



Oral reading fluency of Chinese second language learners

Shuyi Yang^{1,2} 

Accepted: 16 October 2020 / Published online: 4 November 2020
© Springer Nature B.V. 2020

Abstract

The present study examined the components of oral reading fluency (ORF) via various indices and their relationships with comprehension and learner-perceived difficulty of oral reading among Chinese second language (L2) learners. One hundred participants read aloud paragraphs, completed the comprehension test, and rated the difficulty of the oral reading task. Nine indices were generated to measure ORF: the percentage of omissions, substitutions, repetitions, insertions, self-corrections, and segmental problems out of the total produced characters, the number of pauses per minute, the number of actually produced characters per minute, and the number of correctly read characters per minute. The results showed that the nine indices tapped the accuracy, speed, chunking, and monitoring components of the ORF construct. The accuracy, speed, and chunking dimensions effectively predicted comprehension and the accuracy and speed dimension was closely related to learner-perceived difficulty. These findings suggested that ORF is a multi-faceted construct and can be used to assess comprehension and detect learner difficulty in Chinese L2 reading.

Keywords Chinese second language reading · Comprehension · Learner-perceived difficulty · Oral reading fluency

Introduction

Reading in a second language (L2) is a complex process requiring a vast array of intricately interacting sub-processes. The challenge of reading in Chinese is particularly daunting for L2 readers with an alphabetic first language (L1) background because Chinese is a logographic language with unique linguistic features that necessitate specific reading processes and skills.

✉ Shuyi Yang
shuyi.yang@jhu.edu

¹ Department of Asian and Slavic Languages and Literatures, University of Iowa, Iowa City, USA

² Present Address: Johns Hopkins University, Center for Language Education, 3400 N. Charles Street, Baltimore, MD 21218, USA

Oral reading fluency (ORF) is an indicator of the quality of the orchestration of multiple sub-processes and sub-skills that are evidently affected by language-specific features. Fluent processing of the written material is indispensable for efficient comprehension and a critical component of reading development. ORF has become a pillar of reading curriculum and an effective tool to measure comprehension and detect reading difficulty (Kuhn, Schwanenflugel, & Meisinger, 2010), driving major pedagogical decisions. The implementation of fluency instruction and assessment builds upon the sufficient understanding of the ORF construct and its role in skilled reading. Despite the sustained interest in ORF (Berninger et al., 2010; Kuhn, Schwanenflugel, & Meisinger, 2010; Rasinski & Samuels, 2011; Valencia, Smith, Reece, Li, Wixson, & Newman, 2010), the conceptualization of the construct is still incomplete. Most previous literature investigated the accuracy and speed aspects of ORF, with considerably less attention to other aspects.

The importance of ORF extends to L2 reading. Furthermore, since the development of ORF is constrained by language-specific features that are probably different in L1 and L2, understanding L2 ORF would point out the areas of L2 language competence that require targeted teaching and practicing. The accuracy and speed dimensions of ORF in L2 has garnered empirical attention (Crosson & Lesaux, 2010; Jeon, 2012; Jiang, 2016; Jiang, Sawaki, & Sabatini, 2012; McTague, Lems, Butler, & Carmona, 2012), whereas the L2 ORF construct has yet to be explored (Pey, Min, & Wah, 2014). Prior studies focused on alphabetic L2, and little is known about ORF in logographic Chinese. It remains unclear what constitutes ORF in L2 Chinese, and how the multiple tiers of ORF are associated with comprehension and learners' reading difficulty. The present study examined the ORF construct and its relation with comprehension and learner-perceived difficulty for Chinese L2 readers.

Literature review

Oral reading fluency construct

Historically, ORF is hard to define. Despite the conceptual fluidity, a consensus has been reached that ORF is the oral reproduction of connected text with speed and accuracy. ORF demonstrates the ability to read rapidly with ease and accuracy, and with appropriate prosodic expression (Rasinski & Samuels, 2011).

The most thoroughly researched aspects of ORF are accuracy and speed, which are commonly measured by a combined index, the number of words read correctly per minute. It exhibits the speed and correctness with which written material is decoded and reproduced into spoken language. This accurate rate index has been extensively used to as a proxy of ORF among L1 readers (Arnesen et al., 2016; Morris, Pennell, Perney, & Trathen, 2018; Sabatini, Wang, & O'Reilly, 2018; Wayman, Wallace, Wiley, Ticha, & Espin, 2007) and has also increasingly been adopted for L2 ORF (Crosson & Lesaux, 2010; Jeon, 2012; Jiang, 2016; Jiang, Sawaki, & Sabatini, 2012; McTague, Lems, Butler, & Carmona, 2012). However, thus far, the accurate rate has been rarely used in L2 Chinese reading (Lv, 2016).

Miscues, the oral errors produced during text reading (Goodman & Goodman, 1994), constitute a bundle of indices for the ORF accuracy dimension. Non-target substitutions, omissions, and insertions of additional words demonstrate the deviations from the original text and have been the focus of the miscue analysis research. Fluent reading is characteristic of few substitutions, omissions, and insertions (Laing, 2002). Moreover, the patterns of substitutions provide a lens into the ongoing reading processes (Beatty & Care, 2009; Briceño & Klein, 2018; Chang, 2015; Kucer, 2009; Wu & Anderson, 2007; Yan & Wang, 2011). Only a handful of studies (Wang, 2006) have described Chinese L2 readers' miscues.

Repetitions reflect readers' struggles with words/phrases and serve as a speed index of ORF. Frequent repetitions slow the reading down, leaving it choppy and stumbled. Fluent reading, by contrast, characterizes smoothness with few "rough spots" caused by repetitions.

Intrusive pauses within words and inappropriate phrasing of text into fragmented sections disrupt the "prosodic phrasing and contours of the text" (Grabe, 2009, p. 292) and hence capture the chunking/prosody component of ORF. Pauses are measured by objective indices of pausal duration, frequency, and position, whereas phrasing is generally evaluated via subjective scales (Godde, Bosse, & Bailly, 2020). Researchers have found that fluent readers signal major boundaries with pauses and strategically process texts in meaningful chunks (Álvarez-Cañizo, Suárez-Coalla, & Cuetos, 2018; Kuhn & Schwanenflugel, 2018). Less fluent readers either halt inside meaningful units or separate words in ways that deviate from the natural phrasing (Binder, Tighe, Jiang, Kaftanski, Qi, & Ardoin, 2013; Miller & Schwanenflugel, 2006; Valle, Binder, Walsh, Nemier, & Bangs, 2013). It worth mentioning that pause misplacement is more symptomatic of decoding than chunking difficulty, as it is a compensation invoked when word recognition fails (Kuhn, Schwanenflugel, & Meisinger, 2010). Few studies have investigated L2 reading chunking/prosody (Pey, Min, & Wah, 2014) and even fewer for Chinese L2 readers, which was only touched upon in the rough description of within-word pauses and incorrect segmentations in miscues (Wang, 2006).

Self-correction is an overt manifestation of the monitoring processes (Kormos, 1999) and thus taps the monitoring aspect of ORF. As a self-initiated, self-completed repair, self-correction in oral reading comes out when the reader detects the error and executes a correction. Self-correction enhances accuracy, while frequent self-correction would rend the reading less smooth and coherent. Studies have shown that the occurrence of self-correction decreased as readers became more sophisticated (Kucer, 2017; Share, 1990), because fewer errors were made and the majority of miscues did not cause meaning breakdown. Meanwhile, more proficient readers (McGee, Kim, Nelson, & Fried, 2015; Nguyen, Pickren, Saha, & Cutting, 2020) and readers with faster progress (D'Agostino, Kelly, & Rodgers, 2019) displayed a greater tendency to self-correct, revealing a heightened awareness of self-monitoring. L2 readers are less likely to correct themselves, either because their unstable L2 knowledge fails to perceive errors or their insufficient skills of processing the written input leave little cognitive capacity for self-monitoring (Francis, 1999). Chinese L2 reading researchers observed rare

self-correction among novice readers and increasing repair behaviors as learners' proficiency improved (Liu, 1999; Wu & Anderson, 2007).

Finally, it was noted that ORF has an extra element of verbal output compared to silent reading fluency. Reading aloud is a sequential process comprising visual recognition, the conversion of orthographic representations into phonological codes, and the activation of the corresponding articulatory-motor program for overt production (Timmer & Schiller, 2012). In this regard, ORF involves phonological recoding, which imposes another burden for L2 readers. Due to the less developed oral proficiency in L2, even successful word recognition does not guarantee proper articulation for reasons such as failure to sound out the phonemes and mispronunciation (Lems, 2006).

Oral reading fluency and comprehension

The vast majority of ORF studies focused on the accuracy and speed aspects and employed the accurate rate index (number of correctly read words/min). A strong correlation between ORF and comprehension has been established among young L1 readers (Arnesen et al., 2016; Jenkins, Fuchs, Van den Broek, Espin, & Deno, 2003; Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008; Sabatini, Wang, & O'Reilly, 2018; Schilling, Carlisle, Scott, & Zeng, 2007; Silbergliitt, Burns, Madyun, & Lail, 2006; Wanzek, Roberts, Linan-Thompson, Vaughn, Woodruff, & Murray, 2010) and L2 learners (Crosson & Lesaux, 2010; Jeon, 2012; Jiang, 2016; Jiang, Sawaki, & Sabatini, 2012; McTague, Lems, Butler, & Carmona, 2012).

The line of miscue analysis research probed into the relationship between different types of miscues and comprehension. Children who omitted less frequently (Laing, 2002) or made a smaller total number of miscues as well as fewer substitutions with meaning change tended to recall more and better of the text (Beatty & Care, 2009; Kucer, 2009; Wu & Anderson, 2007). The similar pattern has been observed among L2 learners (Wang, 2006). Nevertheless, the small sample sizes of these miscue studies make it impossible to statistically connect miscues with comprehension.

As they further explored, scholars found that prosody, the indicator of efficient chunking, predicts comprehension performance (Arcand et al., 2014; Benjamin & Schwanenflugel, 2010; Calet, Gutiérrez-Palma, & Defior, 2017; Fernandes, Querido, & Verhaeghe, 2018; Groen, Veenendaal, & Verhoeven, 2019; Klauda & Guthrie, 2008; Rasinski, Reutzel, Chard, & Linan-Thompson, 2011). However, prosody does not add much explanatory power to comprehension over accuracy and speed (Riedel, 2007; Sabatini, Wang, & O'Reilly, 2018; Schwanenflugel, Hamilton, Kuhn, Wisenbaker, & Stahl, 2004). Regarding L2 readers, the role of chunking/prosody in comprehension remains underexplored (Pey, Min, & Wah, 2014).

Monitoring, an awareness of the ongoing reading processes, is closely associated with the maintaining of a coherent textual representation. Monitoring enables readers to constantly check whether or not comprehension is occurring and strategically address the possible problems, thus plays a key role in integrating the constructed meaning (Kim, Vorstius, & Radach, 2018). The relationship between

self-correction, a monitoring behavior, and comprehension, has been rarely investigated. Most studies calculated and compared self-correction instances of readers at different levels and yielded inconclusive results (Kucer, 2017; Liu, 1999; McGee, Kim, Nelson, & Fried, 2015; Nguyen, Pickren, Saha, & Cutting, 2020; Share, 1990; Wu & Anderson, 2007). In their inquiry of the association between self-correction and comprehension improvement, D'Agostino, Kelly, & Rodgers, (2019) reported that self-correction positively predicted the progress in comprehension among struggling L1 readers, indicating its facilitative effect.

Generally speaking, the relationship between ORF, especially its accuracy and speed components, and comprehension is robust. The contribution of ORF to comprehension draws theoretical support from Automaticity Theory (DeKeyser, 2001; LaBerge & Samuels, 1974; Segalowitz, 2003; Segalowitz & Segalowitz, 1993). The key premise is that the lower- and higher-level processing compete for the limited cognitive capacity and automaticity of the former ultimately leads to comprehension. If too many cognitive resources are consumed by the lower-level processes such as word recognition and chunking, few will be available for higher-level processes and comprehension will break down. By contrast, if readers perform lower-level processes rapidly and non-deliberately, a large amount of conscious attention will be shifted toward higher-order processes to build comprehension. In other words, reading depends upon the automatic execution of myriad lower-level sub-skills to enable higher-level processes.

ORF, with components of accuracy, speed, chunking/prosody, and monitoring, represents the complex, multi-faceted skill set that entails processes of accessing word meaning, segmenting words into meaningful units, and monitoring the establishment of a coherent representation (Breznitz, 2006). Except the monitoring aspect, ORF is mainly a marker of the efficient orchestration of lower-level processing (Grabe, 2010; Rasinski, Reutzel, Chard, & Linan-Thompson, 2011), which allows adequate cognitive resources to focus on higher-level processes of meaning construction. Therefore, ORF is a prerequisite for comprehension.

The critical role of ORF in comprehension renders it a sensitive tool to diagnose reading problems. The accuracy and speed dimensions of ORF (Fuchs, Fuchs, & Compton, 2004; Hintze & Silbergliitt, 2005; Keh, 2016; Parker et al., 2015; Wise, Sevcik, Morris, Lovett, Wolf, & Schwanenflugel, 2010) have been used to identify weak readers and pinpoint difficulties in word recognition. Much less attention, however, was dedicated to reader-perceived difficulty of oral reading and its connection with performance on different aspects of ORF.

Chinese-specific processes in oral reading fluency

As previously stated, ORF reflects the efficiency of lower-level, linguistic processing, which is constrained by language-specific features. Chinese unique linguistic features necessitate specific lower-level processes that cause difficulties for L2 readers and highlight certain components of ORF.

As a logographic orthography, Chinese lacks script–sound correspondence (Shen, 2013). The basic graphic unit, character, does not provide route to phonological

representation. It imposes great challenges on L2 learners who are accustomed to assembling word pronunciation from phonemes represented in letters. In fact, the difficulty of pronouncing characters/words despite successful lexical access has been frequently mentioned by L2 learners (Hu, 2010). The phonological processing in Chinese is post-lexical: the articulation of written symbols is reliant on the access to meaning. It is possible to know the meaning of a character without knowing its sound, contrasting with being able to pronounce a word without knowing its meaning in alphabetic languages. Decoding without comprehension, a phenomenon common in alphabetic languages, does not work in Chinese; instead, the successful pronunciation indicates successful lexical access. This post-lexical nature of phonological processing allows ORF to be a more robust measure in Chinese than in alphabetic languages.

Another unique feature of Chinese is the lack of spatial demarcation for word boundaries, which necessitates an extra, explicitly operating word segmentation process. The array of characters is continuously parsed into meaningful words as reading proceeds, enabling the specification of individual spatially unmarked words based on which other types of processing can occur (Li, Rayner, & Cave, 2009). It is vital to have existing word representations to segment words, because knowing characters that likely comprise a word helps identify and separate them within the continuous character strings. In this regard, word boundaries naturally arise from successful lexical access. However, under other circumstances, segmentation is implemented in order to recognize words. When encountering unfamiliar words, readers need to determine word boundaries to access the weakly represented entries in mental lexicon. When ambiguity occurs, that is, a character combines with its adjacent characters in different ways to form different words, correct segmentation is essential for the construction of contextually appropriate meaning (Shen & Jiang, 2013). The unspaced layout of Chinese print makes word segmentation an uneasy task for L2 readers whose L1 writing systems have visible word boundaries. L2 readers have a hard time isolating meaningful words in the running text even at advanced levels (Bassetti, 2005; Lee-Thompson, 2008; Shen, 2008; Yang & Jiang, 2012). Distinguished from the phrase-level chunking in reading alphabetic languages, word segmentation in reading Chinese is a fundamental process of word recognition and hence is a more prominent component of ORF.

Previous literature on ORF showed some gaps. First, thus far, research on L2 ORF was scarce and almost exclusively focused on an alphabetic L2. Little is known about ORF in L2 Chinese that activates language-specific processes such as the post-lexical phonological processing in word/character recognition and the explicit word segmentation. Moreover, most studies on ORF – comprehension relationship examined the accuracy and speed aspects, with other aspects largely neglected. Given the crucial role of ORF in reading, it requires exploration of what the ORF construct is and how its different components are associated with comprehension and reader-perceived difficulty among Chinese L2 learners.

The present study was designed to fill the gaps by addressing the following research questions:

1. What are the components of the ORF construct among Chinese L2 learners?
2. How are different components of ORF related to comprehension?

3. How are different components of ORF related to learner-perceived difficulty of oral reading?

Methods

Participants

One hundred L2 Chinese learners from four universities in the U.S. and two universities in China participated in this study (53 males and 47 females; age range: 19 to 26 years old). Among them, 74 were native English speakers, and the rest were from alphabetic European language backgrounds. The racial/ethnic breakdown was as follows: 76% white, 20% Asian, 2% Black, and 2% Hispanic. Convenient sampling was adopted; L2 learners from First ($n=35$), Second ($n=47$), and Third ($n=18$) Year Chinese classes were recruited. A character recognition test was administered to further determine participants' proficiency level.

Instruments

Given that the number of known characters is an important benchmark specified in Chinese proficiency guidelines, a character recognition test (see "Appendix A") was designed to check participants' proficiency level. The test is composed of 50 characters retrieved from the most frequent 1500 Chinese characters listed in the *Modern Chinese Character Frequency List* (National Working Committee for Language and Characters, 1992), which form a hierarchy of characters L2 learners across proficiency levels should master. One character is selected from every 30 characters. Participants read characters aloud as quickly as possible and were given about five minutes to complete this test.

The reading comprehension section of an AP Chinese test was used to measure comprehension. The AP Chinese test has been widely utilized in the U.S. as a designated proficiency assessment for the two-year foreign language requirement at the college level. The test contains six texts of various genres, ranging from 133 to 382 characters in length. It consists of 25 comprehension questions that mainly assess readers' abilities to understand the main idea and specific details, draw inferences, and interpret the purposes of a text. Participants read the texts silently and answered the comprehension questions. The whole process lasted about 40 min.

The oral reading task aimed at eliciting participants' ORF performance. One paragraph was randomly retrieved from each of the six texts in the AP test (see "Appendix B"). The length of the six paragraphs ranged from 81 to 190 characters, with an average of 130 characters. Participants read the paragraphs aloud at their natural pace and they generally spent 10 to 15 min on this task.

A questionnaire (see "Appendix C") was developed to collect participants' evaluation of the oral reading's difficulty level. Participants rated the task difficulty on a 10-point scale, with 1 representing the easiest and 10 representing the hardest. They were also required to explain what aspect of oral reading they felt difficult.

Data collection

The participants were randomly assigned to one of two groups. One group ($n=50$) first took the AP test and character recognition test and then completed the oral reading and the questionnaire. The other group ($n=50$) followed the reverse order. The procedure was counterbalanced to control the practice effect resulting from the use of the same materials twice in the AP test and oral reading. The entire process lasted 60 min.

Scoring

Character recognition test

Each correct naming of a character was marked as correct. A divergence only in tone was treated as correct. Incorrect responses, unintelligible expressions or pronunciations, and skipped or indicated as unknown characters were marked as errors. The accuracy score was the percentage of correctly read characters. The Cronbach's coefficient alpha value was 0.95, indicating high consistency. A second rater scored 15% of the character recognition data (15 participants). The inter-rater reliability was high at $r=0.99$ ($p<.01$). The few discrepancies in coding were negotiated and resolved. Participants' proficiency level was determined following the criteria of previous literature (Zhang, 2018): below 60% accuracy was classified as low proficiency, 60% to 80% accuracy as intermediate proficiency, and above 80% accuracy as high proficiency.

AP test

The accuracy score was the percentage of correctly answered items. The Cronbach's coefficient alpha value showed a reliability estimate of 0.83, indicating relatively high consistency.

Oral reading

Seven types of reading behaviors were identified (see Table 1). An *omission* was considered no response to certain characters/words. A *substitution* was the actually produced, non-target character/word/sound. A *repetition* took place when participants repeated themselves on characters/words. A *pause* was the hesitation for more than two seconds within a word. An *insertion* referred to extra characters/words/sounds that were inserted into the text. A *self-correction* occurred when participants made an error, realized it, and self-corrected successfully. A *segmenting problem* was incorrect grouping of characters into non-words, which resulted in unnatural prosody.

The percentage of omissions, substitutions, repetitions, insertions, self-corrections, and segmental problems out of the total produced characters was calculated.

Table 1 Seven types of oral reading behaviors

	Correct reading	Incorrect reading
Omission (OM)	课临时改到上午 <i>Classes are temporarily changed to morning.</i>	*课(OM)时改到上午
Substitution (SU)	参观我们的学校 <i>Visit our school</i>	*参加我们的学校
Repetition (RP)	然后还可以去图书馆 <i>Then you can go to the library.</i>	*然后还然后还可以去图书馆
Pause (P)	课临时改到上午 <i>Classes are temporarily changed to morning.</i>	*课临时改到上(P)午
Insertion (IN)	参观学校 <i>Visit school</i>	*参观我学校
Self-correction (SC)	西安 <i>Xi'an</i>	*Read 西安 as 山安, then make a correction
Segmental Problem (SP)	人文/科学 <i>humanities</i>	*人/文科/学

The number of pauses per minute was also counted. Another two indices were developed. The unpruned rate was the number of actually produced characters per minute. The pruned accurate rate was the number of correctly read characters per minute, excluding self-corrections, repetitions, and segmental problems. Altogether nine ORF indices were derived. A second rater scored 15% of the oral reading data (15 participants). The inter-rater agreement reached 98%. The few discrepancies in coding were negotiated and resolved.

Questionnaire

The participants' difficulty ratings were recorded as the indicator of learner-perceived difficulty of oral reading. Their comments on the locus of difficulty were collected and analyzed.

Results

RQ1: what are the components of the ORF construct?

The descriptive statistics of character recognition test, AP comprehension test, difficulty rating, and nine ORF indices are displayed in Table 2.

First, the average character recognition accuracy (31.96%) revealed that the participants in this study were at the lower end of the proficiency spectrum (below 60%). Only 7 out of the 100 participants achieved above 60% accuracy.

Table 2 Descriptive Statistics (n = 100)

	<i>M</i>	<i>SD</i>	Min	Max
Character recognition accuracy (%)	31.96	18.62	5.00	89.00
AP Chinese Test Accuracy (%)	52.68	21.76	7.69	96.00
Text reading difficulty	6.94	1.85	3.00	10.00
Unpruned Rate (#characters/min)	55.87	30.69	8.10	137.25
Pruned Accurate Rate (#characters/min)	44.77	24.56	3.56	110.68
Omission (%)	15.13	14.10	0.00	74.80
Substitution (%)	13.39	12.59	0.70	59.67
Repetition (%)	3.07	3.20	0.00	15.90
Pause (#instance/min)	0.92	0.80	0.00	4.36
Insertion (%)	0.34	0.61	0.00	3.14
Self-correction (%)	0.33	0.49	0.00	2.56
Segmental Problem (%)	1.60	1.47	0.00	7.56

Table 3 Factor Loadings for Factor Analysis

	Factor 1: Accurate Rate	Factor 2: Chunking	Factor 3: Monitoring
UR	-.93	.03	.13
PR	-.93	-.03	.05
OM	.72	-.23	-.03
SU	.82	.09	.02
RP	.60	-.31	.23
P	-.21	.83	-.29
IN	.52	.41	.11
SC	-.08	.30	.92
SP	.30	.85	-.03

UR unpruned rate, *PR* pruned accurate rate, *OM* omission, *SU* substitution, *RP* repetition, *P* pause, *IN* insertion, *SC* self-correction, *SP* segmental problem

The data were transformed in response to the non-linearity problems. The percentage scores were transformed with the empirical logit function. The rate scores were log-transformed.

To explore the structure of the ORF construct measured by the nine indices, a factor analysis was run (see Table 3). Factors with an eigenvalue greater than one were retained. Accordingly, a well-defined three factor solution was generated. The three factors had eigenvalue of 3.67, 1.82, and 1.02 respectively, and explained 40.80%, 20.21%, and 11.30% of the total variance. In total, the three factors accounted for 72.31% of the variance. The indicators with a factor loading greater than .45 were considered in the interpretation of the factor. Given the six indices (UR, PR, OM, SU, RP, IN) that loaded highly on the first factor

Table 4 Correlation Matrix between ORF Components and Comprehension

	1. RC	2. AR	3. C	4. M
1. RC	–	.57**	–.25*	.03
2. AR		–	–.29**	.09
3. C			–	.19
4. M				–

RC reading comprehension measured by AP test accuracy, AR accurate rate measured by pruned accurate rate, C chunking measured by segmental problem, M monitoring measured by self-correction

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

Table 5 Hierarchical Regression Analysis for ORF Components

Model	<i>b</i>	<i>SE</i>	β	<i>T</i>	<i>sig</i>	<i>df</i>	R^2	Adj R^2	ΔR^2	ΔF	Sig. <i>F</i> change
1 C	–7.21	2.84	–.25	–2.54	.01	1, 98	.06	.05	.06*	6.46	.01
2 AR	1.94	.31	.55	6.29	.00	2, 97	.33	.32	.27**	39.50	.00
C	–2.61	2.51	–.09	–1.04	.30						

AR accurate rate, C chunking

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

represented accuracy and speed, the factor was labeled the *accurate rate factor*. The second factor was labeled the *chunking factor* because the two indices (P, SP) that loaded highly on it measured the processes of chunking text into meaningful units. One index (SC) tapping monitoring loaded highly on the third factor and thus the factor was labeled the *monitoring factor*.

RQ2: how are different components of ORF related to comprehension?

Three indices that represented each of the three ORF components/factors were selected. They were: pruned accurate rate for the *accurate rate factor*, segmental problem for the *chunking factor*, and self-correction for the *monitoring factor*. Pruned accurate rate was selected because it is a combined index of accuracy and speed and has been widely used in previous literature. Segmental problem was chosen over pause because pause more reflected the deficiency in decoding than chunking (Kuhn, Schwanenflugel, & Meisinger, 2010) and thus was less appropriate for indexing chunking.

Zero-order correlations were run to establish the direction and strength of the relationships between the three ORF components and comprehension (see Table 4).

The students’ comprehension performance had a moderate, significant correlation with accurate rate and a weak albeit significant correlation with chunking. In contrast, the correlation between comprehension and monitoring was negligible.

The accurate rate and chunking components were both significantly related to comprehension. However, considering the high inter-correlations, it was not clear

Table 6 Correlation Matrix between ORF Components and Difficulty Rating

	1. D	2. AR	3. C	4. M
1. D	–	-.42**	.12	-.12
2. AR		–	-.29**	.09
3. C			–	.19
4. M				–

D difficulty rating, *R* accurate rate, *C* chunking, *M* monitoring

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

Table 7 Oral reading difficulty comments

	Comments
Character/ word Recognition	<i>The characters I didn't recognize, making my speaking more choppy and less fluent</i> <i>Without knowing the characters, my fluency was much slower</i>
Pronunciation	<i>There were many words that I could not remember how to pronounce</i>
Chunking	<i>The grammar was also complex, so it was harder to understand the meaning of the sentence</i> <i>The lines are very long, so sometimes when I'm reading I don't know where to pause or break the line to make the sentences more meaningful</i>
Affective	<i>Reading aloud is difficult for me. I am always trying to hurry up and I end up making a lot of mistakes</i> <i>But some of them I can't remember under pressure</i>

whether they would explain the unique variance in comprehension. To investigate the individual and collective contributions of the two ORF components to comprehension, multiple regression analyses were conducted.

As shown in Table 5, chunking was significant when it was the only predictor in the model, accounting for 5% of the variance in comprehension. However, the standardized beta estimate of chunking became non-significant when accurate rate was entered into the equation. Accurate rate, on the other hand, made an independent, significant contribution to comprehension ($p < .01$) after controlling for chunking. Together the two components explained 32% of the variance in comprehension.

RQ3: how are different components of ORF related to learner-perceived difficulty?

Zero-order correlation analyses were operated to investigate the relationship between the three ORF components and learners' difficulty rating (see Table 6).

The learners' difficulty rating had a moderate, significant negative correlation with accurate rate, whereas the correlations between the other two components and difficulty rating were weak and insignificant.

Table 7 summarized participants' remarks on the difficulties in oral reading. Four categories were identified. *Character/word recognition* was considered as the biggest obstacle, reflected in up to 67 statements on this issue. Another problem involved the *pronunciation*. Learners felt hard to pronounce characters/words correctly, although they knew the meaning. The difficulty on *chunking* included segmenting characters into words and parsing complex grammatical structures. The *affective* difficulty referred to the anxiety caused by the task of reading aloud.

Discussion

Components of ORF

According to the factor analysis, Chinese L2 ORF consisted of three components: accurate rate, chunking, and monitoring.

It has been widely accepted that ORF primarily involves speed and accuracy (Kuhn, Schwanenflugel, & Meisinger, 2010; Morris, Pennell, Perney, & Trathen, 2018). Fluent readers can smoothly and effortlessly read at an appropriate pace, whereas slow reading with many halts and repetitions does not represent fluency. Accuracy is assumed in the concept of ORF, and it is meaningless to read rapidly but produce nonsense words. Omission, substitution, and insertion reflected the reader's failure of providing the target responses and captured the accuracy. Unpruned rate outlined the overall pace of oral reading and repetition demonstrated the reader's struggling with processing the characters/words, both representing the speed. Pruned accurate rate combined speed and accuracy.

The results also indicated that chunking constitutes an important dimension of ORF. Chunking, the ability to separate text into meaningful units, is one of the fundamental processes for fluent reading (Kuhn & Schwanenflugel, 2018). Without chunking, readers cannot read smoothly with the appropriate prosodic expression even if they have mastered word decoding skills. Chunking is particularly crucial in reading visually unspaced Chinese, since it enables the grouping of seemingly disjoint characters into larger blocks, namely words, instead of following the slow character-by-character manner. In this study, chunking was measured by within-word pauses and segmental problems. The former exhibited the failed grouping of identified individual characters into words and the latter signified the wrong parsing of characters that yields meaningless units that do not fit the ongoing text.

Monitoring is another component underlying ORF. Monitoring involves checking the coherence of what has so far been constructed and repairing possible breakdowns and errors (Kim, Vorstius, & Radach, 2018), which was manifested in self-correction.

The relationship between ORF and comprehension

One major finding is that the accurate rate component of ORF uniquely and robustly predicted comprehension. Chinese L2 learners who read aloud correctly and rapidly tended to have better textual understanding. It confirmed that accurate and fluent oral reading was related to enhanced comprehension in reading alphabetic L1 (Arnesen et al., 2016; Jenkins, Fuchs, Van den Broek, Espin, & Deno, 2003; Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008; Sabatini, Wang, & O'Reilly, 2018; Schilling, Carlisle, Scott, & Zeng, 2007; Silberglitt et al., 2006; Wanzen et al., 2010) and L2 (Crosson & Lesaux, 2010; Jeon, 2012; Jiang, 2016; Jiang, Sawaki, & Sabatini, 2012; McTague, Lems, Butler, & Carmona, 2012). Such a strong connection extends to reading a logographic L2 (Chinese), suggesting the cross-linguistically universal role of the accuracy and speed aspects of ORF in comprehension.

Another finding is the significant correlation between chunking and comprehension. Chinese L2 readers who made more chunking errors in oral reading performed less well in comprehension. It was in line with prior research which showed that the chunking/prosody performance is an indicator of comprehension (Arcand et al., 2014; Benjamin & Schwanenflugel, 2010; Calet, Gutiérrez-Palma, & Defior, 2017; Fernandes, Querido, & Verhaeghe, 2018; Groen et al., 2019; Klauda & Guthrie, 2008; Rasinski, Reutzel, Chard, & Linan-Thompson, 2011; Pey, Min, & Wah, 2014). The lack of visually marked word space in Chinese renders chunking an explicitly activated process that operates at word level instead of phrase level in alphabetic languages. Word-level chunking in Chinese contributes to comprehension through lexical access (Shen & Jiang, 2013). On the one hand, word boundaries naturally take shape when readers decipher the meaning of individual words within the connected text. Therefore, successful word segmentation to some extent depends on accurate and rapid word recognition. On the other hand, finding the correct word boundary helps pack relevant characters into contextually appropriate words and elicits lexical access, if word recognition fails or ambiguity occurs. Lexical access then provides data for higher-order processes to work on to build comprehension. Inaccurate lexical access may cause difficulty in comprehension, which in turn informs readers of the inappropriateness of an initial segmenting decision and further promotes a re-examination and new combination that leads to a change in meaning.

However, chunking did not add significantly to comprehension when controlling accurate rate, consistent with previous studies that chunking/prosody does not uniquely predict comprehension beyond accurate rate (Riedel, 2007; Sabatini, Wang, & O'Reilly, 2018; Schwanenflugel, Hamilton, Kuhn, Wisenbaker, & Stahl, 2004). The accurate and fluent oral reading requires the synchronization of multiple sub-skills (Breznitz, 2006), including segmenting the text into meaningful chunks and preserving the natural prosody. It is uncommon for one to read rapidly and accurately without appropriate prosody. Hence, it is reasonable that accurate rate captured variance associated with chunking/prosody.

The critical role of the accuracy, speed, and chunking aspects of ORF in comprehension highlights the importance of low-level word recognition and segmentation

processes. Low-level processing competes with high-level processing for the limited cognitive capacity. Lack of efficiency in low-level processing places additional demands on readers' cognitive resources, leaving few for high-level processing to generate comprehension. If readers have difficulty in identifying and segmenting words, their reading will be slow and laborious, which further disrupts the processes of making and maintaining connections among ideas within a text. Consequently, the understanding of textual meaning will be hampered. As readers move toward automatic low-level processing, more cognitive resources will be freed up for text processing to gain better comprehension. For L2 readers, the already limited cognitive resources are devoted to low-level processing and thus little remains for higher-level processes. Comprehension will be challenging as a result.

A very weak relationship between self-correction and comprehension was observed, probably due to the floor effect: self-correction seldom occurred and the individual variability was rather small. The rare self-correction was in parallel with prior studies on Chinese L2 low-level readers (Liu, 1999; Wu & Anderson, 2007). The reason L2 readers did not intend to self-correct might be the failure to either realize the presence of errors or provide a repair. L2 oral reading is constrained by learners' unstable language system and insufficiently developed processing skills (Segalowitz, 2010). With cognitive resources exhausted by recognizing, segmenting, and pronouncing words, L2 readers have no additional attention to detect and fix errors despite sensing the non-comprehension (Francis, 1999). Apart from not having the competence to identify and correct certain errors, L2 readers are also likely to be uncertain whether their oral reading is error free due to their limited metalinguistic awareness (Kormos, 1999). Another possibility is that L2 readers have their own perceptions as to what extent their errors may result in meaning breakdown and are reluctant to correct errors because they still make sense from the text (Kucer, 2017; Share, 1990).

The relationship between ORF and learner-perceived difficulty

Oral reading was viewed as a moderately difficult task ($M=6.47$) and the difficulty rating was negatively correlated with accurate rate performance. Participants who read texts slowly and arduously were more likely to consider the task difficult. The accuracy and speed dimension of ORF marks the efficiency in word recognition, and its relation with learner-perceived difficulty provided evidence that stumbles at word recognition are at the core of reading difficulties (Stanovich, 2000). It has been extensively reported that at-risk L1 readers unexceptionally encountered difficulty in word recognition (Parker et al., 2015). Most of L2 reading problems also lie in the processing of L2-specific linguistic forms such as words (Yamashita, 2001). It was further supported by participant statements that the unrecognized characters/words constituted the main hurdle in oral reading.

The correlation between difficulty rating and chunking performance did not reach significance. It seemed that the implementation of segmenting adjacent

character strings into meaningful words did not affect learners' evaluation of the oral reading's difficulty level. Nevertheless, a number of participants commented that struggling with chunking imposed burden on oral reading, increasing the task's perceived-difficulty.

Another challenge that was frequently mentioned by participants was the hindered pronunciation retrieval. It has been posited that comprehension without recoding is more common in L2 reading (Lems, 2006; Jeon, 2012). Knowledge of a word does not ensure a successful recoding due to the deficiency in pronunciation. It is especially true in Chinese: the lack of script–sound correspondence makes pronunciation retrieval of a word a challenging task (Hu, 2010).

Conclusion

The current study showed that L2 Chinese ORF consisted of accuracy, speed, chunking, and monitoring components. The accuracy, speed, and chunking components were robustly related to comprehension, with the accuracy and speed emerging as the stronger predictor. The accuracy and speed dimension was also an indicator of learner-perceived difficulty of oral reading. The findings offered strong support for ORF as an effective tool to assess comprehension and diagnose reading problems in Chinese L2 reading. The accuracy, speed, and chunking aspects of ORF are sufficiently accurate to reflect comprehension performance and sensitive to detect reading difficulties. The scoring can be conducted promptly and reliably.

Another implication of this research lies in the systematic development of ORF. Given the essential role of ORF in comprehension, ORF training should be incorporated into Chinese L2 reading curricula. Effective pedagogical practices include the specially-designed exercises targeting character/word recognition and segmentation, as well as integrated activities such as reading aloud, extensive reading, and repeated reading.

There are several limitations of the present study. The majority of the participants were at lower proficiency level. It remains a question how the relationship between ORF and comprehension evolves as learners' overall language competence develops. Future studies that recruited advanced learners would generate more insights into the interaction between ORF and comprehension at varying proficiency levels. It is also a worthwhile avenue of further research to probe the benchmark of ORF a reader should reach to obtain adequate comprehension.

Author's contributions The author contributed to the study conception and design. The author performed material preparation, data collection, and analysis, as well as wrote the draft. The author reviewed and approved the final manuscript.

Funding Not applicable.

Availability of data and material The data that support the findings of this study are available on request from the corresponding author, S Yang. The data are not publicly available due to their containing information that could compromise the privacy of research participants.

Code availability Not applicable.

Compliance with ethical standard

Conflict of interest The authors declare that they have no conflict of interests.

Appendix A: Character Recognition Test

	1	2	3	4	5	6	7	8	9	10
1	了	利	路	片	网	免	钟	稳	符	辩
2	能	门	清	专	创	穿	饭	途	厚	矿
3	法	平	服	士	夜	顾	博	灯	咱	肩
4	但	务	击	严	游	野	孙	圆	恨	撒
5	三	展	示	米	衣	缺	祖	阻	暂	昨

Appendix B: Text Reading Example

我下午来找你，可是你不在家。我想邀请你明天跟我去参加我校举办的一年一度的郊游野餐，每位参加者要带一份十元以内的小礼物。餐后有抽奖活动，我想一定很好玩。你回来后尽快给我打个电话。

I came for you this afternoon, but you were not at home. I want to invite you to participate in the annual picnic held by my school tomorrow. Each participant needs to bring a small gift with a value less than 10 RMB. After the picnic there will be a lottery draw. I think it must be fun. Please call me as soon as you come back.

Appendix C: Questionnaire

Please rate the difficulty level of the oral reading task from 1 to 10: 1 as the least difficult and 10 as the most difficult. Please explain what aspect of the task you think is difficult.

1 2 3 4 5 6 7 8 9 10.

Which aspect is difficult? Explain:

References

- Álvarez-Cañizo, M., Suárez-Coalla, P., & Cuetos, F. (2018). Reading prosody development in Spanish children. *Reading and Writing: An Interdisciplinary Journal*, *31*, 35–52. <https://doi.org/10.1007/s11145-017-9768-7>.
- Arcand, M.-S., Dion, E., Lemire-Théberge, L., Guay, M.-H., Barrette, A., Gagnon, V., et al. (2014). Segmenting texts into meaningful word groups: Beginning readers' prosody and comprehension. *Scientific Studies of Reading*, *18*, 208–223. <https://doi.org/10.1080/10888438.2013.864658>.
- Arnesen, A., Braeken, J., Baker, S., Meek-Hansen, W., Ogden, T., & Melby-Lervåg, M. (2016). Growth in oral reading fluency in a semitransparent orthography: Concurrent and predictive relations with reading proficiency in Norwegian, grades 2–5. *Reading Research Quarterly*, *52*, 177–201. <https://doi.org/10.1002/rrq.159>.
- Bassetti, B. (2005). Effects of writing systems on second language awareness: Word awareness in English learners of Chinese as a foreign language. In V. Cook & B. Bassetti (Eds.), *Second language writing systems* (pp. 335–356). Clevedon: Multilingual Matters.
- Beatty, L., & Care, E. (2009). Learning from their miscues: Differences across reading ability and text difficulty. *Australian Journal of Language and Literacy*, *32*, 226–224.
- Benjamin, R. G., & Schwanenflugel, P. J. (2010). Text complexity and oral prosody in young readers. *Reading Research Quarterly*, *45*, 388–404. <https://doi.org/10.1598/RPQ.45.4.2>.
- Berninger, V. W., et al. (2010). Applying the multiple dimensions of reading fluency to assessment and instruction. *Journal of Psychoeducational Assessment*, *28*, 3–18. <https://doi.org/10.1177/0734282909336083>.
- Binder, K. S., Tighe, E., Jiang, Y., Kaftanski, K., Qi, C., & Ardoin, S. P. (2013). Reading expressively and understanding thoroughly: An examination of prosody in adults with low literacy skills. *Reading and Writing: An Interdisciplinary Journal*, *26*, 665–680. <https://doi.org/10.1007/s11145-012-9382-7>.
- Breznitz, Z. (2006). *Fluency in reading: Synchronization of process*. Mahwah: Lawrence Erlbaum.
- Briceno, A., & Klein, A. F. (2018). Running records and first grade English learners: An analysis of language related errors. *Reading Psychology*, *39*, 335–361. <https://doi.org/10.1080/02702711.2018.1432514>.
- Calet, N., Gutiérrez-Palma, N., & Defior, S. (2017). Effects of fluency training on reading competence in primary school children: The role of prosody. *Learning and Instruction*, *52*, 59–68. <https://doi.org/10.1016/j.learninstruc.2017.04.006>.
- Chang, L. (2015). Proficient readers' reading behaviors in Taiwan: The study of young Chinese readers. *Universal Journal of Educational Research*, *3*, 283–287.
- Crosson, A. C., & Lesaux, N. K. (2010). Revisiting assumptions about the relationship of fluent reading to comprehension: Spanish-speakers' text reading fluency in English. *Reading and Writing: An Interdisciplinary Journal*, *23*, 475–494. <https://doi.org/10.1007/s11145-009-9168-8>.
- D'Agostino, J. V., Kelly, R. H., & Rodgers, E. (2019). Self-corrections and the reading progress of struggling beginning readers. *Reading Psychology*, *40*, 525–550. <https://doi.org/10.1080/02702711.2019.1629518>.
- DeKeyser, R. M. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125–151). Cambridge: Cambridge University Press.
- Fernandes, S., Querido, L., & Verhaeghe, A. (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*, S102–S129. <https://doi.org/10.1111/1467-9817.12248>.
- Francis, N. (1999). Self-correction patterns and metalinguistic awareness: A proposed typology for studying text-processing strategies of proficient readers. *Journal of Research in Reading*, *22*, 304–310. <https://doi.org/10.1111/1467-9817.00094>.
- Fuchs, D., Fuchs, L. S., & Compton, D. L. (2004). Identifying reading disability by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly*, *22*, 216–222. <https://doi.org/10.2307/1593674>.
- Godde, E., Bosse, M., & Bailly, G. (2020). A review of reading prosody acquisition and development. *Reading and Writing: An Interdisciplinary Journal*, *33*, 399–426. <https://doi.org/10.1007/s11145-019-09968-1>.

- Goodman, Y. M., & Goodman, K. S. (1994). To err is human: Learning about language processes by analyzing miscues. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (pp. 104–123). Newark: International Reading Association.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge: Cambridge University Press.
- Grabe, W. (2010). Fluency in reading: Thirty five years later. *Reading in a Foreign Language*, 22, 71–83.
- Groen, M. A., Veenendaal, N. J., & Verhoeven, L. (2019). The role of prosody in reading comprehension: Evidence from poor comprehenders. *Journal of Research in Reading*, 42, 37–57. <https://doi.org/10.1111/1467-9817.12133>.
- Hintze, J. M., & Silbergitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review*, 34, 372–386.
- Hu, B. (2010). The challenge of Chinese: a preliminary study of UK learners' perceptions of difficulty. *Language Learning Journal*, 38, 99–118. <https://doi.org/10.1080/09571731003620721>.
- Jenkins, J. R., Fuchs, L. S., Van den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, 95, 719–729. <https://doi.org/10.1037/0022-0663.95.4.719>.
- Jeon, E. H. (2012). Oral reading fluency in second language reading. *Reading in a Foreign Language*, 24, 186–208.
- Jiang, X. (2016). The role of oral reading fluency in ESL reading comprehension among learners of different first language backgrounds. *Reading Matrix: An International Online Journal*, 16, 227–242.
- Jiang, X., Sawaki, Y., & Sabatini, J. (2012). Word reading efficiency, text reading fluency, and reading comprehension among Chinese learners of English. *Reading Psychology*, 33, 323–349. <https://doi.org/10.1080/02702711.2010.526051>.
- Keh, M. L. (2016). Understanding and evaluating English learners' oral reading with miscue analysis. *Journal of Adolescent and Adult Literacy*, 60, 643–653. <https://doi.org/10.1002/jaal.625>.
- Kim, Y. G., Vorstius, C., & Radach, R. (2018). Does online comprehension monitoring make a unique contribution to reading comprehension in beginning readers? Evidence from eye movements. *Scientific Studies of Reading*, 22, 367–383. <https://doi.org/10.1080/10888438.2018.1457680>.
- Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology*, 100, 310–321. <https://doi.org/10.1037/0022-0663.100.2.310>.
- Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning*, 49, 303–342. <https://doi.org/10.1111/0023-8333.00090>.
- Kucer, S. B. (2009). Examining the relationship between text processing and text comprehension in fourth grade readers. *Reading Psychology*, 30, 340–358. <https://doi.org/10.1080/02702710802411604>.
- Kucer, S. B. (2017). The monitoring and responding behaviors of proficiency fourth grade readers to miscues on a complex scientific text. *Literacy*, 51, 154–161. <https://doi.org/10.1111/lit.12116>.
- Kuhn, M. R., & Schwanenflugel, P. J. (2018). Prosody, pacing, and situational fluency. *Journal of Adolescent & Adult Literacy*, 62, 363–368. <https://doi.org/10.1002/jaal.867>.
- 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, 230–251. <https://doi.org/10.1598/RPQ.45.2.4>.
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293–323. [https://doi.org/10.1016/0010-0285\(74\)90015-2](https://doi.org/10.1016/0010-0285(74)90015-2).
- Laing, S. P. (2002). Miscue analysis in school-age children. *American Journal of Speech-Language Pathology*, 11, 407–416. <https://doi.org/10.1044/1058-0360>.
- Lee-Thompson, L. C. (2008). An investigation of reading strategies applied by American learners of Chinese as a foreign language. *Foreign Language Annals*, 41, 702–721. <https://doi.org/10.1111/j.1944-9720.2008.tb03326.x>.
- Lems, K. (2006). Reading fluency and comprehension in adult English language learners. In T. Rasinski, C. Blachowicz, & K. Lems (Eds.), *Fluency instruction: Research based best practices* (pp. 231–252). New York: Guilford Press.
- Li, X., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology*, 58, 525–552. <https://doi.org/10.1016/j.cogpsych.2009.02.003>.
- Liu, S. (1999). A miscue analysis of nine Japanese-speaking students of Chinese. *Language Teaching and Linguistics Studies*, 2, 97–109.

- Lv, C. (2016). Chinese reading development among young learners in a Chinese immersion program and a Chinese heritage language school. *Chinese Teaching in the World*, 30, 550–562.
- McGee, L. M., Kim, H., Nelson, K. S., & Fried, M. D. (2015). Change over time in first graders' strategic use of information at point of difficulty in reading. *Reading Research Quarterly*, 50, 263–291. <https://doi.org/10.1002/rrq.98>.
- McTague, B., Lems, K., Butler, D., & Carmona, E. (2012). ELL fluency scores: What can they tell us? In T. Rasinski, C. Blachowicz, & K. Lems (Eds.), *Fluency instruction: Research-based best practices* (pp. 278–288). New York: Guilford Press.
- Miller, J., & Schwanenflugel, P. J. (2006). Prosody of syntactically complex sentences in the oral reading of young children. *Journal of Educational Psychology*, 98, 839–843. <https://doi.org/10.1037/0022-0663.98.4.839>.
- Morris, D., Pennell, A. M., Perney, J., & Trathen, W. (2018). Using subjective and objective measures to predict level of reading fluency at the end of first grade. *Reading Psychology*, 39, 253–270. <https://doi.org/10.1080/02702711.2017.1418466>.
- National Working Committee for Language and Characters. (1992). *Modern Chinese character frequency list*. Beijing: Language & Culture Press.
- Nguyen, T. Q., Pickren, S. E., Saha, N. M., & Cutting, L. E. (2020). Executive functions and components of oral reading fluency through the lens of text complexity. *Reading and Writing: An Interdisciplinary Journal*, 33, 1037–1073. <https://doi.org/10.1007/s11145-020-10020-w>.
- Parker, D. C., Zaslofsky, A. F., Burns, M. K., Kanive, R., Hodgson, J., Scholin, S. E., et al. (2015). A brief report of the diagnostic accuracy of oral reading fluency and reading inventory levels for reading failure risk among second- and third-grade students. *Reading and Writing Quarterly*, 31, 56–67. <https://doi.org/10.1080/10573569.2013.857970>.
- Pey, K. C., Min, L. H., & Wah, L. K. (2014). Relationship between oral reading fluency and reading comprehension among ESL students. *GEMA Online Journal of Language Studies*, 14, 19–32. <https://doi.org/10.17576/GEMA-2014-1403-02>.
- Rasinski, T. V., Reutzel, D. R., Chard, D., & Linan-Thompson, S. (2011). Reading fluency. In M. L. Kamil, P. D. Pearson, E. BirrMoje, & P. P. Afflerbach (Eds.), *Handbook of reading research: Volume IV* (pp. 286–319). New York: Routledge.
- Rasinski, T. V., & Samuels, S. J. (2011). Reading fluency: What it is and what it is not. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about reading instruction* (pp. 94–114). Newark: International Reading Association.
- Riedel, B. W. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban, first grade students. *Reading Research Quarterly*, 42, 460–466. <https://doi.org/10.1598/RPQ.42.4.5>.
- Roehrig, A. D., Petscher, Y., Nettles, S. M., Hudson, R. F., & Torgesen, J. K. (2008). Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology*, 46, 343–366. <https://doi.org/10.1016/j.jsp.2007.06.006>.
- Sabatini, J., Wang, Z., & O'Reilly, T. (2018). Relating reading comprehension to oral reading performance in the NAEP Fourth-grade special study of oral reading. *Reading Research Quarterly*, 54, 253–271. <https://doi.org/10.1002/rrq.226>.
- Schilling, S. G., Carlisle, J. F., Scott, S. E., & Zeng, J. (2007). Are fluency measures accurate predictors of reading achievement? *The Elementary School Journal*, 107, 429–448. <https://doi.org/10.1086/518622>.
- Schwanenflugel, P., Hamilton, M., Kuhn, M., Wisenbaker, J., & Stahl, S. (2004). Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers. *Journal of Educational Psychology*, 96, 119–129. <https://doi.org/10.1037/0022-0663.96.1.119>.
- Segalowitz, N. S. (2003). Automaticity and second language learning. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 382–408). Oxford: Blackwell.
- Segalowitz, N. S. (2010). *Cognitive bases of second language fluency*. New York: Routledge.
- Segalowitz, N. S., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics*, 14, 369–385. <https://doi.org/10.1017/S0142716400010845>.
- Share, D. L. (1990). Self-correction rates in oral reading: Indices of efficient reading or artifact of text difficulty? *Educational Psychology*, 10, 181–187. <https://doi.org/10.1080/0144341900100205>.
- Shen, H. H. (2008). An analysis of word decision strategies among learners of Chinese. *Foreign Language Annals*, 41, 501–524. <https://doi.org/10.1111/j.1944-9720.2008.tb03309.x>.

- Shen, H. H. (2013). Chinese L2 literacy development: Cognitive characteristics, learning strategies, and pedagogical interventions. *Language and Linguistics Compass*, 7, 371–387. <https://doi.org/10.1111/lnc3.12034>.
- Shen, H. H., & Jiang, X. (2013). Character reading fluency, word segmentation accuracy, and reading comprehension in L2 Chinese. *Reading in a Foreign Language*, 25, 1–25.
- Silbergliitt, B., Burns, M. K., Madyun, N. H., & Lail, K. E. (2006). Relationship of reading fluency assessment data with state accountability test scores: A longitudinal comparison of grade levels. *Psychology in the Schools*, 43, 527–535. <https://doi.org/10.1002/pits.20175>.
- Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Press.
- Timmer, K., & Schiller, N. O. (2012). The role of orthography and phonology in English: An ERP study on first and second language reading aloud. *Brain Research*, 1483, 39–53. <https://doi.org/10.1016/j.brainres.2012.09.004>.
- Valencia, S. W., Smith, A. T., Reece, A. M., Li, M., Wixson, K. K., & Newman, H. (2010). Oral reading fluency assessment: Issues of construct, criterion, and consequential validity. *Reading Research Quarterly*, 45, 270–291. <https://doi.org/10.1598/RPQ.45.3.1>.
- Valle, A., Binder, K. S., Walsh, C. B., Nemier, C., & Bangs, K. E. (2013). Eye movements, prosody, and word frequency among average- and high-skilled second-grade readers. *School Psychology Review*, 42, 171–190. <https://doi.org/10.1080/02796015.2013.12087483>.
- Wang, S. (2006). *A socio-psycholinguistic study on L2 Chinese readers' behavior while reading orally*. (Unpublished doctoral dissertation). University of Arizona, USA.
- Wanzek, J., Roberts, G., Linan-Thompson, S., Vaughn, S., Woodruff, A., & Murray, C. S. (2010). Differences in the relationship of oral reading fluency and high-stakes measures of reading comprehension. *Assessment for Effective Intervention*, 35, 67–77. <https://doi.org/10.1177/1534508409339917>.
- Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education*, 41, 85–120. <https://doi.org/10.1177/00224669070410020401>.
- Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., Wolf, M., & Schwanenflugel, P. (2010). The relationship between different measures of oral reading fluency and reading comprehensions in second-grade students who evidence different oral reading fluency difficulties. *Language, Speech, and Hearing Services in Schools*, 41, 340–348. <https://doi.org/10.1044/0161-1461>.
- Wu, X., & Anderson, R. C. (2007). Reading strategies revealed in Chinese children's oral reading. *Literacy Teaching and Learning*, 12, 47–72.
- Yamashita, J. (2001). Transfer of L1 reading ability to L2 reading: An elaboration of the linguistic threshold. *Studies in Language and Culture*, 23, 189–200. <https://doi.org/10.18999/stulc.23.1.189>.
- Yan, Q., & Wang, J. (2011). Investigating the miscue-reflected EFL oral reading process: A case study. *Chinese Journal of Applied Linguistics (Quarterly)*, 34, 62–82. <https://doi.org/10.1515/cjal.2011.015>.
- Yang, W., & Jiang, L. (2012). Segmentation ambiguity in Chinese words by elementary and intermediate learners of Chinese: An experimental study. *Language Teaching and Linguistics Studies*, 3, 10–17.
- Zhang, H. (2018). The influence of different L2 Chinese proficiency measurements on the results of CSL research. *Language Teaching and Linguistics Studies*, 6, 14–23.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.