Using TAASSC to Investigate Fine-Grained Grammatical Complexity in Reading Texts of Two High-Stakes English Tests in China

Shengshu LIN

School of Foreign Languages, Nanjing University of Posts and Telecommunications, China
shengshu2016@hotmail.com

Abstract: Fine-grained grammatical complexity measures are better predictors than large-grained indices of register variation and writing of different levels in that they provide in-depth explanation of what accounts for syntactic complexity. Computer technologies have made it possible to study these fine-grained measures based on large corpora. Using the Tool for Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC) (Kyle, 2016), the present study investigates phrasal and clausal complexity in reading texts of two high-stakes English tests in China (CET-4 and NETEM) different in difficulty level. The results showed that clausal and phrasal complexity indices could predict the two tests. More importantly, in combined analysis of the two types of measures, the results demonstrated that noun phrases, particularly those with phrasal modifiers, were better predictors of the more difficult test NETEM. These findings support the importance of noun phrases and the diverse modifiers they take in informational texts. Findings of this research will help college students advance their English learning with a specific purpose and in a right direction.

Keywords: clausal complexity, phrasal complexity, reading texts, tests

1. Introduction

Reading comprehension has always been an important task in time-limited English tests. Yet grammatical complexity is given scant attention in reading accuracy despite its wide attention in writing. Grammatical complexity in spoken and written English and its historical changes were fully documented in Biber and Gray (2016). Recent research showed that fine-grained grammatical complexity measures were more effective in accounting for sentence length and writing quality (Kyle & Crossley, 2018; Zhang & Lu, 2022; Zhang & Liu, 2022; Gray, Geluso, & Nguyen, 2019). According to Biber, Gray and Poonpon (2011), advanced phrasal features are not acquired naturally and not always successfully acquired. The empirical research by Priven (2020) showed that fine-grained phrasal complexity pervasive in university textbooks caused difficulty to ESL readers. In Bychkovska (2021), the improved quality of writing by ESL writers after explicit teaching of phrasal features also lent support to the importance of fine-grained complexity features in reading. Due to the washback effects of tests, reading texts of a test are a critical source of language input for learners. For example, Chinese learners in preparation for NETEM (the National Entrance Test of English for MA/MS) use past years' real tests as major learning materials. The present study is conducted with the purpose of shedding light on the fine-grained linguistic features in reading texts of two important tests in China and helping test takers advance their learning in a right direction.

2. Research Background

2.1 Fine-Grained Measures

Measures of grammatical complexity have undergone dramatic changes with mounting attention to fine-grained measures because they can account for what contributes to the length of length-/ratio-based large-grained measures.

In this line of research, Biber and colleagues have always employed form-function framework to define linguistic features. They believe only by taking structural types and their grammatical functions into consideration can register variations be fully described, because the same structural type may have different functions in different registers. By now, the eight form-function categories have included over 130 specific linguistic features and proved fruitful. Register studies showed that academic discourse was complex in the dense use of phrasal features as nominal modifiers, while spoken conversation was clausally complex. Language learners progress along a sequence of linguistic features acquiring grammar in conversation first and that of writing later (Biber et al., 2011). Their theory was supported by many researchers (Staples et al., 2016; Parkinson & Musgrave, 2014; Gray et al., 2019).

Although fine-grained measures based on form-function framework are easy to understand, the tools used to tag and count linguistic features are not publicly available. From the usage-based theory of language acquisition, Kyle (2016) developed the Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC) by utilizing Stanford Neural Network Dependency Parser (version 3.5.1), making it possible to analyze a large number of fine-grained linguistic units. Research conducted with TAASSC showed that fine-grained measures were better predictors than large-grained measures of learner writing (Kyle & Crossley, 2018; Zhang & Liu, 2022; Zhang & Lu, 2022). Kyle and Crossley (2018) was the first research that directly compared the predictive power of large-grained indices (calculated by L2SCA), clausal and phrasal complexity in TOEFL independent writing by L2 writers from different language backgrounds. Large-grained indices explained 5.8% of variance in holistic essay scores while fine-grained clausal complexity explained 3.0% and fine-grained phrasal complexity explained 18.9%. The large predictive power was supported by recent studies (Zhang & Liu, 2022; Zhang & Lu, 2022). Zhang and Liu (2022) examined Chinese university students' English expository writing. Their results showed that fine-grained indices accounted for 22.2% of the variance in students' writing quality, significantly higher than 7.2% by large-grained indices. Zhang and Lu (2022) investigated grammatical complexity in application letters and argumentative essays by Chinese students. Traditional indices explained 20.2% of the variance in letter scores, and 15.7% of the variance in essay scores while fined-grained measures explained 31.9% and 30.6% respectively.

The research aforementioned focused on the predictive power of fine-grained measures in writing while investigation of reading texts in tests has been ignored by now. Some fined-grained measures such as (multi-)noun as premodifiers and postmodifying prepositional phrases are compressed with dense information, and may cause confusion in reading (Priven, 2020). Therefore, utilizing TAASSC, the present study makes contributions to this area of research by investigating the reading texts of two important English tests in China.

2.2 CET-4 and NETEM in China

College English Test Band 4 (CET-4) as a proficiency test for university students in China is influential among college students in China. In the third year or even earlier, millions of college students will start to prepare for NETEM held in December annually to pursue postgraduate studies. Reading tasks common in CET-4 and NETEM require test takers to answer five multiple-choice questions after reading a passage. Although reading texts in these two tests are extracts from some publications in English-speaking countries including *The Economist, The Guardian* or other academic reading materials, CET-4 is easier for its simpler vocabulary, shorter sentence length and text length. Since NETEM is regarded as the most challenging test for non-English majors, some publications encourage test takers to use complex clauses to obtain higher scores without mentioning phrasal features (Zhou, 2022). The status quo reflects the stereotype of academic writing as clausally complex and clausal elaboration as a focus in most EAP teaching (Parkinson, 2020). To reveal grammatical complexity in the reading texts of the two tests, the present study addresses three research questions:

RQ1. What clausal complexity measures can predict CET-4 texts and NETEM texts?
RQ2. What phrasal complexity measures can predict CET-4 texts and NETEM texts?
RQ3. What clausal and phrasal complexity measures when considered together can predict CET-4 texts and NETEM texts?

3. Methodology

3.1 Corpus

Two corpora (NETEM and CET-4) used in this study were built by the author. NETEM is held annually with four reading passages in each real test paper of the previous 13 years from 2011 to 2023. 52 texts were collected with a total of 23,023 words. CET-4 reading texts were more easily available because it is administered twice a year with 6 different reading passages each time. In order to control the effect of time on CET-4 language use, we randomly selected four texts in each year to ensure that CET-4 reading passages are evenly distributed from 2010 to 2022. 52 CET-4 texts totaled 18,411 words. All texts were saved in a .txt format with the filename "year_test_text_number" that can be processed by TAASSC.

3.2 Tools and Procedures

The current study uses TAASSC (Kyle, 2016; Kyle & Crossley, 2018) to investigate clausal and phrasal complexity. From usage-based theory perspective, the clausal complexity in TAASSC refers to a main verb and its associated structures. TAASSC creates a dependency representation of each sentence, that is, the functional categories and their grammatical relations. It includes 31 clausal indices, of which 2 are general indices of clausal complexity and 29 calculate the average number of particular units per clause. The phrasal complexity involves 7 noun phrase types and 10 dependent types. Phrasal measures produced by TAASSC have pronoun version and no-pronoun version, each including 66 indices. This study used no-pronoun version because pronoun-headed phrases are not common (Lan et al., 2019). Of the 66 phrasal indices, 16 calculate the average number of dependents per each phrase type and for all phrase types, and their standard deviations (SDs). 10 indices calculate the occurrence of particular dependent types regardless of the type of noun phrases they occur in. 40 indices calculates the average occurrence of particular dependent types in particular types of noun phrases. A higher value indicates a higher frequency of a linguistic unit. SDs are used to measure syntactic variation.

These 97 complexity indices collected by TAASSC were entered in SPSS as independent variables. Since rareness of some linguistic structures is inherent in fine-grained analysis (Kyle & Crossly, 2018), we removed those that displayed "0" value. Due to nonnormal distributions of the variables, non-parametric independent samples Mann-Whitney U test was performed to identify the indices on which CET-4 and NETEM differed significantly ($p < 0.5$). This step identified the independent variables that were used for binary logistic regression. Binary logistic regression was performed on the significantly different clausal complexity indices and phrasal complexity indices respectively to determine which were predictors of reading texts of the two tests. The results from these two steps answered RQ1 and RQ2. To answer RQ3, a combined analysis of the significantly different clausal and phrasal indices was conducted to determine which indices entered the final regression model.

4. Results and Discussion

Table1, Table 2 and Table 3 include the variables that entered into binary logistic regression. The variables with asterisks were predictors of texts in the two tests at the statistical level of

0.05. Exp(B) values more than 1 indicated the complexity indices had predictive effects on texts in NETEM, while Exp(B) values less than one predicted CET-4 texts.

Table 1. *Logistic Regression of Clausal Complexity*

		B	S.E.	Wald	Sig.	Exp(B)
1	Adjective complements per clause	-3.652	5.262	0.482	0.488	0.026
2	Conjunctions per clause	6.568	4.955	1.757	0.185	712.143
3	Direct objects per clause	4.534	2.592	3.060	0.080	93.152
4	Bare noun phrase temporal modifiers per clause	48.906	19.923	6.026	0.014*	1.736
5	Auxiliaries per clause	7.367	3.322	4.919	0.027*	1583.482
	Constant	-4.215	1.549	7.401	0.007	0.015

Table 2. *Logistic Regression of Phrasal Complexity*

		B	S.E.	Wald	Sig.	Exp(B)
1	Dependents per nominal	-1.645	2.685	0.375	0.540	0.193
2	Dependents per object of the preposition	3.254	2.047	2.527	0.112	25.882
3	Dependents per nominal subjects (SD)	3.965	1.572	6.363	0.012*	52.720
4	Dependents per passive nominal subject (SD)	0.678	0.547	1.538	0.215	1.970
5	Dependents per agent(SD)	1.191	0.649	3.369	0.066	3.290
6	Verbal modifiers per nominal	-31.467	13.879	5.141	0.023*	0.000
7	Nouns as nominal dependents per nominal	4.992	4.275	1.364	0.243	147.287
8	Possessives per direct object	8.531	3.531	5.839	0.016*	5071.844
	Constant	-7.178	2.487	8.329	0.004	0.01

Table 3. *Logistic Regression of Clausal Complexity and Phrasal Complexity Combined*

	Independent variables	B	S.E.	Wald	Sig.	Exp(B)
1	Dependents per nominal	-3.475	3.233	1.155	0.282	0.031
2	Dependents per object of the preposition	4.801	2.432	3.896	0.048*	121.638
3	Dependents per nominal subject (SD)	4.775	1.795	7.074	0.008*	118.500
4	Dependents per passive nominal subject (SD)	0.544	0.672	0.655	0.418	1.724
5	Dependents per agent (SD)	1.724	0.811	4.522	0.033*	5.607
6	Verbal modifiers per nominal	-37.935	16.829	5.081	0.024*	0.000
7	Nouns as nominal dependents per nominal	1.801	4.518	0.159	0.690	6.054
8	Possessives per direct object	10.437	4.230	6.089	0.014*	34083.305
9	Adjective complements per clause	-2.666	6.630	0.162	0.688	0.070
10	Conjunctions per clause	3.663	6.374	0.330	0.565	38.992
11	Direct objects per clause	5.389	3.179	2.873	0.090	218.883
12	Bare noun phrase temporal modifiers per clause	78.275	28.267	7.668	0.006*	9.873
13	Auxiliary per clause	3.541	4.187	0.715	0.398	34.494
	Constant	-10.874	3.435			

To answer RQ1, Table 1 shows that bare noun phrase temporal modifiers and auxiliaries are more common in NETEM texts, and the more use of these features indicates the higher probability that reading texts are from NETME. Many of NETME texts are extracted from top newspapers, where temporal words are important linguistic resources in narrating events. The ability to capture the temporal dimension is essential for understanding news events. Auxiliaries can convey much grammatical information such as aspect, tense, voice and modality. More importantly, modal auxiliaries express a variety of meanings related to concepts such as ability, permission, necessity, and obligation (Biber et al., 1999). It is very common in NETEM reading comprehension that questions are asked about attitudes of the news writer or people involved in news events. NETME is more complicated than CET-4 in the use of temporal words to connect events and auxiliaries to express grammatical information and attitudes, though both tests use news extracts as a major source of reading texts.

With regard to RQ2, Table 2 shows that NETEM employs diverse modifying structures in noun phrases as a subject, as is illustrated by Excerpt 1. Excerpt 1 includes adjective "personal", (multi-)noun "grievance" and "Employment Relations", of-preposition "of New Zealand's...2000", and possessive "New Zealand's".

Excerpt 1: The personal grievance provisions of New Zealand's Employment Relations Act 2000 (ERA) prevent an employer from firing an employee without good cause. (2022_NETEM_text_4)

A subject with various types of modifiers are quite common in NETEM reading texts because the extended subject packages dense information into phrasal premodifiers. Another index that makes reading texts more likely to be NETEM texts is the use of more possessives modifying a noun in direct object positions. Possessives, nouns and adjectives are the most compressed modifiers in noun phrases (Biber & Gray, 2016). However, using possessives, especially genitive 's is not encouraged among Chinese students because it may strengthen the stereotype of "Chinese English" since its counterpart structure in Chinese is prevalent. In this sense, our findings counter this stereotype and highlight the importance of these features in reading. According to Biber and Gray (2016), the use of verbal modifiers is viewed as more structurally elaborate rather than compressed, and is more common in spoken conversation. This is supported by our finding that more use of verbal modifiers of a noun will increase the likelihood of reading being CET-4 texts. Instead, phrasal modifiers are a common feature in passages with information focus (Biber & Gray, 2016).

Results in Table 3 provide answers to RQ3. As can be seen in Table 3, when clausal and phrasal complexity were combined, two new phrasal indices entered the regression model: dependents per object of prepositional phrases and dependents per agent (SD), while only bare noun phrase temporal modifiers per clause was retained as clausal index. Modifying devices of agents are also more diverse in NETEM reading. The index of dependents of prepositional objects further affirmed that NETEM were more likely to use extended noun phrases with embedded modification as illustrated by Excerpt 2. The prepositional phrase "of human activity" also takes an embedded modifier "in recent climate change", which again takes an adjective "recent" and a noun "change" as its modifiers.

Excerpt 2: These would emphasise the primacy of human activity in recent climate change and encourage discussion of mitigation measures. (2023_NETEM_text_1)

Based on Biber et al.'s (2011) language development theory, extended prepositional phrases as post-modification are one of the most advanced features of academic language, and also the most difficult for both native and EFL language learners. NETEM is more likely to use diverse dependents to add information in compressed structures. According to Kyle and Crossley (2018), Zhang and Lu (2022), phrasal complexity is more predictive of higher level learner writing. This study further affirmed that, in English news, fine-grained indices, phrasal level features in particular, correlated significantly with the degree of difficulty and revealed what specific patterns contributed to difficulty of textual sentences in NETME. It is noteworthy that all the indices included in the regression model of combined analysis were realized linguistically by noun phrases that may take a variety of modifiers. This lends support to Biber and Gray's (2016) claim about noun phrases that they are prevalent in information writing, especially those with phrasal modifiers. In view of empirical evidence that extended noun phrases may cause difficulty in reading university textbooks to L2 learners (Priven, 2020), noticing and understanding complex nominal modifiers should be emphasized in teaching advanced English learners. For those who prepare for NETEM or plan their learning after passing CET-4, using these compressed phrasal features to convey dense amounts of information should be a central area of learning.

5. Conclusion

Utilizing TAASSC, the present study provides insights into what fine-grained measures are predictive of reading texts of two important tests for university students in China. Findings will help millions of students focus on the linguistic features that are more typical of NETEM rather

than just spending time on clausal subordination. Future research can include vocabulary as a predictor variable, comprehensively looking at what vocabulary characteristics and linguistic features can predict reading texts in tests for different purposes and at different stages of English learning.

Acknowledgements

References

Biber, D., & Gray, B. (2016). *Grammatical Complexity in Academic English: Linguistic Change in Writing*. Cambridge University Press.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, *45*(1), 5-35.

Biber, D., Johansson, S., Leech, G., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Pearson Education Limited.

Bychkovska, T. (2021). Effects of explicit instruction on noun phrase production in L2 undergraduate writing. *Journal of English for Academic Purpose*, *54*,1-16.

Gray, B., Geluso, J., & Nguyen, P. (2019). The longitudinal development of grammatical complexity at the phrasal and clausal levels in spoken and written responses to the TOEFL iBT® test. *ETS Research Report Series*, (1), 1-51.

Kyle, K. (2016). *Measuring Syntactic Development in L2 Writing: Fine-grained Indices of Syntactic Complexity and Usage-based Indices of Syntactic Sophistication*. Unpublished doctoral dissertation. Georgia State University, America.

Kyle, K., & Crossley, S. A. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *The Modern Language Journal*, *102*(2), 333-349.

Lan, G., Lucas, K., & Sun., Y. (2019). Does L2 writing proficiency influence noun phrase complexity? A case analysis of argumentative essays written by Chinese students in a first-year composition course. *System*, *85*, 1-13.

Parkinson, J. (2020). Book review. *Journal of English for Academic Purposes*, *45*,1-2.

Parkinson, J., & Musgrave, J. (2014). Development of noun phrase complexity in the writing of English for Academic Purposes students. *Journal of English for Academic Purposes*, *14*, 48-59.

Priven, D. (2020). "All These Nouns Together Just Don't Make Sense!": An Investigation of EAP students' challenges with complex Noun Phrases in First-Year college-level textbooks. *Canadian Journal of Applied Linguistics*, *23*(1), 93-116.

Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication*, *33*(2), 140-183.

Zhang, X., & Liu, W. (2022). Predictive effects of syntactic complexity on the quality of Chinese university students' English expository writing. *Modern Foreign Languages*, *45*(3), 331-343.

Zhang, X., & Lu, X. (2022). Revisiting the predictive power of traditional vs. fine-grained syntactic complexity indices for L2 writing quality: Then case of two genres. *Assessing Writing*, 51, 1-14.

Zhou, S. (2022). *An Interpretation of the Syllabus for National Entrance Test of English (1) for MA/MS*. Beijing: Higher Education Press.