(O)

2021 2023

Spoken and Written Vocabulary in High-stakes Assessments: The Impact of Lexical Choices at Different IELTS Proficiency Levels.

Higginbotham, George

3,200,000

2 200 L2 6

IELTS

IELTS 2 5

IELTS IELTS 5.
5 7.5

This research project, in collaboration with Queen Mary University of London, created a database of over 200 L2 learners written and spoken vocabulary and analysed the data in two main ways. One part of the analysis focused on students written essays, using them to validate 6 measures of lexical diversity. As well as finding that all the measures (to varying degrees) related to IELTS proficiency scores, another finding was that it is important from a methodological perspective to be clear on how the words are counted: either as lemmas or flemmas. The other part of this project focused on students spoken data (seminar discussions) to see how students use of multi-word-units affected their fluency. The findings demonstrate that students who have higher IELTS spoken grades are also the students that are making effective use of multi-word-units. As well as looking at collocations (2 word units) this part of the project extended previous studies by exploring 3, 4 and 5 word units.

productive vocabulary

vocabulary lexical diversity multi word units

Background to the research

This research project was designed to understand the vocabulary requirements for written and spoken high-stakes English language assessments (IELTS) in order to better help second language learners prepare for these types of assessments. As such tests are gateways to overseas universities, it is vital to understand their precise linguistic requirements, at each level, if Japan is to achieve the goal of developing human resources capable of working within a global context. Previous studies of vocabulary knowledge have tended to focus on one mode of production, making it difficult to ascertain how vocabulary acquisition in one mode of production impacts the learners' ability to use that item across other modes of production (Clenton & Booth, 2020). Papers such as Treffers-Daller et al. (2018) help to discriminate between writers from second language learners of English, whose work has been assessed as belonging to different levels of the Common European Framework of Reference (CEFR), help in some regard by detailing vocabulary knowledge at generic proficiency levels. Similarly, when examining speaking, papers (Clenton et al. 2020, De Jong et al. 2012, Tavakoli & Uchihara, 2020) detail the spoken vocabulary knowledge of different second language users of different proficiency levels. The current project therefore attempts to bridge these two strands of research by detailing the written as well as the spoken vocabulary knowledge at different levels. As the Principal Investigator (PI) has a strong connection with a language program at a UK university (Queen Mary University of London) that regularly collects essays and records seminars for a large body of L2 students from diverse L1 backgrounds, it was therefore proposed that we collect a multimodal spoken as well as written corpus of student generated texts of academic English, which could be used to analyze the written and spoken vocabulary at various proficiency levels.

Purpose of the study

The initial aim of the project was to develop a multimodal spoken as well as written corpus of student generated texts of academic English that could be used to analyze the vocabulary that second language learners produced. Specifically, we were interested in understanding the differences in the lexical complexity of learner produced texts over different levels of proficiency and across different modes of production. With two modes of production (written and spoken), this study can be seen to have two main threads. With the written thread we could use the corpus to analyze the extent to which automated (Lexical Diversity) writing measures were able to distinguish between different generic proficiency levels. With the spoken thread we could use the corpus to explore how fluency was related to proficiency and other factors, such as how efficiently L2 learners used multi word units (MWUs).

Methods of research

In the first phase of the project, in collaboration with teachers and program managers at Queen Mary University of London (QMUL), students on a pre-sessional language course were invited to join the research project. Having signed ethical consent forms, and in compliance with the QMUL ethics guidelines, essays and recordings of seminar discussions were collected to form the corpus.

In the second phase of the project (to explore vocabulary produced in writing), a sample of essays were selected from the L2 learner corpus. Selected participants (n = 105) were 55 (52.38%) male and 50 (47.62%) female students. To limit the potential variability in the data, the participants selected were all from the same L1 background (Chinese speakers) and all Humanities and Social Science majors. Their written language proficiency ranged from IELTS bands 6.5 (n = 29), 7 (n = 43), and 7.5 (n = 33). First the essays were cleaned. By this we mean: in-text direct citations were excluded; proper names, acronyms, and cardinal numbers were removed; spelling errors were corrected; contracted forms were transformed into full forms they were analyzed against 6 measures of lexical diversity. This was achieved with a python program for natural language processing. For the cleaned essays, LD scores were computed for three basic LD measures (Types, TTR, Guiraud's Index), and three sophisticated LD measures (D, MTLD, HD-D). As we were concerned that *how words are counted* might be an issue, this was repeated in 3 ways, for non-lemmatized texts, lemmatized texts and flemmatized texts.

In the third phase of the project (to explore spoken vocabulary) a sample of seminar discussion recordings were selected from the L2 learner corpus. Our participants (n = 150) were again (mostly) Chinese L1 and again an equal mix of genders. The learners this time were, however, from a wide range of disciplines: humanities, law, science, technology, engineering, and mathematics majors. Their spoken language proficiency ranged from IELTS bands 6.5 (n = 50), 7 (n = 50), and 7.5 (n = 50). From the 15-minute seminar discussions, 3-minute segments (starting at the 30 second mark) were transcribed partly using speech to text software and partly through traditional (non-automated) transcription. When the recordings had been transcribed (and cleaned in a similar way to the essay data), they were analyzed against

3 measures of fluency (speed, breakdown, and repair fluency), to see the effect short (2 or 3 word) lexical bundle use and longer (4 or 5 word) lexical bundle use has on fluency.

Research Achievements

The first achievement is the creation of a large corpus of both written and spoken academic L2 English that can be analyzed from a range of proficiency levels (IELTS levels 5 – 8), which has, so far, been analyzed in two main ways.

Our exploration of the written corpus found that all of the 6 LD measures investigated were able to predict the different proficiency levels, although their predictive strength was determined by the method used to count words: simple count, lemma or flemma. Types and D methods for measuring lexical diversity were, for example, more powerful measures when flemmatized data was used. TTR and Guiraud's Index on the other hand were stronger discriminators of writing levels on lemmatized data. In addition, when a simple count method was used the sophisticated LD measures (such as D) were more predictive. Importantly, our investigations underline the need for researchers to be aware that how they count words will affect the analysis. For a more detailed discussion of our findings, refer to Maw et al. (2022), an international journal article co-authored by the PI and one of the co-investigators in this project.

Our exploration of the spoken corpus found an increase in lexical bundle (LB) usage, particularly in bigram and trigram frequency, as proficiency levels increased. This trend highlights the connection between proficiency development and the use of shorter LBs. Significant trends also emerged in longer LBs, with the mean values for three-to-five-word usage and MI scores increasing notably. Confirmed by ANOVA, these results support our hypothesis of a positive link between proficiency and increased use of both shorter and longer LBs. These trends suggest that higher proficiency learners are not only employing longer and more complex LBs but are doing so with greater sophistication, which is crucial for advanced language proficiency. For a more detailed discussion of our findings and the pedagogic implications, refer to Hougham et al. (2024a; 2024b), both of which are international journal articles co-authored by the PI and one of the co-investigators in this project.

As well as the specific findings from our experimental work, outlined above, this project has been able to partially support 2 PhD theses, we therefore anticipate further investigations into productive vocabulary in the near future. Consequently, the database that we created for this project is still being used in our ongoing exploration of the vocabulary that L2 students produce at various proficiencies.

5. References
Clenton, J., & Booth, P. (Eds.). (2020). *Vocabulary and the four skills: Pedagogy, practice, and implications for teaching vocabulary.* Routledge.
Clenton, J., de Jong, N. H., Clingwall, D., & Fraser, S. (2020). Investigating the extent to which vocabulary knowledge and skills can predict aspects of fluency for a small group of pre-intermediate Japanese L1 users of English (L2). *In Vocabulary and the four skills* (pp. 126-145). Routledge.
De Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2012). Facets of speaking proficiency. Studies in Second Language Acquisition, 34, 5-34.
Hougham, D., Clenton, J., & Uchihara, T. (2024). Disentangling the contributions of shorter vs. longer lexical bundles to L2 oral fluency. *System*, 121, 103243. https://doi.org/10.1016/j.system.2024.103243
Hougham, D., Uchihara.T, Clenton J. & Higginbotham, G. (2024) The Impact of Lexical Bundle Length on L2 Oral Proficiency. *Languages.* (accepted for publication, in press)
Maw, T. M. M., Clenton, J., & Higginbotham, G. (2022). Investigating whether a flemma count is a more distinctive measurement of lexical diversity. *Assessing Writing,* 53, 100640. https://doi.org/10.1016/j.asw.2022.100640
Tavakoli, P., & Uchihara, T. (2020). To what extent are multiword sequences associated with oral fluency?. *Language Learning*, 70(2), 506-547.
Treffers-Daller, J., Parslow, P., & Williams, S. (2018). Back to basics: How measures of lexical diversity can help discriminate between CEFR levels. *Applied Linguistics*, 39(3), 302–327. https://doi.org/10.1093/applin/amw009

	4	4	3	0
Myint Maw, T. M, Clenton, J., & Higginbotham, G.				53
Investigating whether a Flemma Count is a More Distinctive Measurement of Lexical Diversity				2022
Assessing Writing				100640
DOI 10.1016/j.asw.2022.100640				

Myint Maw, T. M	March
Investigating Factors Affecting Lexical Diversity Measure Predictions of Writing and Speaking Proficiency: Word-Counting Criteria, L1 Background, Language Proficiency, and Text Length	2023
Hiroshima University, PhD thesis	1-278
DOI	

Hougham Dan Clenton Jon Uchihara Takumi	121
Disentangling the contributions of shorter vs. longer lexical bundles to L2 oral fluency	2024
System	103243 103243
DOI 10.1016/j.system.2024.103243	

Hougham D, Clenton J, Uchihara T, Higginbotham G.	"-"
The Impact of Lexical Bundle Length on L2 Oral Proficiency.	2024
Languages	"-"
DOI	

5 5 3

Myint Maw, T. M, Clenton, J., & Higginbotham, G.
Can Lexical Diversity measures discriminate between IELTS bands?
Hiroshima Lexical Research Forum
2022

Higginbotham, G, Myint Maw, T. M, Hougham, D
Investigating Productive Vocabulary
QMUL Language Centre Learning and Teaching Conference
2022 2023

Myint Maw, T. M, Clenton, J., & Higginbotham, G.
Can Lexical Diversity measures discriminate between IELTS bands?
Hiroshima Lexical Research Forum
2021

Hougham, D, Clenton, J., & Higginbotham, G.
How Does Learner Use of Longer MWUs Relate to Fluency?
Vocabulary@Vic Conference, Victoria University of Wellington
2023 2024

Researching Online: issues with data collection.		
Higginbotham, G.		
QMUL Language Centre Learning and Teaching Conference		
2022		

0

(Clenton Jonathon) (80762434)	(15401)	
(Brooks Gavin) (10610818)	(34304)	
(Fraser Simon) (10403510)	(15401)	

2

Hiroshima Lexical Research Forum (HLRF) Online Conference	2021	2024
QMUL Language Centre Conference: Theme ' Facing up to face-to-face teaching'	2022	2022

	Queen Mary University of London			
	Hiroshima University			
	Eikei University of Hiroshima			