42*, 342–356. doi: 10.1177/0165025417727872

Xue, J., Shu, H., Li, H., Li, W., & Tian, X. (2013). The stability of literacy-related cognitive contributions to Chinese character naming and reading fluency. *Journal of Psycholinguistic Research, 42*(5), 433-450. doi: 10.1007/s10936-012-9228-0

Yeung, P. S., Ho, C. S. H., Chan, D. W. O., & Chung, K. K. H. (2016). A componential model of reading in Chinese. *Learning and Individual Differences*, *45*, 11-24. doi: 10.1016/j.lindif.2015.11.007

Zhang, Y., Zhang, L., Shu, H., Xi, J., Wu, H., Zhang, Y., & Li, P. (2012). Universality of categorical perception deficit in developmental dyslexia: an investigation of Mandarin Chinese tones. *Journal of Child Psychology and Psychiatry*, *53*(8), 874-882. doi: 10.1111/j.1469-7610.2012.02528.x

Zhou, H., Li, Y., Liang, M., Guan, C. Q., Zhang, L., Shu, H., & Zhang, Y. (2017). Mandarin-speaking children's speech recognition: Developmental changes in the influences of semantic context and F_0 contours. *Frontiers in Psychology*, *8*, 1090. doi: 10.3389/fpsyg.2017.01090