A Corpus-Based Analysis of Text Complexity in NMET Reading and Its Implication on Reading Instruction

Liyan Huang, Jiaying Wang, and Xiaochen Hua(✉)

Language Cognition and Assessment Centre, School of Foreign Studies, South China Normal University, No. 55, Zhongshan Avenue West, Guangzhou, Guangdong, China
frankhxc@scnu.edu.cn

Abstract. This study builds a small-scale corpus of 54 reading texts of the National Matriculation English Test from 2016 to 2019 and uses natural language processing tools, Python and Coh-Metrix, to investigate text complexity with 35 quantitative indices from five aspects, including length, readability, vocabulary, syntax and cohesion. Results reveal the major characteristics of these texts, showing they have appropriate length, similar difficulty as the texts in high school English textbooks and a higher level of lexical diversity and stem overlap among all sentences than those in textbooks. The findings offer support for more precise and effective reading instructional practice in high school from the perspectives of the WHAT and the HOW in classroom teaching.

Keywords: Text complexity · NMET reading · Reading instruction in high school · Python · Coh-Metrix

1 Introduction

Reading is an essential competence for EFL learners. It is a challenge for many Chinese EFL learners, and the reasons are two-fold. First, reading comprehension comprises the largest portion of almost every English high-stakes test in China, e.g., the National Matriculation of English Test (hereafter, NMET). Second, reading authentic English texts is the most common approach used by Chinese EFL learners in both rural and urban areas to gain access to examples of authentic English. Reading instruction is therefore an extremely important component of English teaching in schools and so is the promotion of teaching effectiveness. In order to help students learn in the best possible way, teachers should keep many points in mind while designing and implementing a reading class, e.g., setting teaching goals, selecting appropriate materials, and employing effective teaching strategies. Currently, teachers are expected to adhere to the principles and requirements published as the National English Curriculum Standards (Ministry of Education of the People's Republic of China 2018) while planning teaching. The literature on the topic of turning the Standards into reading instructional practice is quite voluminous (e.g., Wang et al. 2019); however, few studies have analyzed the NMET and drawn out any

L.-K. Lee et al. (Eds.): ICTE 2020, CCIS 1302, pp. 23–34, 2020.
https://doi.org/10.1007/978-981-33-4594-2_3

implications for classroom teaching. Consequently, in this study we focus on the reading component of the NMET and conduct a corpus-based analysis of one of the test text's features, i.e., text complexity, and put forward a series of suggestions regarding the WHAT and the HOW issues of reading instruction in high school based on the statistical results.

2 Reading and Reading Instruction

Reading as an essential competence, both its instruction and its assessment, have long been researched. By its nature, the concept of reading is goal-directed, multidimensional and developmental (Fox and Alexander 2011). Specifically, reading is defined as a meaning-making process initiated by a reader's intention to interact with texts, which requires an integration of cognitive, motivational and socio-contextual factors, and it is also a dynamic accumulative development process across a reader's lifespan (Alexander 2012; Pearson 2009). Research on second language (L2) reading "has evolved both as an extension of first language (L1) reading and as a branch of second language acquisition" (Koda 2012, p. 304). Verhoeven (2011) pinpointed that L2 reading involves a reader's linguistic knowledge (e.g., vocabulary, grammar), the proficient use of a complex set of reading skills in both cognitive processes (e.g., working memory, schema, motivation) and social processes (e.g., technology-assisted learning, communicative learning). L2 reading also relates to the text itself in that rhetoric, genre and text complexity (e.g., lexical, syntactical, semantic features) may influence a reader's understanding (Hornof 2008). In this complex process, L2 reading involves three major operations: (a) extracting linguistic information from written language, (b) decoding the extracted information into words, sentences and paragraphs, and (c) mapping the new information onto prior knowledge (Koda 2012).

EFL learners' reading competence may gradually improve through chronic practice, with individual readers, especially struggling learners, requiring customized reading instruction to facilitate the development of their reading comprehension (Pearson 2009). Reading instruction generally includes macro reading models and micro reading strategies. Reading strategies are those deliberate, goal-directed skills such as inferencing, summarizing and comparing that attempt to support the reader's efforts to recognize words, process sentences and construct meanings out of text (Paris et al. 1984). Reading models are packages or structures of a series of reading strategies in a specific order. Popular reading models such as the top-down model, the bottom-up model, the interactive model, etc. have long been investigated to help students improve their reading competence (Shrum and Glisan 2010). Implementing such a reading pedagogy, however, does not necessarily guarantee every student success in EFL reading competence.

As reading comprehension is the consequence of an extended amount of engaged reading (Guthrie et al. 2013), selecting reading materials that will engage EFL readers plays a pivotal role in both teaching and self-learning. The criteria or guidelines for selecting such readings for EFL readers stress factors such as the reader's prior knowledge, topics of interest, cultural differences, value of information and text complexity (Guthrie et al. 2013). Especially for intermediate level readers or middle school students, there is also the factor of text complexity which is intrinsic to reading text in a

foreign language (Laufer and Nation 2012). Therefore, to deepen our understanding of L2 reading and to contribute to the development of EFL reading models, it is necessary to draw on implications from the analysis of text complexity in reading materials. In light of the widely noted contribution made by the study of syntactic complexity to the development of writing instruction (e.g., Lu 2017), the current study is an attempt to propose modifications to the models of contemporary reading instruction based on our findings of text complexity.

3 Text Complexity

Broadly speaking, text complexity refers to the inherent difficulty of reading and comprehending a text combined with consideration of reader and task variables (NGA/CCSSO 2010). In a narrower sense, it is conceived as textual elements or factors that can be analyzed, studied or manipulated (Mesmer et al. 2012). In the current study, we use this term in its narrower sense.

Text complexity has been considered a critical task-related variable that may affect test-takers' reading performance, along with other variables such as genre (e.g., narratives vs. informational tests) (Gardner 2004) and topic (e.g., a more vs. less familiar topic) (Pulido 2004).

Previous studies of the effect of text complexity have examined three features of individual texts, namely words, syntax and discourse structure (Mesmer et al. 2012). Empirical studies have documented a strong relationship between vocabulary knowledge and comprehension, finding that vocabulary difficulty has influence on the understanding of text by L1 readers as well as L2 readers (Wright and Cervetti 2016). Regarding syntax, researchers have considered the ability to parse sentence syntax an essential component of reading comprehension (e.g., Graesser et al. 1996). It has been reported that the opacity and heaviness of the constituent structure of sentences make texts harder to process because readers might have difficulty in parsing the sentences (Berman 1984). Regarding discourse structure, previous investigations of text complexity have been limited to text length and cohesion. Hiebert (1999) observed that text length was one of the features that distinguished between texts of different difficulty. Several studies have reported that increasing the cohesion of a text makes it easier for readers to comprehend it (Graesser et al. 2003).

Regarding studies of text complexity of NMET texts set for reading, previous corpus-based studies have mainly examined lexical complexity. Jin et al. (2017) investigated the lexical profiles of reading texts in the NMET on a large scale, using Nation's (2006) fourteen 1K word-family lists based on the British National Corpus. They established two sets of benchmarks using the 95% and 98% text coverage targets, which provides implications for selecting and adapting reading texts in high-stakes tests. Making use of the corpus tools Wordsmith and Treetagger, Wang (2018) reported the lexical complexity of reading comprehension texts set by NMET (Jiangsu Province version) from three aspects, i.e., type-token ratio, vocabulary distribution and function words, finding that more attention should be paid to the width of vocabulary, lexical chunks and function words. Other studies have evaluated the quality of the test design of NMET's reading section by investigating text complexity on the basis of text length and readability (e.g.,

Peng et al. 2017). Nevertheless, the text complexity of the reading section in NMET has not yet been examined fully, especially for syntax and discourse structure. With the current study we aim to fill these research gaps.

4 The Present Study

In this study we investigate the text complexity of the reading component in the NMET. Specifically, we seek to address the two research questions: What are the characteristics of the text complexity of the NMET's reading section? WHAT does the analysis of text complexity tell us and HOW should reading instruction at high school change?

4.1 The Corpus

The current study built a small-scale corpus of 54 reading texts of the NMET, including a nationwide version and a local version (i.e., Zhejiang Province version) for the academic years 2016 to 2018.

4.2 Data Collection and Analysis

We analyzed the text complexity of reading texts in the NMET corpus with 35 measures of text length, readability, lexical complexity, syntactic complexity and cohesion, using the natural language processing tools Python and Coh-Metrix (see below). We chose these features because they have been reported to have an effect on the reading process of language learners (Alderson 2000) and on the performance of test-takers in the testing context (Bachman and Palmer 1996). These measures comprehensively reflect the text complexity of reading texts in vocabulary, syntax and discourse, providing a theoretical framework for the study of text complexity.

As mentioned above, we used two tools, Python and Coh-Metrix, because they are complementary: Python for the analysis of text length, readability and lexical complexity and Coh-Metrix for the analysis of syntactic complexity and cohesion of texts (McNamara et al. 2014). The advantage of using Python was that it integrated various functions of text normalization (e.g., cleaning texts, removing special characters) and textual analysis (e.g., counting words, computing formulas), while Coh-Metrix provided a series of reliable measures of syntactic complexity and cohesion of texts, thus having advantages at syntax and discourse levels. In addition, we included Coh-Metrix L2 readability for the reason that it was designed especially for L2 learners. The measures of text complexity and procedure were as follows (Table 1).

Text length was assessed via six measures. Before the counting of text length, texts were normalized through Python, for instance, via text tokenization, text cleaning, removing of special characters and case conversions.

Readability was assessed via three measures, including two traditional readability formulas based on sentence length and average number of syllables in words, as well as an L2 readability formula based on word frequency, syntactic similarity and referential cohesion.

Table 1. Measures of text complexity

Category	Measures
Text length	Number of words (excluding numbers or punctuation)
	Number of sentences
	Number of paragraphs
	Length of words (average number of letters in words)
	Length of sentences (average number of words in sentences)
	Length of paragraphs (average number of sentences in paragraphs)
Readability	Flesch Reading Ease
	Flesch-Kincaid Grade Level
	Coh-Metrix L2 Readability
Lexical complexity	Lexical density (ratio of content words to function words)
	Lexical diversity (type-token ratio)
	Vocabulary coverage of the British National Corpus (1K word family)
	Vocabulary coverage of the British National Corpus (2K word family)
	Vocabulary coverage of the British National Corpus (3K word family)
	Vocabulary coverage of the NMET syllabus glossary
Syntactic complexity	Number of words before main verb
	Number of modifiers per noun phrase
	Minimal Edit Distance based on part of speech
	Minimal Edit Distance based on all words
	Minimal Edit Distance based on lemmas
	Sentence syntax similarity between adjacent sentences
	Sentence syntax similarity of all combinations
Cohesion	Noun overlap between adjacent sentences
	Argument overlap between adjacent sentences
	Stem overlap between adjacent sentences
	Content word overlap between adjacent sentences
	Noun overlap of all sentences
	Argument overlap of all sentences
	Stem overlap of all sentences
	Content word overlap of all sentences
	LSA overlap between adjacent sentences
	LSA overlap of all sentences

(*continued*)

Table 1. (*continued*)

Category	Measures
	LSA overlap between adjacent paragraphs
	LSA Given-New
	Incidence score (occurrence per 1,000 words) for all connectives

Lexical complexity was assessed via six measures. Among them, vocabulary coverage of the British National Corpus (BNC) refers to the ratio of words belonging to the first 3K word families to words in the text and vocabulary coverage of the NMET syllabus glossary refers to the ratio of words belonging to the NMET glossary to words in the text. To be specific, vocabulary coverage of the BNC was calculated on the basis of three vocabulary levels of the fourteen 1K word-family lists based on the BNC 14K lists (Nation 2006), namely 1K, 2K and 3K word families, eliminating the effect of proper nouns such as personal names or place names. The BNC 14K list was based on the concept of word family, which was deemed suitable as a unit of measurement for the vocabulary requirements of receptive skills such as listening and reading (Nation and Beglar 2007). In order to maintain consistent statistical results, to calculate the vocabulary coverage of the NMET syllabus glossary we established a corresponding relation between the NMET syllabus glossary and the BNC 14K list on the basis of word family, eliminating the effect of proper nouns.

Syntactic complexity, that is, the degree of sophistication and variation of the structures produced, has been operationalized in many ways (Lu 2017). We adopted seven indices incorporated in Coh-Metrix. A higher value in the first four measures is associated with a higher degree of syntactic sophistication, whereas a higher degree of sentence syntax similarity is associated with a lower degree of syntactic variation (McNamara et al. 2014). Besides, a combination of semantic and syntactic dissimilarity measuring the uniformity and consistency of sentence construction in the text was based on the notion of a Minimal Edit Distance (McCarthy et al. 2009). It calculates the average minimal edit or the distance that parts of speech, words or lemmas are from one another between consecutive sentences in a text.

Cohesion features are explicit characteristics in a text that help create cohesive links between ideas and clauses (McNamara et al. 2014). We assessed the cohesion of the texts using eight referential cohesion indices, four Latent Semantic Analysis (LSA) indices and one incidence score in Coh-Metrix. Referential cohesion refers to overlap in content words between local sentences, or coreference, which is a linguistic cue that can aid readers in making connections between propositions, clauses and sentences in their understanding (Halliday and Hasan 1976). Referential gaps can occur when the words or concepts in a sentence do not overlap with other sentences in the text, affecting readers' comprehension and reading time according to their abilities (O'Reilly and McNamara 2007). LSA provides measures of semantic overlap between sentences or between paragraphs (Landauer et al. 2007). The LSA indices range from 0 to 1, with a higher value indicating greater cohesion (McNamara et al. 2014). Coh-Metrix also provides an incidence score (occurrence per 1,000 words) for all connectives, which

play an important role in the cohesive links between ideas and clauses and provide clues about text organization (Cain and Nash 2011).

In addition, given the fact that no systematic or complete benchmarks were proposed for all the measures of text complexity, some of the measures used the values of a small-scale corpus of high school students' textbooks for reference. In order to establish the benchmarks of the textbook corpus, two experienced teachers were invited to select reading texts with similar difficulty as set by the NMET from high school students' English textbooks (People's Education Press version), including one text of practical writing, one of expository writing, one of narrative writing and one of argumentative writing. The text complexity data of 35 measures covering text length, readability, lexical complexity, syntactic complexity and cohesion were obtained through the descriptive statistics of SPSS (23.0) and used as benchmarks in a comparison of the NMET reading text corpus to determine the level of text complexity of the NMET, which is discussed in the next section.

5 Results and Discussion

The results of text complexity were as follows (Table 2).

For text length, according to the NMET syllabus and the text length requirements in reading tests put forward by scholars such as Alderson (2000), the reading texts of the NMET were of relatively moderate text length.

Based on the result, the text difficulty of the reading section in the NMET was near Grade 10 level of native speakers. Compared with the L2 Readability of the textbook corpus ($M = 16.70$, $SD = 8.25$), the reading texts of the NMET were more difficult than the textbooks for high school students.

As regards lexical complexity, the Mann-Whitney U test showed that the lexical density of NMET reading texts was not significantly different from that of the textbooks ($p = 0.41$) but the lexical diversity of the NMET reading texts was significantly higher than that of the textbooks ($p = 0.00$). The vocabulary coverage of BNC 3K word families was higher than 95%, indicating that a large number of words used in the reading texts of the NMET were high-frequency words and the texts would not present a heavy burden for test takers' reading (Laufer 1989). Furthermore, the vocabulary coverage of the NMET syllabus glossary exceeded 96%, showing that the NMET reading texts were based on the NMET syllabus glossary.

Among the results of syntactic complexity, a Mann-Whitney U test showed that only one measure, the Minimal Edit Distance based on all words, of the NMET reading texts had a significantly higher value than that of the textbook corpus ($p = 0.03$); other measures showed no significant difference. It indicated that the syntactic complexity of the NMET reading texts was at a similar level as high school textbooks.

The results of cohesion features showed that, with the textbook corpus assessed via a Mann-Whitney U test, the NMET reading texts had a significantly higher value only in one measure, that is, stem overlap of all sentences ($p = 0.04$), while other measures revealed no significant difference. This showed that the levels of cohesion features of the NMET reading texts were close to those of the textbooks of high school students, implying that the cohesion features of the NMET reading texts were suitable for testing high school students.

Table 2. Results of text complexity

Measures	*N*	*M*	*SD*
Number of words	54	282.39	10.00
Number of sentences	54	16.31	2.08
Number of paragraphs	54	4.94	0.83
Length of words	54	4.62	0.40
Length of sentences	54	18.59	2.40
Length of paragraphs	54	3.52	0.45
Flesch Reading Ease	54	58.65	5.76
Flesch-Kincaid Grade Level	54	9.70	1.27
Coh-Metrix L2 Readability	54	14.51	3.56
Lexical density	54	0.83	0.34
Lexical diversity	54	0.61	0.03
Vocabulary coverage of the BNC (1K word family)	54	0.83	0.03
Vocabulary coverage of the BNC (2K word family)	54	0.93	0.01
Vocabulary coverage of the BNC (3K word family)	54	0.96	0.01
Vocabulary coverage of the NMET syllabus glossary	54	0.96	0.01
Number of words before main verb	54	3.96	0.91
Number of modifiers per noun phrase	54	0.89	0.10
Minimal Edit Distance based on part of speech	54	0.68	0.03
Minimal Edit Distance based on all words	54	0.91	0.02
Minimal Edit Distance based on lemmas	54	0.89	0.02
Sentence syntax similarity between adjacent sentences	54	0.09	0.02
Sentence syntax similarity of all combinations	54	0.09	0.01
Noun overlap between adjacent sentences	54	0.32	0.10
Argument overlap between adjacent sentences	54	0.47	0.09
Stem overlap between adjacent sentences	54	0.42	0.10
Content word overlap between adjacent sentences	54	0.08	0.02
Noun overlap of all sentences	54	0.28	0.09
Argument overlap of all sentences	54	0.40	0.07
Stem overlap of all sentences	54	0.36	0.11
Content word overlap of all sentences	54	0.06	0.01
LSA overlap between adjacent sentences	54	0.18	0.04
LSA overlap of all sentences	54	0.17	0.05

(*continued*)

Table 2. (*continued*)

Measures	*N*	*M*	*SD*
LSA overlap between adjacent paragraphs	54	0.33	0.06
LSA Given-New	54	0.29	0.03
Incidence score for all connectives	54	82.56	8.67

6 Suggestions for Reading Instruction at High School

In this study, we took the NMET reading section as an example of the reading task and analyzed it for 35 features such as text length, readability, lexical complexity, syntactic complexity and cohesion with the natural language processing tools Python and Coh-Metrix. Our results indicate that the reading section of the NMET had proper text length, meeting the requirements of the NMET syllabus and general reading tests; however, it was of somewhat greater difficulty than high school textbooks. In terms of lexical complexity, the vocabulary in the reading section of the NMET was mostly high-frequency words and the vocabulary of the NMET syllabus glossary was also mostly of high-frequency words. In terms of lexical diversity, Minimal Edit Distance based on all words of the NMET reading texts and one measure of cohesion (i.e., stem overlap of all sentences) showed higher standards than usual in the NMET reading section. Lexical density, syntactic complexity and other cohesion measures showed no significant difference with the teaching materials of high school students in China.

Based on the summarized characteristics of the text complexity of the NMET's reading section, we, thus, believed that the answer of our first question can cast light on our further inquires. That is, what and how the analysis of text complexity can enlighten the reform of reading instruction in Chinese high schools? In the followed sections, we proposed specific implications.

6.1 The WHAT

Local education departments generally suggest that schools in the same district use the same textbook. However, due to a mismatch between existing textbooks and contextual factors such as teachers' belifes of teaching, students' competence level and individual differences, it is not unusual for teachers wanting to develop self-edited materials as supplementary teaching resources when preparing lessons. Moreover, the findings of the present study indicate that there are some differences between NMET reading texts and the reading texts in the students' textbook in terms of text complexity. Therefore, a textbook cannot be the only source for daily teaching and NMET preparation. To improve students' English competence and their performance in the NMET, teachers are advised to take account of the NMET reading texts while designing classroom instruction.

Given that developing criteria for evaluating, selecting and adapting materials is fundamental, the results of the analysis of text complexity in the study served the aforementioned issue in two aspects. First of all, it is an indispensable part of designing a reading class to set teaching goals, and to select and adapt the reading texts. Teachers

might refer to our framework of text complexity, which was used to analyze 35 features of complexity, such as text length, readability, lexical complexity, syntactic complexity and cohesion, to help evaluate the quality of the reading texts and then select and edit them. Traditionally, reading instruction in high school aims at explaining vocabulary, gist, details and syntax while cohesion is ignored. To reduce the students' processing burden, it is advised that teachers should include lexical complexity, syntactic complexity and cohesion in their teaching. With regard to the selection and adaption of teaching materials, teachers might refer to the statistical results of this study and develop a set of criteria. For example, the number of words in each text could be around 280, the Flesch Reading Ease score could be approximately 60, the Coh-Metrix L2 Readability score could be around 15, the vocabulary coverage of BNC 3K word families and the NMET syllabus glossary could reach 96%, lexical density could be around 1 and the incidence of all connectives per 1,000 words could be about 85.

Secondly, the study offers possible methods for the development of reading texts or banking items for formative assessment during teaching. Every natural language processing tool has its advantages. As mentioned earlier, the use of Python reduces statistical differences due to its powerful functions while Coh-Metrix shines at syntactic and discourse levels. Changes in text complexity to suit teacher and students could integrate the relative advantages of Python and Coh-Metrix, and the teacher could conduct an in-depth analysis of selected reading texts by constructing 35 measures covering five aspects: text length, readability, lexical complexity, syntactic complexity and cohesion. Teachers could build an item banking system based on these measures, making the text complexity of reading materials for teaching and testing more sound by being scientifically based.

6.2 The HOW

The washback effect of the NMET should be taken into consideration in the design of teaching. Examinations, especially high-stakes tests like the NMET, have a great impact on the teaching in Chinese high schools. Through the analysis of the text complexity of the NMET reading section, we have primarily become aware of the fact that textual features could make a difference to high school teaching goals, teaching content and teaching methods. Thus, during the teaching process, more attention should be paid to the textual features of reading texts, for instance, the number of words, sentences and paragraphs.

Secondly, the teaching of long sentences and paragraphs should be preceded by a logical analysis of their meanings. As regards vocabulary, teachers ought to put more stress on lexical diversity and strengthen students' ability to comprehend word meanings by analyzing what it means to use a variety of words in different contexts, something that has been neglected in high school teaching.

Thirdly, since this study has been demonstrated that the ability to analyze syntax and cohesion is also required in the NMET, teachers cannot ignore the need to provide an explanation for and the learning of words modifying the main verb, syntactic patterns, overlaps among words and so on.

To sum up, while the analysis of text complexity provides an objective assessment of the difficulty of reading texts, the overall difficulty of the NMET test is also affected by

factors such as test takers' proficiency, familiarity with the topic, cognition and thinking ability. In any case, an objective statistical analysis of text complexity is an important step to assist teachers in their selection of teaching materials and teaching decision. Future research could explore how to use technology to analyze the reading source text and develop more specific teaching strategies for devising or selecting reading texts at different levels of complexity.

Acknowledgment. This work was supported by grants from the National Social Science Fund of China (15BYY080) to Huang, by the Chinese MOE Research Grant for Humanities and Social Sciences (20YJC880032) to the corresponding author Hua, and the Innovation Project of Graduate School of South China Normal University to Wang.

References

Alderson, J.C.: Assessing Reading. Cambridge University Press, London (2000)

Alexander, P.: Reading into the future: competence for the 21st century. Educ. Psychol. **47**(4), 259–280 (2012)

Bachman, L.F., Palmer, A.S.: Language Testing in Practice. Oxford University Press, Oxford (1996)

Berman, R.A.: Syntactic components of the foreign language reading process. In: Alderson, J.C., Urquhart, A.H. (eds.) Reading in a Foreign Language, pp. 139–156. Longman, London (1984)

Cain, K., Nash, H.M.: The influence of connectives on young readers' processing and comprehension of text. J. Educ. Psychol. **103**(2), 429–441 (2011)

Fox, E., Alexander, P.A.: Learning to read. In: Alexander, P.A., Mayer, R. (eds.) Handbook of Research on Learning and Instruction, pp. 7–31. Routledge, New York (2011)

Gardner, D.: Vocabulary input through extensive reading: a comparison of words found in children's narrative and informational reading materials. Appl. Linguist. **25**(1), 1–37 (2004)

Graesser, A.C., McNamara, D.S., Louwerse, M.M.: What do readers need to learn in order to process coherence relations in narrative and expository text? In: Sweet, A.P., Snow, C.E. (eds.) Rethinking Reading Comprehension, pp. 82–98. Guilford, New York (2003)

Graesser, A.C., Swamer, S.S., Baggett, W.B., Sell, M.A.: New models of deep comprehension. In: Britton, B.K., Graesser, A.C. (eds.) Models of Understanding Text, pp. 1–32. Erlbaum, Mahwah (1996)

Guthrie, J.T., Klauda, S.L., Ho, A.N.: Modeling the relationships among reading instruction, motivation, engagement, and achievement for adolescents. Read. Res. Q. **48**, 9–26 (2013)

Halliday, M.A.K., Hasan, R.: Cohesion in English. Longman, London (1976)

Hiebert, E.H.: Text matters in learning to read. Read. Teach. **52**(6), 552–566 (1999)

Hornof, M.: Reading tests as a genre study. Read. Teach. **62**(1), 69–73 (2008)

Jin, T., Guo, K., Mak, B., Wu, Q.: Lexical profiles of reading texts in high-stakes tests: where are the benchmarks? Int. J. Comput. Assist. Lang. Learn. Teach. **7**(1), 34–49 (2017)

Koda, K.: Development of second language reading skills: cross-linguistic perspective. In: Gass, S., Mackey, A. (eds.) The Routledge Handbook of Second Language Acquisition, pp. 303–318. Routledge, London and New York (2012)

Landauer, T., McNamara, D.S., Dennis, S., Kintsch, W.: Handbook of Latent Semantic Analysis. Erlbaum, Mahwah (2007)

Laufer, B., Nation, I.S.P.: Vocabulary. In: Gass, S., Mackey, A. (eds.) The Routledge Handbook of Second Language Acquisition, pp. 163–175. Routledge, London and New York (2012)

Lu, X.: Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment. Lang. Test. **34**(4), 493–511 (2017)
McCarthy, P.M., Guess, R.H., McNamara, D.S.: The components of paraphrase evaluation. Behav. Res. Methods **41**(3), 682–690 (2009)
McNamara, D.S., Graesser, A.C., McCarthy, P.M., Cai, Z.: Automated Evaluation of Text and Discourse with Coh-Metrix. Cambridge University Press, New York (2014)
Mesmer, H.A., Cunningham, J.W., Hiebert, E.H.: Toward a theoretical model of text complexity for the early grades: learning from the past, anticipating the future. Read. Res. Q. **47**(3), 235–258 (2012)
Ministry of Education of the People's Republic of China: National English Curriculum Standards for General High School, 2017 edn. People's Education Press, Beijing (2018)
Nation, I.S.P.: How large a vocabulary is needed for reading and listening? Can. Mod. Lang. Rev. **63**(1), 59–82 (2006)
Nation, I.S.P., Beglar, D.: A vocabulary size test. Lang. Teach. **31**(7), 9–13 (2007)
National Governors Association Center for Best Practices & Council of Chief State School Officers (NGA/CCSSO): The Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects, Appendix A: Research Supporting Key Elements of the Standards, Washington, DC: Author (2010)
O'Reilly, T., McNamara, D.S.: Reversing the reverse cohesion effect: good texts can be better for strategic, high-knowledge readers. Discourse Process. **43**, 121–151 (2007)
Paris, S.G., Cross, D.R., Lipson, M.Y.: Informed strategies for learning: a program to improve children's reading awareness and comprehension. J. Educ. Psychol. **76**(6), 1239–1252 (1984)
Pearson, P.D.: The Roots of Reading Comprehension Instruction. In: Israel, S., Duffy, G. (eds.) Handbook of Research on Reading Comprehension, pp. 3–30. Taylor & Francis, New York (2009)
Peng, Y., Gu, X., Huang, J.: The research on the test-designed quality of the reading comprehension among test papers of NMET national version and provincial autonomous ones in 2016. Educ. Meas. Eval. **3**, 57–64 (2017)
Pulido, D.: The relationship between text comprehension and second language incidental vocabulary acquisition: a matter of topic familiarity? Lang. Learn. **54**(3), 469–523 (2004)
Shrum, J., Glisan, E.: Teacher's Handbook: Contextualized Language Instruction, 4th edn. Heinle Cengage Learning, Canada (2010)
Verhoeven, L.: Second Language Reading Acquisition. In: Kamil, M., Pearson, P., Moje, E., Afflerbach, P. (eds.) Handbook of Reading Research, vol. IV, pp. 661–683. Routeldge, New York (2011)
Wang, R.: Study on the vocabulary of reading comprehension texts in Jiangsu National Matriculation English Test from 2008 to 2017. Educ. Meas. Eval. **1**, 19–25 (2018)
Wang, Q., Qian, X., Zhou, M.: Text analysis in English language teaching—importance and methods. Foreign Lang. Educ. China **2**(2), 40–47 (2019)
Wright, T.S., Cervetti, G.N.: A systematic review of the research on vocabulary instruction that impacts text comprehension. Read. Res. Q. **52**(2), 203–226 (2016)