



Investigating the Impact of Dialogic and Trialogic Interactive Factors on Chinese Advanced L2 learners' Vocabulary Use in Spoken Contexts

Wang-Taylor, Yixin; Clenton, Jon; Ren, Yinna

Languages

DOI:
[10.3390/languages9080266](https://doi.org/10.3390/languages9080266)

Published: 30/07/2024

[Cyswllt i'r cyhoeddiad / Link to publication](#)

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):
Wang-Taylor, Y., Clenton, J., & Ren, Y. (2024). Investigating the Impact of Dialogic and Trialogic Interactive Factors on Chinese Advanced L2 learners' Vocabulary Use in Spoken Contexts. *Languages*, 9(8). <https://doi.org/10.3390/languages9080266>

Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Article

Investigating the Impact of Dialogic and Trialogic Interactive Factors on Chinese Advanced L2 learners' Vocabulary Use in Spoken Contexts

Yixin Wang-Taylor ^{1,*} , Jon Clenton ²  and Yinna Ren ³¹ School of Arts, Culture and Language Bangor University, Bangor LL57 2DG, UK² Graduate School of Humanities and Social Sciences, Hiroshima University, Hiroshima 739-8527, Japan³ College of Software, Nankai University, Tianjin 300350, China; ryinna@163.com

* Correspondence: y.wang-taylor@bangor.ac.uk

Abstract: The main objective of this study is to investigate how interactive factors affect the vocabulary usage of second language learners in their spoken language. Participants were 24 L1 Chinese undergraduate students of L2 English at an advanced level. L2 learners' vocabulary use was assessed via tokens, lexical diversity, and frequency-based lexical sophistication. Participants provided speech data in response to seven persuasive speaking tasks across three speaking modes: two monologic, two dialogic, and three trialogic. This study showed that the interactive factor has a varied effect on L2 learners' vocabulary usage. It positively influences the use of advanced vocabulary but does not affect the total number of words produced or the diversity of words used. Second, of all three speaking modes, the dialogic speaking mode is the best speaking condition to trigger L2 learners' use of advanced words. Third, the vocabulary employed in dialogues and trialogues can vary due to the inherent disparities between the two modes of speech. Therefore, we propose the use of the dialogic interactive factor and trialogic interactive factor instead of the term "interactive factor" to encompass two specific conditions in which there was a noticeable difference in the performance of L2 learners.

Keywords: vocabulary use; interactive factor; tokens; lexical diversity; lexical sophistication; dialogues; trialogues



Citation: Wang-Taylor, Yixin, Jon Clenton, and Yinna Ren. 2024.

Investigating the Impact of Dialogic and Trialogic Interactive Factors on Chinese Advanced L2 learners' Vocabulary Use in Spoken Contexts. *Languages* 9: 266. <https://doi.org/10.3390/languages9080266>

Academic Editor: Antony Kunnan

Received: 26 April 2024

Revised: 27 June 2024

Accepted: 17 July 2024

Published: 30 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Vocabulary usage is a crucial factor in scoring high-stakes speaking assessments (e.g., IELTS, TOEFL iBT, TOEIC), and language ability descriptors (e.g., CEFR). The adjacent descriptors of "vocabulary" in the marking rubrics of four dominant international English-speaking tests, IELTS (2024), TOEFL iBT (2024), TOEIC (2024), and Cambridge Speaking Test Suite (2024, CEFR vocabulary descriptors), see Table 1, show that human raters' assessment of speaking is guided by a perception of how vocabulary is used by test-takers and that the aspects of vocabulary use cover accuracy, range, frequency, and accessibility.

Based on empirical data, studies have shown that human raters' holistic judgements are significantly associated with lexical diversity (Appel et al. 2019; Lu 2012; Noreillie et al. 2020) or with lexical sophistication (Kyle and Crossley 2015; Li and Lorenzo-Dus 2014; Saito et al. 2016). Several studies have found that the evaluation of speaking skills by human raters can be affected by multiple factors related to vocabulary usage, such as the variety of words used, the total number of words spoken, and the frequency of specific words (Crossley et al. 2015; Iwashita et al. 2008). Crossley et al. (2015), for instance, reported that lexical diversity, word frequency, collocation accuracy, and word concreteness together explained 91% of the variance in the holistic evaluation of lexical proficiency. Examining 200 speaking samples collected from five monologues from five speaking proficiency levels, Iwashita et al. (2008) reported that increases in level were associated with an increase in tokens and types. Comparing L2 learners' vocabulary use with native speakers reported

that native speakers with more lexical proficiency use a wider range of vocabulary and longer speech (e.g., [Noreillie et al. 2020](#); [Skehan 2009a](#); [Tavakoli and Foster 2008](#)).

Table 1. Adjacent words to “vocabulary” in four dominant international spoken proficiency tests.

Speaking Tests	Accuracy	Range	Frequency	Accessibility
A. IELTS	Precise, Precision	Wide, Wide enough, Insufficient	Less common, Simple	Readily, Flexibly, Flexibility
B. TOEFL iBT Independent/ Integrated	Inaccurate	Limited, Severely limited	Simple	Automatic, Effective
C. TOEIC	Accurate, Precise, Imprecise	Limited, Severely limited, Insufficient		
D. Cambridge test suite	Appropriate, High level of accuracy	Wide	Less common, Simple, Elementary	

The above emphasizes the significant role of vocabulary use assessment in speaking. It also emphasizes a major drawback in the dominance of monologic speech studies. Of the above leading studies, only two ([Noreillie et al. 2020](#); [Crossley et al. 2015](#)) elicited speech in interactional contexts. Despite not being widely studied in vocabulary research, interactive speaking tests are a significant part of the speaking test market. According to [Cambridge Language Assessment \(2023\)](#), the design of Cambridge ESOL speaking tests allocates 58–64% of the time to paired interactive tasks for candidates at the B1-B2 level, and 50–85% of the time for candidates at the C1-C2 level. Local speaking proficiency tests, like the Chinese national CET 4 and 6 Speaking Test ([WENR 2023](#)) and the speaking component of the Hong Kong Diploma of Secondary Education ([HKDSE 2023](#)), allocate approximately 50% and 70% of the speaking time to paired discussion, respectively. The current study is therefore crucial as it fills a gap in research by exploring the scarcity of interactive design in speaking tests. Our study aims to examine the impact of the interactive factor on L2 learners’ vocabulary use in speaking.

1.1. Vocabulary Use Measures

The current study analyses how L2 learners use vocabulary in terms of lexical diversity and sophistication. Previous research (e.g., [Crossley et al. 2015](#)) has established a strong correlation between human raters’ evaluation and these two factors. The literature has scrutinized these factors with written but not spoken texts ([Gregori-Signes and Clavel-Arroitia 2015](#); [Higginbotham and Reid 2019](#); [Kim et al. 2018](#); [Kojima and Yamashita 2014](#); [Kuiken and Vedder 2007](#); [Treffers-Daller et al. 2018](#)). These two aspects aside, we also consider the total number of words (tokens) produced in the speech output since this variable has been identified to be closely related to speaking performance in interactional contexts ([Gan 2012](#); [Iwashita et al. 2008](#); [Noreillie et al. 2020](#)).

1.1.1. Lexical Diversity

Lexical diversity, broadly speaking, pertains to the variety of words employed in written or spoken texts. The term “diversity” has, however, been questioned for its multi-faceted meaning. [Laufer and Nation \(1995\)](#) and [Durán et al. \(2004\)](#) defined lexical diversity as a two-layer construct, including both the range of words and how these words are deployed in texts. [Jarvis \(2013\)](#) defines lexical diversity as a multi-layer construct covering six properties: variability, volume, evenness, rarity, dispersion, and disparity. [Jarvis \(2013\)](#) argued that the prevailing concept of lexical diversity, which examines the range of words used in texts, only considers one component of a multi-layer model. The present study adheres to this definition, known as lexical variability.

To measure lexical diversity, researchers have proposed some basic measurements, including the number of types, TTR (Type–Token Ratio, [Templin 1957](#)), and [Guiraud \(1960\)](#), as well as some advanced measurements, including [Maas \(1971\)](#), D ([Malvern et al. 2004](#)), HD-D (Hypergeometric Distribution-D, [McCarthy and Jarvis 2007](#)), MTLT (Measure of Textual Lexical Diversity, [McCarthy 2005](#); [McCarthy and Jarvis 2010](#)), and MATTR (Moving Average Type–Token Ratio, [Covington and McFall 2010](#)). The goal is to use various mathematical metrics to create a reliable and valid index that accurately represents the “d” in texts, while ensuring that the indices are not affected by the length of the text. Although positive evidence can be found in supporting the use of basic measures, such as types ([Iwashita et al. 2008](#); [Treffers-Daller et al. 2018](#)) and Guiraud ([Daller and Xue 2007](#)), studies working with oral spoken data favour advanced measures because they better compensate text length effects that the basic measure cannot achieve due to their intrinsic quality ([Fergadiotis et al. 2013](#); [Koizumi and In’nami 2012](#); [Kyle et al. 2024](#); [McCarthy and Jarvis 2007, 2010](#)). Through several validity studies that compare advanced and basic measures, researchers have determined that MTLT is the most effective measure for representing diversity in spoken output, as noted by [Koizumi and In’nami \(2012\)](#). Additionally, MATTR has also been recognized as a viable option for measuring diversity, as reported by [Fergadiotis et al. \(2013\)](#) and [Kyle et al. \(2024\)](#). Studies reporting the effectiveness of the D measure are mixed with some studies as supportive ([Foster and Tavakoli 2009](#); [Noreillie et al. 2020](#); [Tavakoli and Foster 2011](#)) and some not, especially when texts are shorter than 150 tokens ([Koizumi and In’nami 2012](#)).

1.1.2. Lexical Sophistication

The level of lexical sophistication is connected to the percentage of infrequent words in the texts being analyzed ([Read 2000](#); [Laufer and Nation 1995](#)). The language acquisition literature provides evidence for this concept, as studies have found that learners typically acquire words in a frequency-based order ([Ellis 2002](#); [Milton 2009](#)). More recently, defining lexical sophistication has been defined as a multidimensional construct including macro-features and micro-features ([Kyle and Crossley 2015](#); [Kim et al. 2018](#)). Following this definition, word frequency is treated as one of the macro-features together with other components, such as contextual distinctiveness, psycholinguistic norms, and word neighbourhood; in the meantime, it can include several micro-features such as content word frequency and function word frequency. According to this definition, word frequency is one of the macro-features, along with contextual distinctiveness, psycholinguistic norms, and word neighbourhood. In this study, we adhere to the common notion of lexical complexity, which focuses solely on the percentage of infrequent words in the provided texts and examines this aspect on a macro-level.

Lexical sophistication measures that rely on frequency depend heavily on extensive corpus data. With the use of different calculation methods, they are generally categorized into count-based indices, such as CELEX norms ([Baayen et al. 1995](#), cited in [Crossley et al. 2013](#)); and band-based indices, such as the LFP ([Laufer and Nation 1995](#)), Lambda ([Meara and Bell 2001](#)), and Advanced Guiraud ([Daller et al. 2003](#)). Count-based measures assign a corpus-based frequency count to each word, and calculate the average frequency count from all words in a text to determine its sophistication index. Band-based measures involve assigning each word a frequency count and then placing it into a frequency band. This can be seen with the K1 band, which includes the first 1000 most frequent words, and the K2 band, which includes the second 1000 most frequent words, and so on. Words classified as K1 and K2 are high-frequency, while words classified as K3 and beyond are advanced. Following this principle, several word lists have been developed in a combination of various corpuses built for different purposes, such as the General Service List (GSL, [West 1953](#)), the Academic Word List (AWL, [Coxhead 2000](#)), and more recently the Academic Spoken Word List (ASWL, [Dang et al. 2017](#)).

1.2. Interactive Factor and L2 Vocabulary Use

Many complexity studies have discussed the influence of interactive factors on vocabulary usage in L2 performance, along with other aspects of language performance, like accuracy, fluency, and grammatical complexity (Gan 2012; Michel et al. 2007; Robinson 2001; Skehan 2009b). The debate about the relationship between interactive factors and linguistic complexity boils down to two opposing views on how attention is divided when cognitive demand increases in tasks. One side argues that vocabulary use is negatively affected because L2 learners have limited attentional capacity (Skehan and Foster 2001), with the other side suggesting a positive impact due to the heightened attention and noticing brought by interaction (Robinson 2001, 2005; Schmidt 1990; Pica 1994).

In contrast to situations where second language (L2) learners only need to be able to produce language, L2 learners in interactive contexts are also expected to possess the ability to understand their conversational partners' responses and engage in negotiation, turn-taking, and other activities that naturally occur in interactive communication. Interactive speaking tasks are therefore commonly regarded as more mentally challenging compared to monologic speaking tasks (Robinson 2001, 2005).

1.2.1. Negative Relationship between Interactive Factor and Vocabulary Use

According to papers on limited working memory (Carter 1998; Gathercole and Baddeley 1993) and Van Patten's (1990) study showing that learners' limited attentional resources allocated to either communication content or language form, Skehan and Foster (Skehan 1998, 2001, 2003; Skehan and Foster 2001) proposed L2 learners' Limited Attentional Capacity Model. The model suggests that when L2 learners are given cognitively demanding tasks, their attention is more likely to be focused on the meaning of the content rather than the forms of the language. As a result, they choose a simpler language that they have already automated, rather than using more complex language. Following this perspective, speaking exercises that lack interaction may prompt learners to utilize less common and more diverse vocabulary, while interactive exercises tend to have the opposite outcome.

Some studies (e.g., Gan 2012; Skehan 2009b) have found evidence supporting a negative correlation between interactive factors and vocabulary use. Drawing on Foster's and Skehan's earlier studies (Foster and Skehan 1996, 2013; Skehan and Foster 1999), Skehan (2009b) compared pre-intermediate to intermediate L2 learners' use of less frequent words in a monologic context (narrative tasks) and dialogic context (decision-making tasks). The consistently higher Lambda score achieved in narrative tasks during the studies shows that narratives encourage L2 learners to utilize less common words, supporting the limited attentional capacity theory. The study, though, mentioned nothing about the diversity of vocabulary usage.

Gan (2012) conducted a study comparing the speaking performance of 30 secondary school students. The study focused on two tasks: individual presentations and group discussions with three or four members. To measure grammatical complexity, Gan employed various measures, including basic production unit measures, like the total number of words spoken. The findings showed that when it came to producing more words and showing grammatical complexity, individual presentations were more effective. As a result, the author concluded that monologic presentations challenged learners in terms of grammatical complexity and lexical processing. The study did not include a thorough assessment of vocabulary usage.

In contrast to the previous two studies, Skehan (2009a) was the only one to consider both lexical sophistication, measured by Lambda, and lexical diversity, measured by D, when discussing lexical complexity. When comparing the vocabulary usage of second language (L2) learners with L1 speakers, the study found that L1 speakers consistently scored higher in both the D score and Lambda score across all tasks, including monologues and dialogues. However, the study did not mention whether there was a significant difference in the vocabulary use of L2 learners in various speaking modes. Nevertheless, the research

highlighted that narrative tasks yielded more consistent lexical discrepancies between L1 and L2 speakers, as opposed to interactive tasks.

1.2.2. Positive Relationship between Interactive Factor and Vocabulary Use

Holding different views towards the limited attentional capacity, researchers who believe in the positive impact from interaction suggest that L2 learners' attention can be directed to language forms based on the shared notice that eventually can help improve L2 complexity (Robinson 2001, 2005; Schmidt 1990; Pica 1994).

According to Robinson's (2001, 2005) Cognition Hypothesis or Triadic Componential Framework, interaction played a significant role in the theory's development. The hypothesis suggests that there are two dimensions of task complexity: resource-directing, which focuses on increasing conceptual demands, and resource dispersing, which focuses on enhancing the ability to access and apply prior knowledge. According to the hypothesis, second language development can be enhanced by a series of carefully planned tasks that gradually increase in complexity, both conceptually and linguistically. Learners' existing knowledge can be activated and accessed during real-time communication through interaction, which can help differentiate L2 performance. Following this hypothesis, by manipulating task complexity along the resource-directing dimension, learners' attention can be directed to language forms and structures so as to trigger more accurate and complex speech, as is found in Robinson's (2001) work, where a complex task generated more use of different words, as measured by the Token–Type Ratio.

Robinson's (2005) hypothesis did not provide clear predictions about how interactive factors affect L2 performance, specifically in terms of vocabulary use. This lack of clarity could be attributed to the intricate nature of the task design, which involves multiple dimensions. Michel et al. (2007) conducted an empirical study to respond to this gap by incorporating two tasks (simple and complex) and implementing both tasks in two different speaking modes (monologues and dialogues). The study included 44 participants who were at the B1 and B2 levels of CEFR in their L2 Dutch language proficiency. The topic of linguistic complexity was examined by considering two aspects: structural complexity (measured by clauses per AS unit and subordination index) and lexical complexity (measured by the percentage of lexical words and Guiraud's index). According to the findings, the level of task complexity had a significant effect on the use of lexical words, resulting in more diverse speech patterns. This observation supports an earlier study conducted by Robinson (2001). However, the condition of the task did not show any substantial difference in lexical complexity, although it did have an impact on measures associated with the complexity of the structure.

According to Robinson (2005), interactive factors have a positive influence on L2 performance. He recognizes that adjusting the difficulty of the task in terms of distributing resources could result in a reduction in the complexity of the language used. When tasks become more complex, interactions tend to increase, creating challenges for learners who are trying to use their language skills due to interruptions from other participants.

So far, we have summarized the few existing empirical studies that focus on vocabulary usage in interactive situations. The depiction of vocabulary in these studies is limited in its perspective. In contrast to the extensive examination of grammatical complexity measures in numerous studies (as noted in Gan 2012), lexical complexity is frequently assessed briefly (e.g., using TTR as described by Robinson 2001), and the discussion on the lexical sophistication of L2 learners' vocabulary use is severely underrepresented. To respond to the two opposing views on the impact of increased interaction on language use, our research seeks to examine the potential effects of enhanced interaction on vocabulary usage, specifically in terms of length, diversity, and sophistication.

1.3. Rationale of the Current Study

Our study addresses the research gaps we have identified in the literature. First, we aim to respond to the gap in understanding vocabulary use in interactional contexts by

collecting data in three speaking modes: monologic, dialogic, and triologic. We propose examining how dialogic and triologic interactive factors, particularly in terms of vocabulary use, affect linguistic complexity. This is an area that has been largely overlooked in previous research, but is a practical concern to language teachers and speaking assessment organizations (e.g., Cambridge ESOL speaking tests, CET 4/6 speaking test, and the speaking component of HKDSE). Second, since there is a lack of research on how L2 learners use complex vocabulary in interactive situations, this study examines both the diversity and sophistication of vocabulary as key factors. We hope the results can add more empirical evidence to the discussion on the impact of the dialogic interactive factor on linguistic complexity, with a particular emphasis on vocabulary usage. The research questions for the current study are therefore the following:

RQ1: To what extent does the interactive factor influence L2 learners' vocabulary use in relation to tokens?

RQ2: To what extent does the interactive factor influence L2 learners' vocabulary use in relation to lexical diversity?

RQ3: To what extent does the interactive factor influence L2 learners' vocabulary use in relation to lexical sophistication?

2. Materials and Methods

2.1. Participants

This study included 24 Chinese undergraduate students who were proficient in English at an advanced level, based on their performance on College English Test Level 4 (CET 4). The average score on CET 4 was 612, which corresponds to the C1-C2 level according to the CEFR (2001). Participants received L2 English instruction from L1 Chinese L2 English teachers for 90 min per week before this study, lasting between two and four semesters. There were 16 weeks of teaching in every semester. The participants did not use English regularly outside of the learning context.

2.2. Speaking Tasks

The speech data were collected from participants responding to a total of 7 persuasive speaking tasks in three conditions (2 for monologues; 2 for dialogues; and 3 for triologues) (see Appendix A for speaking tasks). All speech data were digitally recorded and orthographically transcribed and analyzed using Praat (Boersma and Weenink 2021). The choice of the seven speaking tasks was guided by two main criteria: (a) a persuasive topic from or adapted from a published IELTS writing test; (b) the difficulty of the speaking tasks should be around Level 3 out of 5 on a Likert Scale (with 1 as the easiest and 5 as the hardest). Overall, seven speaking tasks were finalized based on the feedback from 27 students who share a similar education background to the participants.

For the monologic tasks, we gave participants 30 s to prepare and 2 min to complete each given task, as in the work by de Jong and Mora (2019) and Clenton et al. (2021). For the interactive tasks (dialogic and triologic), the first speaker of each group was given 30 s to prepare, while the other participants waited and prepared their responses naturally. Given the number of participants involved in the interactive mode, we allocated 4 min for each dialogic discussion and 6 min for each triologic discussion, allowing 2 min on average for each participant. The moderator reminded participants if there was a clear dominance in the interactive speaking mode.

To reduce the prime effect from the first speaker, participants were required to be the first speaker in turns. Researchers then pruned the transcripts to make sure fillers (e.g., *en*, *err*, *umm*) were removed, abbreviations (e.g., *can't*, *isn't*) were extended, spelling mistakes (e.g., *government*) were corrected, and other discourse features (e.g., Chinese words, researcher's utterances) were deleted.

2.3. Vocabulary Use Measures

The vocabulary use was measured in terms of tokens, diversity, and sophistication. Tokens refer to the number of words produced by participants. To assess the diversity of participants' spoken output, we employed both basic measures, such as types and Guiraud, in line with positive findings found in previous research (Daller and Xue 2007; Iwashita et al. 2008; Treffers-Daller et al. 2018); and advanced measures, such as MTL and MATTR, because of their reported resistance to text length in general. Since the average speech length was above 150 tokens in the current study, passing the threshold of text length sensitivity for the D measure (Koizumi and In'nami 2012), the D measure was also included in the diversity measures. Of all diversity measures, a higher value of lexical variation indicates lower word repetition and, therefore, less recycling of words.

To evaluate how sophisticated participants' spoken output was, we employed count-based and band-based frequency measures of lexical sophistication. Count-based measures were found to offer greater accuracy (Crossley et al. 2013), while band-based measures provided clearer and more easily interpretable results. For the count-based frequency measure, each participant's spoken output was given three indices: (1) spoken-frequency-log British National Corpus (BNC), short for Spoken BNC; (2) spoken-frequency-log Corpus of Contemporary American English (COCA), short for Spoken COCA; and (3) academic spoken-frequency-log COCA, short for Academic COCA.

Each participant's spoken output was assigned two vocabulary profiles in the band-based frequency measure. One profile included the first 1000 frequent words (K1), and the second included 1000 frequent words (K2), and the Academic Words List (AWL). The second profile targets specifically on words used in academic spoken contexts outlined by ASWL (Dang et al. 2017). According to this profile, words are divided into four frequency bands (Level 1, 2, 3, and 4) in line with BNC/COCA frequency levels. We kept scores for AWL, and ASWL Level 3 and Level 4, which include low-frequency words, for the lexical sophistication analysis.

3. Results

The three research questions address the impact of the interactive factor on L2 learners' vocabulary use in relation to three aspects, tokens, lexical diversity, and lexical sophistication. To answer these questions, one-way ANOVA was conducted and the results are shown in Table 2. M1 refers to monologues, M2 refers to dialogues, and M3 refers to triologues.

The results showed that there were no significant differences found in tokens ($p = 0.057$). The same results were also found with five diversity measures, represented by types ($p = 0.110$), Guiraud ($p = 0.398$), D ($p = 0.288$), MATTR ($p = 0.907$), and MTL ($p = 0.417$).

The words belonging to K1 and K2 levels from LFP amounted to 163.33 in M1, 183.27 in M2, and 189.57 in M3, forming approximately 93% in M1, 93.8% in M2, and 96% in M3, in reference to the total words produced. There were significant differences across three speaking modes for K1 ($F(2, 69) = 3.55, p < 0.034$) and K2 ($F(2, 69) = 10.49, p < 0.001$). ASWL Level 1 and ASWL Level 2 also had a similar situation, with 151.73 (93.5%) in M1, 171.13 (93.3%) in M2, and 179.4 (94.6%) in M3, aligning with the ASWL profile. Significant differences were also reported across three speaking modes for ASWL Level 1 ($F(2, 69) = 3.54, p < 0.034$) and ASWL Level 2 ($F(2, 69) = 4.23, p < 0.019$).

In addition, ANOVA reported significant differences with words of advanced levels across three speaking modes, as reflected by one count-based measure, Spoken COCA ($F(2, 69) = 137.06, p < 0.001$), and two band-based measures, AWL ($F(2, 69) = 10.9, p < 0.001$) and ASWL 3 ($F(2, 69) = 12.66, p < 0.001$). To further examine this difference, we conducted a post hoc analysis. Given that all groups had equal sample sizes and homogeneity of variance was met, the Tukey method was selected (Field 2009) and the results are shown in Table 3. Since K1, K2, and ASWL Level 1 and Level 2 are words of a high-frequency level, they were excluded in the post hoc lexical sophistication analysis.

Table 2. ANOVA for vocabulary use measures in three speaking modes.

Vocabulary Use	M1		M2		M3		F	p
	Mean	SD	Mean	SD	Mean	SD		
Tokens	162.33	43.06	183.27	36.36	189.57	46.28	2.98	0.057
Lexical diversity								
Types	71.75	13.67	79.46	11.12	77.29	13.75	2.28	0.110
Guiraud	5.76	0.50	5.89	0.44	5.72	0.46	0.94	0.398
D	43.31	7.92	47.26	7.87	45.30	9.82	1.27	0.288
MATTR	0.66	0.04	0.66	0.04	0.66	0.04	0.10	0.907
MTLT	34.69	9.81	32.68	7.52	31.57	7.02	0.89	0.417
Lexical sophistication								
<i>Count-based</i>								
Spoken BNC	0.25	0.08	0.26	0.07	0.28	0.09	0.92	0.402
Spoken COCA	2.90	0.10	3.23	0.06	3.25	0.08	137.06	0.001
Academic COCA	3.09	0.07	3.04	0.10	3.05	0.07	2.53	0.087
<i>Band-based</i>								
K1	145.12	40.02	165.75	33.95	173.32	43.05	3.55	0.034
K2	6.52	2.45	6.31	2.85	9.40	2.52	10.49	0.001
AWL	2.18	0.60	2.49	0.89	1.41	0.94	10.90	0.001
ASWL Level 1	141.13	39.07	163.40	33.33	168.96	42.12	3.54	0.034
ASWL Level 2	10.60	3.74	7.73	3.43	10.44	4.33	4.23	0.019
ASWL Level 3	3.35	2.28	5.27	2.86	1.81	1.95	12.66	0.001
ASWL Level 4	1.18	1.72	1.94	1.72	2.36	1.31	0.78	0.462

Table 3. Post hoc analysis.

	Speaking Mode	Speaking Mode	Mean Difference	Sig.
Spoken COCA	M1	M2	−3112.90 *	<0.001
		M3	−2762.15 *	<0.001
	M2	M1	3112.90 *	<0.001
		M3	350.74	0.524
	M3	M1	2762.15 *	<0.001
		M2	−350.74	0.524
AWL	M1	M2	−0.31	0.395
		M3	0.76 *	0.005
	M2	M1	0.31	0.395
		M3	1.07 *	<0.001
	M3	M1	−0.76 *	0.005
		M2	−1.07 *	<0.001
ASWL Level 3	M1	M2	−1.92 *	0.019
		M3	1.55	0.071
	M2	M1	1.92 *	0.019
		M3	3.47 *	<0.001
	M3	M1	−1.55	0.071
		M2	−3.47 *	<0.001

The post hoc analysis with the Tukey test showed that the mean score for Spoken COCA, the count-frequency measure, in monologues (M = 2.90, SD = 0.10) was significantly lower ($p < 0.001$) than that in dialogues (M = 3.23, SD = 0.06) and in trialogues (M = 3.25, SD = 0.08). No significant difference was found between two interactive conditions.

In terms of the band-based frequency measure, AWL scores (M = 2.49, SD = 0.89) and ASWL Level 3 scores (M = 5.27, SD = 2.86) in dialogues were both significantly higher ($p < 0.001$) than their scores in trialogues (M = 1.41 and SD = 0.94 for AWL, and M = 1.81 and SD = 1.95 for ASWL Level 3). ASWL Level 3 in M2 (M = 5.27, SD = 2.86) achieved a significantly higher value ($p < 0.05$) than that in M1 (M = 3.35, SD = 2.28) and AWL in M1 achieved a significantly higher value ($p < 0.05$) than that in M3 (M = 1.41, SD = 0.94).

4. Discussion

The following discussion outlines the ANOVA and post hoc analysis results and their implications for the three research questions. The first question concerns the impact of the interactive factor on the total number of words produced. The second research question concerns the impact of the interactive factor on the use of different words, and the third research question concerns the impact of the interactive factor on the use of advanced words.

4.1. Interactive Factor on Vocabulary Use in Relation to Tokens

The data showed that the interactive factor has no impact on the current advanced L2 learners' vocabulary use regarding the total words produced. In all speaking scenarios, monologues, dialogues, and trialogues, current L2 learners consistently used a set number of words. Unlike Gan (2012), who discovered that narrative monologic speaking tasks resulted in higher word production compared to group discussions with three or four members, our findings show contrasting results. Gan (2012) attributed this to the communicative pressure faced by secondary school students with an elementary or low-intermediate level of English. Gan (2012) found that learners experienced higher levels of anxiety in interactive situations due to their limited language skills. The current paper results appear subject to the same reasoning.

In contrast to Gan's (2012) study, where participants had elementary or low-intermediate English proficiency, participants in the present study had an advanced level of English. The latter would have better language knowledge and be in an advantageous position for processing basic language skills in interactive activities. Unlike Gan's (2012) study that included secondary school students, our study focused on adult learners with a higher education background. The current participants may find communicative pressure less daunting compared to younger learners with less life experience.

One more factor that could account for the consistent speech duration is this study's design, which mandated that participants have a fixed speaking time of approximately 2 min, regardless of the speaking modes. Furthermore, participants were reminded if there was a clear preference for the interactive speaking mode. If participants have similar fluency levels in different speaking modes, as noted by Wang-Taylor and Clenton (Forthcoming, under review), it is likely that they will consistently produce a similar number of words within the same time frame.

4.2. Interactive Factor on Vocabulary Use in Relation to Diversity

When examining the impact of the interactive factor on the vocabulary usage of L2 learners, the data indicated that it had no significant influence. This was evident from both basic diversity measures and advanced measures. This is in line with the results of Michel et al. (2007), who found that the task condition (monologues vs. dialogues) has no impact on lexical complexity measured via the percentage of lexical words and Guiraud in their study. Despite that those participants in Michel et al.'s (2007) study were of different proficiency levels to the current study, intermediate (B1-B2) for the former and advanced (C1-C2) for the latter, the current study continues revealing the consistency of L2 learners' choice of different words regardless of the speaking modes.

One possible explanation for this finding is that the interactive factor may have helped second language learners focus on meaningful communication, leading them to continue using the same words they were already familiar with from their monologues. This finding supports Skehan's and Foster's Limited Attentional Capacity Model (Skehan 1998, 2001, 2003; Skehan and Foster 2001).

4.3. Interactive Factor on Vocabulary Use in Relation to Sophistication

This study revealed that the interactive aspect has a strong influence on learners' use of advanced vocabulary, particularly when comparing dialogues and monologues. Results from count-based lexical sophistication measures showed that participants generated

more advanced words in dialogic and triologic interactive tasks compared to monologic tasks, although it was unclear which interactive speaking mode was most effective in eliciting advanced words. Results from band-based lexical sophistication measures showed that dialogues generated more advanced words than monologues, as represented by the significantly higher value of ASWL Level 3 in M2 than that in M1. This is in line with results found with the count-based measure. However, the results indicated that monologues produced more sophisticated vocabulary compared to trialogues, as evidenced by the significantly higher Average Word Length (AWL) in M1 compared to M3. This contradicts the findings observed with the count-based measure.

The consistent results from two lexical sophistication measures, count-based and band-based, revealed that dialogues generated more advanced words than other speaking modes, confirming the positive link between the interactive factor and vocabulary use, as proposed by researchers who believe in the Cognition Hypothesis (Robinson 2001, 2005), Noticing Hypothesis (Schmidt 1990), and heightened attention from negotiation (Pica 1994). The use of advanced vocabulary seems to be heightened through dialogic interaction.

Yet, the situation regarding the use of advanced vocabulary in monologues and trialogues is not straightforward. While a count-based frequency analysis shows a positive influence of interactive factors on the use of advanced words, a band-based frequency analysis presents contradictory results. The distinction between dialogues and trialogues lies in how two interactive speaking modes can be interpreted. While a triologue shares interactive characteristics with dialogues, the inclusion of a third interlocutor can introduce additional factors to consider, such as speech order, which may not be relevant in dialogues. In dialogues, when one person pauses, the other person will immediately perceive the potential for a shift in speaking roles and adjust their actions accordingly. Trialogues require the extensive negotiation of speech order, which may decrease both speaking time and speech quality. The level of interactivity in trialogues may not be as high as it is in dialogues when it comes to using advanced words.

The findings of this study may be restricted by a few limitations. Initially, it is necessary to expand sample size in this study so it has more statistical power. This study is a pilot study that is based on 24 participants. To apply the results to high-stakes assessment that uses or considers using dialogic and triologic interactive speaking tasks, a larger sample size is perhaps needed. Second, it is also necessary to expand the participant pool in this study to encompass a broader spectrum of learners' proficiency. This study focuses exclusively on advanced learners, narrowing down the participants under investigation. Future studies should consider exploring how interactive factors impact vocabulary use and its relationship with vocabulary knowledge across different proficiency levels. Lastly, the use of the speaking tasks is limited to persuasive speaking tasks. Including other speaking task types in future studies may be necessary, as different task types can affect L2 learners' performance (Skehan 2009a).

5. Conclusions

Through the current study, we have gained a better understanding of how vocabulary is utilized in interactive contexts. The main significance lies in illustrating how the choice of words changes not just between monologues and dialogues, but also between two different interactional contexts. Findings of this study identified that, first, the interactive factor has a positive impact on the use of advanced words but it has no impact on the total words produced and the words' diversity, partially supporting the benefit of heightened attention (Robinson 2001, 2005; Schmidt 1990; Pica 1994) and partially supporting L2 learners' limited attentional capacity (Skehan 1998, 2001, 2003; Skehan and Foster 2001). Second, among the three speaking modes, the dialogic speaking mode is the most effective condition for encouraging L2 learners to use advanced vocabulary. Third, the use of vocabulary can differ between dialogues and trialogues, owing to their fundamental differences. To accurately capture the two distinct conditions that led to significant differences in L2

learners' performance, we recommend using the terms 'dialogic interactive factor' and 'trialogic interactive factor' instead of the generic term 'interactive factor'.

Two practical implications arise from the current study's findings. To enhance pedagogy, instructors should promote the use of dialogic interactive speaking modes in classrooms to engage L2 learners and focus on less common vocabulary. In speaking tests with limited resources, it may be beneficial to incorporate interactive speaking modes like dialogues and triologies, particularly for advanced L2 learners who consistently produce a substantial amount of words and use similar vocabulary.

Author Contributions: Conceptualization, Y.W.-T.; methodology, Y.W.-T.; software, Y.W.-T.; validation, Y.W.-T., Y.R. and J.C.; formal analysis, Y.W.-T.; investigation, Y.W.-T.; resources, Y.W.-T.; data curation, Y.W.-T. and Y.R.; writing—original draft preparation, Y.W.-T.; writing—review and editing, Y.W.-T. and J.C.; visualization, Y.W.-T. and J.C.; supervision, Y.W.-T. and J.C.; project administration, Y.W.-T. and Y.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study has been approved with regards to ethical concerns, following consultation with the School of Foreign Studies, Nankai University (4 March 2021)).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to participants' privacy.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

Speaking tasks

Monologic speaking tasks

Task1. People today can shop/work/communicate with others via the Internet instead of face to face. Is this an improvement of the society?

Task2. As a college student, do you think peer pressure is a positive thing or a negative thing?

Dialogic speaking tasks

Task1. Cycling is more environmental friendly than other forms of transportation, but many people still prefer to use cars instead of bikes. How to increase the popularity of cycling in your opinion?

Task2. Many people use social media to get in touch with people. Do you think this is a good habit? To what extent do you think this will affect their communicating ability when people are talking with each other face to face?

Triologic speaking tasks

Task1. Some people think that the best way to be successful in life is to get a university education. To what extent do you agree or disagree with this?

Task2. Any student caught cheating on an examination should be automatically dismissed from college. To what extent do you agree or disagree, why?

Task3. In the future, nobody will buy printed books because they will be able to read everything they want online without paying. To what extent do you agree or disagree with this statement?

References

- Appel, Randy, Pavel Trofimovich, Kazuya Saito, Talia Isaacs, and Stuart Webb. 2019. Lexical aspects of comprehensibility and nativeness from the perspective of native-speaking English raters. *International Journal of Applied Linguistics* 170: 24–52. [CrossRef]
- Baayen, R. H., R. Piepenbrock, and L. Gulikers. 1995. *The CELEX Lexical Database (Release 2)*. Philadelphia: Linguistic Data Consortium.

- Boersma, Paul, and David Weenink. 2021. *Praat: Doing Phonetics by Computer [Computer Program]*, Version 6.1.42. Available online: www.praat.org/ (accessed on 5 November 2021).
- Cambridge Language Assessment. 2023. Cambridge Speaking Test Suite. Available online: <https://www.cambridgeenglish.org/exams-and-tests/> (accessed on 11 May 2023).
- Cambridge Speaking Test Suite. 2024. CEFR Vocabulary Descriptors. Available online: <https://cefrlevels.com/descriptors/vocabulary/> (accessed on 20 April 2024).
- Carter, Rita. 1998. *Mapping the Mind*. London: Weidenfeld and Nicolson.
- CEFR. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Available online: <https://rm.coe.int/16802fc1bf> (accessed on 25 April 2024).
- Clenton, Jon, Nivja H. de Jong, Dion Clingwall, and Simon Fraser. 2021. 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: Pedagogy, Practice, and Implications for Teaching Vocabulary*. Edited by Jon Clenton and Paul Booth. Abingdon and New York: Routledge, pp. 126–45.
- Covington, Michael, and Joe McFall. 2010. Cutting the Gordian knot: The moving-average type-token ratio (MATTR). *Journal of Quantitative Linguistics* 17: 94–100. [CrossRef]
- Coxhead, Averil. 2000. A new academic word list. *TESOL Quarterly* 34: 213–38. [CrossRef]
- Crossley, Scott, Tom Cobb, and Danielle McNamara. 2013. Comparing count-based and band-based indices of word frequency: Implications for active vocabulary research and pedagogical applications. *System* 41: 965–81. [CrossRef]
- Crossley, Scott, Tom Salsbury, and Danielle McNamara. 2015. Assessing lexical proficiency using analytic ratings: A case for collocation accuracy. *Applied Linguistics* 36: 570–90. [CrossRef]
- Daller, Helmut, and Huijuan Xue. 2007. Lexical richness and the oral proficiency of Chinese EFL students. In *Modelling and Assessing Vocabulary Knowledge*. Edited by Helmut Daller and James Milton. Cambridge: Cambridge University Press, pp. 150–64.
- Daller, Helmut, Roeland van Hout, and Jeanine Treffers-Daller. 2003. Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics* 24: 197–222. [CrossRef]
- Dang, Thi Ngoc Yen, Averil Coxhead, and Stuart Webb. 2017. The academic spoken word list. *Language Learning* 67: 959–97. [CrossRef]
- de Jong, Nivja H., and Joan Mora. 2019. Does having good articulatory skills lead to more fluent speech in first and second languages? *Studies in Second Language Acquisition* 41: 227–39. [CrossRef]
- Durán, Pilar, David Malvern, Brian Richards, and Ngoni Chipere. 2004. Developmental trends in lexical diversity. *Applied Linguistics* 25: 220–42. [CrossRef]
- Ellis, Nick. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24: 143–88. [CrossRef]
- Fergadiotis, Gerasimos, Heather Wright, and Thomas M. West. 2013. Measuring lexical diversity in narrative discourse of people with aphasia. *American Journal of Speech-Language Pathology* 22: 1–26. [CrossRef] [PubMed]
- Field, Andy. 2009. *Discovering Statistics Using SPSS*, 3rd ed. London: Sage Publications Ltd.
- Foster, Pauline, and Parvaneh Tavakoli. 2009. Native speakers and task performance: Comparing effects on complexity, fluency, and lexical diversity. *Language Learning* 59: 866–96. [CrossRef]
- Foster, Pauline, and Peter Skehan. 1996. The influence of planning and task type on second language performance. *Studies in Second Language Acquisition* 18: 299–323. [CrossRef]
- Foster, Pauline, and Peter Skehan. 2013. Anticipating a post-task activity: The effects on accuracy, complexity and fluency of second language performance. *The Canadian Modern Language Review* 69: 249–73. [CrossRef]
- Gan, Zhengdong. 2012. Complexity measures, task type, and analytic evaluations of speaking proficiency in a school-based assessment context. *Language Assessment Quarterly* 9: 133–51. [CrossRef]
- Gathercole, Susan, and Ala Baddeley. 1993. *Working Memory and Language*. Hillsdale: Lawrence Erlbaum.
- Gregori-Signes, Carmen, and Begoña Clavel-Arroitia. 2015. Analysing lexical density and lexical diversity in university students' written discourse. *Procedia-Social and Behavioral Sciences* 198: 546–56. [CrossRef]
- Guiraud, Pierre. 1960. *Problèmes et méthodes de la statistique linguistique [Problems and Methods of Linguistic Statistics]*. Paris: Presses Universitaires de France.
- Higginbotham, George, and Jacqui Reid. 2019. The lexical sophistication of second language learners' academic essays. *Journal of English for Academic Purposes* 37: 127–40. [CrossRef]
- HKDSE. 2023. Hong Kong Diploma of Secondary Education. Available online: https://www.hkeaa.edu.hk/en/hkdse/assessment/assessment_framework/ (accessed on 11 April 2024).
- IELTS. 2024. Speaking Band Descriptors. Available online: https://takeielts.britishcouncil.org/sites/default/files/ielts_speaking_band_descriptors.pdf (accessed on 20 April 2024).
- Iwashita, Noriko, Annie Brown, Tim McNamara, and Sally O'hagan. 2008. Assessed levels of second language speaking proficiency: How distinct? *Applied linguistics* 29: 24–49. [CrossRef]
- Jarvis, Scott. 2013. Defining and measuring lexical diversity. In *Vocabulary Knowledge: Human Ratings and Automated Measures*. Edited by Scott Jarvis and Michael Daller. Amsterdam: John Benjamins Publishing Company, pp. 13–43.
- Kim, Minkyung, Scott Crossley, and Kristopher Kyle. 2018. Lexical sophistication as a multidimensional phenomenon: Relations to second language lexical proficiency, development, and writing quality. *The Modern Language Journal* 102: 120–41. [CrossRef]

- Koizumi, Rie, and Yo In'nami. 2012. Effects of text length on lexical diversity measures: Using short texts with less than 200 tokens. *System* 40: 554–64. [CrossRef]
- Kojima, Masumi, and Junko Yamashita. 2014. Reliability of lexical richness measures based on word lists in short second language productions. *System* 42: 23–33. [CrossRef]
- Kuiken, Folkert, and Ineke Vedder. 2007. Task complexity and measures of linguistic performance in L2 writing. *International Review of Applied Linguistics in Language Teaching* 45: 261–84. [CrossRef]
- Kyle, Kristopher, and Scott Crossley. 2015. Automatically assessing lexical sophistication: Indices, tools, findings, and application. *TESOL Quarterly* 49: 757–86. [CrossRef]
- Kyle, Kristopher, Hakyung Sung, Masaki Eguchi, and Fred Zenker. 2024. Evaluating evidence for the reliability and validity of lexical diversity indices in L2 oral task responses. *Studies in Second Language Acquisition* 46: 278–99. [CrossRef]
- Laufer, Batia, and Paul Nation. 1995. Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics* 16: 307–22. [CrossRef]
- Li, Hui, and Nuria Lorenzo-Dus. 2014. Investigating how vocabulary is assessed in a narrative task through raters' verbal protocols. *System* 46: 1–13. [CrossRef]
- Lu, Xiaofei. 2012. The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal* 96: 190–208. [CrossRef]
- Maas, Heinz-Dieter. 1971. Über den Zusammenhang zwischen Wortschatzumfang und Länge eines Textes [On the connection between vocabulary breadth and text length]. *Zeitschrift Für Literaturwissenschaft Und Linguistik* 2: 73–96.
- Malvern, David, Brian Richards, Ngoni Chipere, and Pilar Durán. 2004. *Lexical Diversity and Language Development: Quantification and Assessment*. New York: Palgrave Macmillan.
- McCarthy, Philip. 2005. An Assessment of the Range and Usefulness of Lexical Diversity Measures and the Potential of the Measure of Textual, Lexical Diversity (MTLD). Doctoral dissertation, The University of Memphis, Memphis, TN, USA.
- McCarthy, Philip, and Scott Jarvis. 2007. vocd: A theoretical and empirical evaluation. *Language Testing* 24: 459–88. [CrossRef]
- McCarthy, Philip, and Scott Jarvis. 2010. MTL, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods* 42: 381–92. [CrossRef]
- Meara, Paul, and Huw Bell. 2001. P-Lex: A Simple and Effective Way of Describing the lexical Characteristics of Short L2 Tests. *Prospect* 16: 5–19.
- Michel, Marije, Folkert Kuiken, and Ineke Vedder. 2007. The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics in Language Teaching* 45: 241–59. [CrossRef]
- Milton, James. 2009. *Measuring Second Language Vocabulary Acquisition*. Bristol: Multilingual Matters.
- Noreillie, Ann-Sophie, Piet Desmet, and Elke Peters. 2020. Factors predicting low-intermediate French learners' vocabulary use in speaking tasks. *The Canadian Modern Language Review* 76: 194–217. [CrossRef]
- Pica, Teresa. 1994. Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning* 44: 493–527. [CrossRef]
- Read, John. 2000. *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Robinson, Peter. 2001. Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics* 22: 27–57. [CrossRef]
- Robinson, Peter. 2005. Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching* 43: 1–32. [CrossRef]
- Saito, Kazuya, Stuart Webb, Pavel Trofimovich, and Talia Isaacs. 2016. Lexical profiles of comprehensible second language speech: The role of appropriateness, fluency, variation, sophistication, abstractness, and sense relations. *Studies in Second Language Acquisition* 38: 677–701. [CrossRef]
- Schmidt, Richard. 1990. The role of consciousness in second language learning. *Applied Linguistics* 11: 129–58. [CrossRef]
- Skehan, Peter. 1998. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Skehan, Peter. 2001. Tasks and language performance assessment. In *Researching Pedagogic Tasks, Second Language Learning, Teaching and Testing*. Edited by Martin Bygate and Peter Skehan. Harlow: Longman, pp. 167–85.
- Skehan, Peter. 2003. Task-based instruction. *Language Teaching* 36: 1–14. [CrossRef]
- Skehan, Peter. 2009a. Lexical performance by native and non-native speakers on language-learning tasks. In *Vocabulary Studies in First and Second Language Acquisition*. Edited by Brian Richards, Michael Daller, David Malvern, Paul Meara, James Milton and Jeanine Treffers-Daller. London: Palgrave Macmillan, pp. 107–24.
- Skehan, Peter. 2009b. Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics* 30: 510–32. [CrossRef]
- Skehan, Peter, and Pauline Foster. 1999. The influence of task structure and processing conditions on narrative retelling. *Language Learning* 49: 93–120. [CrossRef]
- Skehan, Peter, and Pauline Foster. 2001. Cognition and tasks. In *Cognition and Second Language Instruction*. Edited by Peter Robinson. Cambridge: Cambridge University Press, pp. 183–205. [CrossRef]
- Tavakoli, Parvaneh, and Pauline Foster. 2008. Task design and second language performance: The effect of narrative type on learner output. *Language Learning* 58: 439–73. [CrossRef]

- Tavakoli, Parvaneh, and Pauline Foster. 2011. Task design and second language performance: The effect of narrative type on learner output. *Language Learning* 61: 37–72. [CrossRef]
- Templin, Mildred. 1957. *Certain Language Skills in Children*. Minneapolis: University of Minneapolis.
- TOEFL iBT. 2024. Speaking Scoring Guide Flyer. Available online: <https://www.ets.org/pdfs/toefl/toefl-ibt-speaking-rubrics.pdf> (accessed on 20 April 2024).
- TOEIC. 2024. TOEIC Speaking and Writing Score Descriptors. Available online: <https://www.ets.org/pdfs/toeic/toeic-speaking-writing-score-descriptors.pdf> (accessed on 20 April 2024).
- Treffers-Daller, Jeanine, Patrick Parslow, and Shirley Williams. 2018. Back to basics: How measures of lexical diversity can help discriminate between CEFR levels. *Applied Linguistics* 39: 302–27. [CrossRef]
- Van Patten, Bill. 1990. Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition* 12: 287–301. [CrossRef]
- Wang-Taylor, Yixin, and Jon Clenton. Forthcoming. Second language learners' vocabulary size and speaking fluency in interactional contexts. *TESOL Quarterly*. under review.
- WENR. 2023. World Education News and Reviews. Available online: <https://wenr.wes.org/2018/08/an-introduction-to-chinas-college-english-test-cet> (accessed on 11 May 2023).
- West, Michael. 1953. *A General Service List of English Words*. London: Longman.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.