

A Corpus-Based Study on the Development of Lexical Diversity and Lexical Sophistication of Argumentative Writings

Ziyi Zhu^{1,a}, Xianwei Pu^{1,b*}

¹School of Foreign Languages, Nanjing University of Science and Technology, Nanjing, 210018, China

^azzyisense@163.com, ^bphil667@163.com

*Corresponding author

Abstract: Based on the Written English Corpus of Chinese Learners (WECCL) and the Louvain Corpus of Native English Essays (LOCNESS), this study investigated the development trends of lexical diversity and lexical sophistication in the argumentative essays written by Chinese English majors from freshman to senior year and undergraduate students speaking English as a native language, as well as examined six lexical diversity measures and six lexical sophistication indices to determine whether they can effectively discriminate among the five different English language proficiency levels. As the results suggested, in terms of lexical diversity development, these indices either showed non-linear trends or almost no trend; in terms of indices validity, no diversity indices can distinguish different adjacent proficiency levels. As for lexical sophistication development, the trends were generally classified into two categories: (1) linear development; and (2) fluctuating development; as for indices validity, K1, AWL, OFF, and S can distinguish English language proficiency. In conclusion, this study indicated that the validity of lexical diversity in assessing English L2 learners' language proficiency was weak, while lexical sophistication can effectively assess English L2 learners' language proficiency.

Keywords: Argumentative writing, Lexical diversity, Lexical sophistication, Development trend

1. Introduction

As writing is one of the fundamental abilities that L2 students were expected to master, it was considered of great significance to language acquisition. On the basis of the increasing interest in second language acquisition (SLA) research field, researchers began to put eyes on second language (L2) writing since the 1960s. It was since the 1980s that more and more scholars, both domestic and international, have focused on L2 writing and conducted numerous studies, such as Astika (1993) and Wang and Wen (2002)^{[1][2]}. On account that L2 writing includes multifaceted aspects, research has been conducted on diverse areas involving genre, feedback, correction, individual differences, writing proficiency, and so on. As pointed out by Kyle and Crossley (2016), it is widely verified that lexical knowledge, as being one of the features of writing proficiency being investigated, has played a pivotal role in successful writing^[3]. Many researchers, such as Cumming (1990) and Leki and Carson (1994) also demonstrated that L2 learners tend to pay considerable attention to lexical issues when constructing essays^{[4][5]}. In order to analyze characteristics of lexical use, Yule (1944) introduced the term lexical richness, which has subsequently been enriched by various researchers^[6]. Different researchers maintained diverse views on the constructs of lexical richness. Laufer (1991) classified lexical variation, also known as lexical diversity, lexical density, lexical sophistication, and lexical originality as the four dimensions to measure lexical richness^[7]. Engber (1995) introduced the error dimension and reclassified the measures into lexical diversity with error and without error, proportion of error, and lexical density^[8]. Read (2000) removed lexical originality and measured richness from four aspects: lexical variation, lexical density, lexical sophistication, and a small number of lexical errors^[9]. Bao (2008) employed lexical diversity, lexical density, lexical sophistication, and lexical originality as measurement indicators^[10]. Although the constructs of lexical richness varied among different researchers, lexical diversity and lexical sophistication have always been considered in the context of lexical richness.

Considering the importance of lexical diversity and sophistication, this paper seeks to explore the developmental trends in lexical diversity and sophistication of argumentative essays written by Chinese English majors from the first year to the fourth year and British or American university students.

2. Literature Review

Lexical diversity links to producing more unique words in a text and is an unambiguous component of lexical richness. Laufer and Nation (1995) referred it to type/token ratio, i.e. the fraction of unique words (types) in the language output to the total number of running words (tokens)^[11]. Jarvis (2013) regarded lexical diversity as the opposite of the word repetition rate in a text since the term was frequently concretized as measures targeting at calculating the proportion of words in a language production that were not repetitions of words that appeared in the preceding section^[12]. A variety of lexical diversity indices have been employed to indicate the lexical expertise of the language producer for skilled writers may master a larger number of lexical items from which to choose, allowing them to utilize more diverse and unique words to accomplish their tasks, while relatively unskilled writers tend to repeat a small number of words to fill in the text. There was a consensus achieved among researchers that the type-token ratio (TTR), also known as the simple type-token ratio, was the most fundamental, well-known, and widely-used lexical diversity measure. But its negative relationship with text length inevitably dragged him down, putting its credibility in a poor situation. To overcome the credibility dilemma, some researchers devoted themselves to transforming TTR into an index with less affection from text length, thus creating several variants, such as Maas' Index (Maas) created by Maas (1972) and Moving Average Type-Token Ratio (MATTR) proposed by Covington and McFall in 2010, which were proven to be less affected by text length^{[13][14]}. Some other researchers sought to develop other lexical diversity measures of high reliability and validity. Measure of Textual Lexical Diversity (MTLD) and HD-D were the results of it, which were created by McCarthy and Jarvis in 2007 and in 2010^{[15][16]}.

Lexical sophistication is also indispensable when mentioning lexical richness. Laufer and Nation (1995) regarded it as the proportion of "advanced" vocabulary applied in writing^[11]. Read (2000) stated that lexical sophistication was about the percentage of low-frequency words (or unusual words) that more precisely suited the topic and style of the text in language outputs, instead of utilizing everyday common words, a claim that has received widespread support from researchers^[9]. During decades of research, there has reached a consensus that frequency was a reliable measure to interpret lexical sophistication. Less frequently occurred words are considered sophisticated while more frequently occurred ones are considered not so advanced. Therefore, texts containing a higher percentage of low-frequency words, whether spoken or written, are regarded to be more sophisticated compared to those with more high-frequency words. Laufer and Nation (1995) created the Lexical Frequency Profile (LFP), which consists of four dimensions generating four scores, namely 1K (the most frequently used 1000 words), 2K (the next most frequently used 1000 words), AWL (Academic word list) and OFF List (words not in the previous three lists), to calculate frequency^[11]. Although it is an effective method of good reliance, researchers have not given up creating more efficient measures. For example, Meara and Bell coined P-Lex in 2001, and Kojima and Yamashita introduced S in 2014^{[17][18]}. The two indices are both based on LFP, but further improved the LFP method, outputting a straightforward single score, thus making data interpretation easier.

Considering all of the shreds of evidence, a considerable amount of studies have put their eyes on L2 learners' lexical diversity and lexical sophistication. But still, there remain several aspects of lexical diversity and sophistication about which relatively little is known. Besides, some relatively new measuring indices still wait to be applied in more studies and have their reliability examined. Thus this paper not only explores the developmental trend in lexical diversity and sophistication of argumentative essays written by Chinese English majors from the first year to fourth year and British or American university students but also investigates several lexical diversity and sophistication indices to select useful ones that are capable of discriminating writings from different English proficiency levels.

3. Methodology

3.1. Research Question

Since the present study aims to trace the development trend of lexical diversity and sophistication reflected in argumentative essays written by both domestic and foreign student writers, and to investigate the applicability of lexical diversity and sophistication measures in discriminating English proficiency levels among different English writers, the research questions were proposed as below:

(1) What are the development patterns of the lexical diversity of argumentative essays across different English proficiency levels? Which lexical diversity indices can significantly distinguish adjacent English proficiency levels?

(2) What are the development patterns of the lexical sophistication of argumentative essays across different English proficiency levels? Which lexical sophistication indices can significantly distinguish adjacent English proficiency levels?

3.2. Indices Employed by the Study

The current study employed Maas_tr, MATTR50, HD-D42, MTLD_original, MTLD_ma_bi, and MTLD_ma_wrap, which are the lexical diversity indices, as well as K1, K2, AWL, OFF-LIST (OFF), P-Lex, and S, which are lexical sophistication indices. The above indices were selected mainly because most chosen measures were relatively new compared to classic measures. Precisely because they are new, their application is relatively underdeveloped, resulting in these indices being less researched and their reliability and validity being relatively unknown. Therefore, these measures were selected in this study with the aim of examining them effectively.

3.3. Corpus

In the present study, 452 argumentative texts on two topics were searched from the Written English Corpus of Chinese Learners (WECCL) and the Louvain Corpus of Native English Essays (LOCNESS). The cohort was divided into five groups according to English proficiency levels, which were Year 1 (Y1, 128 Chinese English major essays from grade 1), Year 2 (Y2, 113 from grade 2), Year 3 (Y3, 110 from grade 3), Year 4 (Y4, 68 from grade 4), and Native group (33 university students from the UK or the US). More information about adopted research texts is presented in Table 1.

WECCL, as a sub-component of Spoken and Written English Corpus of Chinese Learners (SWECCL), is a far-reaching corpus created by Professor Wen Qiufang. As noted by Wen (2006), the WECCL has comprised as many as 3059 time-limited or non-time-limited composition texts composed by English majors as well as some non-English majors who were aged from 18 to 23 and studied at universities at all levels across China, with a total of 17 argumentative topics involved^[19]. The Louvain Corpus of Native English Essays (LOCNESS), which was organized by S. Granger and S. De Cock, collected timed and non-timed 436 argumentative texts on about 100 topics, with participants being native English speakers recruited from different universities, covering the UK and the US.

Table 1: Brief introduction of research texts.

Corpora	Group	Number	Average essay length	Token	Type	Average type
WECCL	Grade1 (Y1)	128	378.52	48,451	21,288	166.31
	Grade2 (Y2)	113	350.79	39,639	18,082	160.02
	Grade3 (Y3)	110	365.20	40,172	18,727	170.25
	Grade4 (Y4)	68	321.28	21,847	10,378	152.62
LOCNESS	Native	33	582.88	19,235	8,118	246.00

3.4. Data Analysis

All the work on the computer was carried out using AntWordProfiler, TAALED, P-Lex (<https://www.lognostics.co.uk>), S (<https://www.kojima-vlab.org>), Microsoft Office Excel, and SPSS 22.0. In the first stage, 452 argumentative essay samples were obtained from WECCL and LOCNESS respectively. Before analyzing the data, document names of all samples were reset (.text files, with no space in filenames) and were compressed into five zip folders (.zip format) according to their groups. In the second stage, the five folders were uploaded into TAALED, AntWordProfiler, P-Lex, and S respectively. For measuring Maas_tr, MATTR50, HD-D42, MTLD_original, MTLD_ma_bi, and MTLD_ma_wrap, TAALED was adopted; for K1, K2, AWL, OFF-LIST (OFF), AntWordProfiler; for P-Lex, P-Lex; for S, S. Once those online tools automatically generated the data results, the five output tables were integrated into one Excel file.

Statistical analysis was completed in the final stage using SPSS software version 22.0 and Microsoft Office Excel. The initial step in this procedure was to do a compare-means analysis to calculate the mean for each index in the Y1, Y2, Y3, Y4, and Native groups. Then Microsoft Office Excel would plot the development trend of the five groups. The last step was to generate a p-value for each index between every two adjacent groups using Factorial ANOVA to investigate which indices can significantly discriminate English proficiency levels.

4. Results and Discussion

4.1. Lexical Diversity

To examine lexical diversity indices, the texts went to TAALED, and the mean value of each index of each essay was calculated by SPSS 22.0. Their developmental plots were generated with Microsoft Office Excel. At last, the F-value and corresponding p-value of each index on the overall difference among groups were generated by Factorial ANOVA using SPSS 22.0.

Table 2: Means, F-values and p-values of lexical diversity measures.

Measure	Y1	Y2	Y3	Y4	Native	F	p
Maas ttr	.053	.052	.050	.052	.048	4.787	.001**
MATTR50	.774	.779	.783	.778	.760	3.495	.008**
HD-D42	.812	.811	.813	.803	.794	4.416	.002**
MTLD original	69.493	70.732	73.540	69.989	62.42	3.196	.013*
MTLD ma bi	66.994	68.856	72.454	68.640	61.488	3.119	.015*
MTLD_ma_wrap	68.188	69.627	72.498	68.722	62.182	2.903	.022*

* p<0.05 ** p<0.01

Table 2 illustrates the mean values, F-values, and p-values of six indices of argumentative essays written by the 5 groups (i.e. Y1, Y2, Y3, Y4, and Native). As we can see, all of the p-values are less than 0.05, meaning that the difference among the five groups is of high significance, thus indicating that all of the lexical diversity indices examined are capable of distinguishing among the five groups. However, their abilities to differentiate adjacent groups are relatively weak. For clear comparison, the development trends of six lexical diversity indices were displayed in Figure 1 and Figure 2.

A cursory glance at Figure 1 exhibited that although there are extremely slight fluctuations, overall the changes of the three broken lines between each adjacent group are not obvious. Despite the MATTR50 line's fluctuation being slightly more obvious than that of the other indices, the three broken lines are almost close to three horizontal lines, with nearly no development trend.

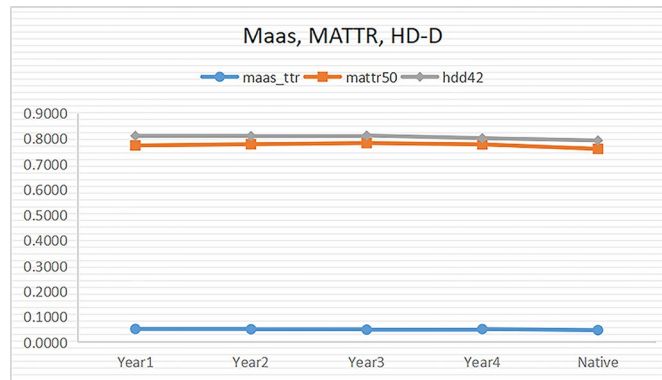


Figure 1: Development trends of Maas_ttr, MATTR, and HD-D.

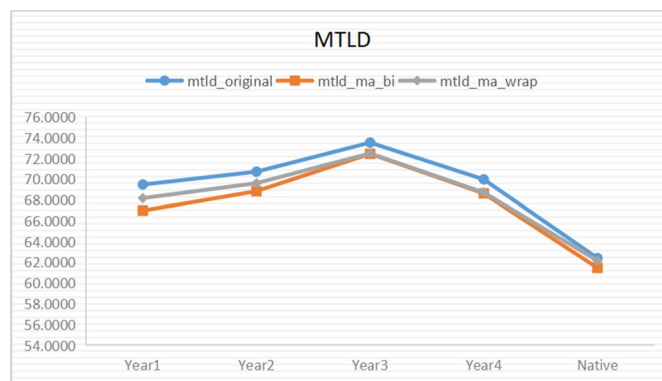


Figure 2: Development trends of MTLD_original, MTLD_ma_bi, and MTLD_ma_wrap.

Figure 2 showed three different variants of MTLD, but reported subtle differences. They shared a

similar non-linear development trend. The broken lines representing these three indices first went through a smooth increase, from freshman year to sophomore year, followed by a relatively sharper increase beginning with the sophomore year, reaching a peak at junior year, then presented a downward trend, first relatively gently declining to senior year, and finally pretty sharply to native speakers. Though the difference between two adjacent groups is relatively large, for example, the MTLTD scores for Y3 are higher than those of Y2, and those for Y2 are higher than Y1, there is no continuous upward linear pattern as the proficiency level increases. This phenomenon is obviously displayed in Figure 2 since it can be seen that while Y3 achieves higher scores than Y2, it also maintains higher scores than Y4 and the native group. On account of the relatively major fluctuation displayed in the development of the MTLTD indices and the lack of an overall linear growth pattern, the effectiveness of the three indices in distinguishing between adjacent English proficiency levels is also quite limited.

To summarize, the development of the six lexical diversity indices can be roughly categorized into two situations: first, for Maas_tr, MATTR50, and HD-D42, there was no apparent development trend. Second, for three MTLTD indices, a fluctuated developmental trend is shown, with an increase from Y1 to Y3, and a decrease from Y3 to the native group. In spite of the two existing situations, the six indices share a general quality, which is low effectiveness in distinguishing adjacent English proficiency levels on the basis of the major fluctuation and the weak performance on differentiating adjacent groups. This finding is corresponding to what Bao found in his 2008 research^[10]. Besides the whole data, the results gained by the Chinese undergraduate groups can provide some other useful insights. As shown in Figure 2, the lexical diversity firstly gently increases from freshman year to sophomore year, then rises more rapidly from sophomore year to junior year, reaching a peak, and then dramatically decreases from junior year to senior year. This phenomenon precisely fits the several hypotheses mentioned in Wen (2006)^[19]. The first one is the Study Gold Period Hypothesis, which suggests that the period between sophomore year and junior year is the best time for college students' learning. Since the teaching methods, contents, requirements, and living environment in universities are quite different from those in middle schools, students are required to adjust themselves from a variety of aspects, thus forming an adjustment period from freshman year to sophomore year, which may contribute to the reasons why the lexical diversity experience a smooth increase from Y1 to Y2. After the adjustment period, students have been more familiar with the teaching and living environment, and have also adapted to the university teaching methods. Therefore, their learning is on the fast lane and may gain greater achievement, which can explain the high-speed lexical diversity rising from Y2 to Y3. However, according to the second hypothesis, the Plateau Hypothesis, after significant progress in skill learning, there will be a certain stage of slowness, stagnancy, or even decline, which may account for the peak of lexical diversity in junior year and the decline in senior year. However, the regression that happened in the senior year may be also related to the Inadequate Skill training Hypothesis, which points out that the reason for the signs of deterioration in the lexical diversity of senior students may be that the English courses are fading away during this grade, while students are engaged in seeking jobs or preparing for postgraduate entrance exams, thus resulting in the decrease.

4.2. Lexical Sophistication

With the goal of investigating lexical sophistication indices, the texts collected from the corpora were input into P-Lex (<https://www.lognostics.co.uk>), S (2014, <https://www.kojima-vlab.org>), and AntWordProfiler. The next step was to generate the mean values of K1, K2, AWL, OFF, P-Lex, and S of each composition using SPSS 22.0, and their developmental plots using Microsoft Office Excel.

Table 3: Means of lexical sophistication measures: K1, K2, AWL, OFF (unit:%), P-Lex, and S.

Measure	Y1	Y2	Y3	Y4	Native
K1	88.27	86.94	85.67	84.90	77.99
K2	4.89	4.92	5.25	4.84	3.00
AWL	1.94	2.37	2.95	3.59	7.63
OFF	4.90	5.77	6.13	6.67	11.38
P-Lex	1.26	1.39	1.52	1.49	1.75
S	2665.33	3119.21	3451.83	3684.73	4571.36

Table 3 illustrates the mean values of K1, K2, AWL, OFF, P-Lex, and S of argumentative essays written by the five groups, Y1, Y2, Y3, Y4, and Native group. The following Figures 3 to 6 more vividly present the development trends of K1, K2, AWL, OFF, P-Lex, and S.

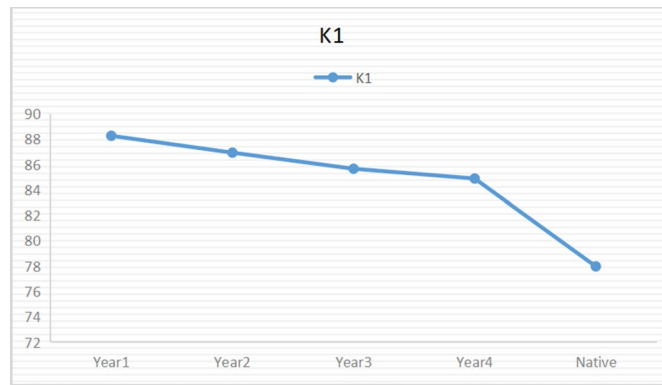


Figure 3: Development trend of K1.

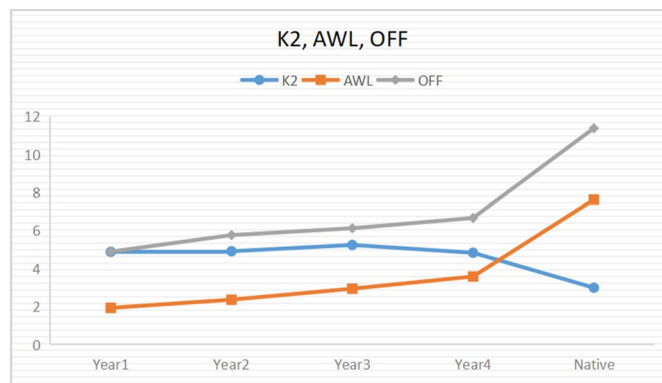


Figure 4: Development trends of K2, AWL, OFF.

A combination of Figures 3 and 4, as well as Table 3 exhibited several findings. Firstly, K1 (most frequently used 1000 words) continued to decline at different proficiency levels, among which the decrease during the period of freshman to senior year was rather gentle, forming an approximately linear pattern, and that from Year 4 to Native speakers was pretty sharp. This result implies a two-fold interpretation. On the one hand, it is a merit that Chinese students tend to utilize less simple K1 words as language proficiency improves, suggesting progress in lexical sophistication. On the other hand, there still remains a large difference between L2 learners and native speakers in terms of the usage of low-frequency words, since native speakers tend to use way fewer K1 words than Y4 students. As for K2 words (the following most frequently used 1000 words), which are also low-frequency words, there was no intensive change among Chinese English majors from four grades. The development of K2 also fluctuated, with no stable upward linear trend. But generally, native speakers also employ fewer K2 words than Chinese L2 learners. What is evident about AWL and OFF (more advanced words) is that although native speakers maintain a far outpaced percentage of those advanced words than Chinese students, there are clear and linear development tendencies for both AWL and OFF among Chinese students, for the percentage has constantly increased as language proficiency progressed. As a fortunate sign, it suggests that Chinese students' lexical sophistication has improved on the basis of the fact that skilled learners tend to apply more complex and sophisticated words in AWL and OFF. The overall results indicated that K1, AWL, and OFF are all valid indices to measure lexical sophistication on account of their linear developmental tendency.

As can be seen from Figure 5, the broken line of lambda was going up in general, but slightly fluctuated at the point of Y3, suggesting lambda's incapacity of generating a linear development trend. As for the index S, it continued to rise from Y1 to Native, with a gradually slow increasing rate in each Chinese adjacent group, and a rather rapid rate from Y4 to Native, suggesting that as Chinese students' language proficiency improved, more sophisticated words were chosen to write and that native speakers maintain more advanced words at their disposal than L2 learners.

Given that the trends of K2 and P-Lex were not linear, therefore being weak in distinguishing between adjacent English proficiency levels, the author excluded these two values in the next stage where the author employed SPSS 22.0 to generate the outcome of Factorial ANOVA Contrasts for K1, AWL, OFF, and S. SPSS 22.0 generated their significance value of the difference between every two adjacent groups, which respectively were Y2-Y1, Y3-Y2, Y4-Y3, and Native-Y4. The results of the Factorial ANOVA

Contrasts are shown in Table 4.

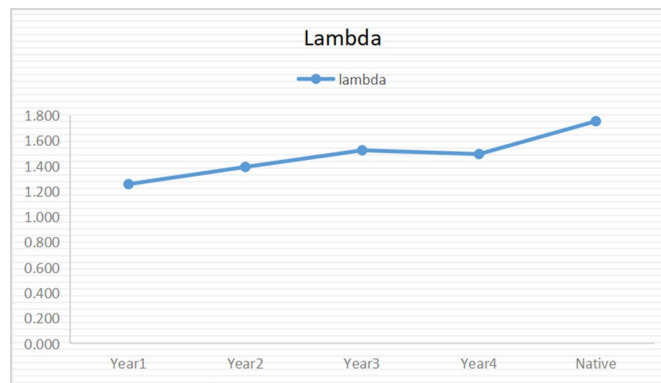


Figure 5: Development trend of Lambda.

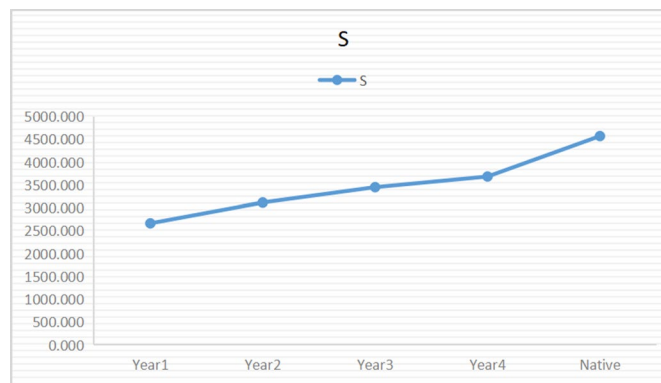


Figure 6: Development trend of S.

According to Table 4, in terms of K1, there were strongly significant differences found in the Y2-Y1, Y3-Y2, and Native-Y4 groups ($p = .003$, $p = .003$, and $p = .000$), but no significant difference existed in Y4-Y3. As for AWL, the significant differences in every two adjacent groups were all very strong ($p = .004$, $p = .001$, $p = .003$, $p = .000$), which showed its great effectiveness as a strong lexical sophistication index in discriminating different language proficiency. With regard to OFF, it was relatively weaker among the four indices since it was only capable of distinguishing between Chinese freshman and sophomore-year students, as well as between fourth-year Chinese students and British or American native speakers ($p = .013$, $p = .000$). As for the index S, the differences in the Y2-Y1, Y3-Y2, and Native-Y4 groups were strongly significant ($p = .000$, $p = .007$, and $p = .000$), which shares a similar pattern with K1.

Table 4: *p*-values gained from the Factorial ANOVA on lexical sophistication measures.

Measure	Y2-Y1	Y3-Y2	Y4-Y3	Native-Y4
K1	.003**	.003**	.102	.000**
AWL	.004**	.001**	.003**	.000**
OFF	.013*	.208	.174	.000**
S	.000**	.007**	.103	.000**
* $p < 0.05$ ** $p < 0.01$				

It can be summarized from the data that K1, AWL, OFF, and S were all able to discriminate among groups, and were all efficient lexical sophistication measures, while OFF is relatively weaker and AWL is the strongest index, being able to distinguish between any two adjacent groups. In addition, the results also prove that lexical sophistication is the most important differentiator in word use between different language proficiency groups, and its developmental trend almost remains linear, even when measured differently, which is in line with Bao's finding in his research in 2008^[10]. As for the specific lexical sophistication measures, the present results also confirmed some previously published papers. Taking the research of Laufer and Nation in 1995 as an example, they demonstrated that less proficient language users applied more K1 words than more proficient ones, and tended to use more K2 words, while for complex words (AWL and OFF), there was a significant difference among users with different language proficiency, which coincides with the present study^[11].

5. Conclusions

In conclusion, this study found that for development trends, the research has found that generally the development patterns of lexical diversity were non-linear, and exhibited two patterns. First, for Maas_ttr, MATTR50, and HD-D42, there was no apparent development trend. Second, for three MTLTD indices, a fluctuated developmental trend is shown, with an increase from Y1 to Y3, and a decrease from Y3 to the native group. In contrast, this study has identified the linear development of four lexical sophistication measures (K1, AWL, OFF, S), while the development of P-Lex and K2 was almost linear with slight fluctuation. With regard to indices efficiency, it was found that no lexical diversity indices employed in the current study were efficient, while lexical sophistication indices including K1, AWL, OFF, and S were effective measures, among which AWL was the strongest, followed by K1 and S, then OFF. This result indicates that lexical sophistication is a better indicator of distinguishing vocabulary differences between different levels than lexical diversity.

References

- [1] Astika G G. *Analytical assessments of foreign students' writing*[J]. *RELC Journal*, 1993, 24(1): 61-70.
- [2] Wang W, Wen Q. *Mother tongue thinking and second language writing, college students' English writing process study*[J]. *Journal of PLA Foreign Languages Institute*, 2002, 25(4): 5.
- [3] Kyle K, Crossley S. *The relationship between lexical sophistication and independent and source-based writing*[J]. *Journal of Second Language Writing*, 2016, 34: 12-24.
- [4] Cumming A. *Expertise in evaluating second language compositions*[J]. *Language Testing*, 1990, 7(1): 31-51.
- [5] Leki I, Carson J G. *Students' perceptions of EAP writing instruction and writing needs across the disciplines* [J]. *TESOL Quarterly*, 1994, 28(1).
- [6] Yule G U. *The statistical study of literary vocabulary*[J]. *Nature*, 1944, 5(27): 91-a.
- [7] Laufer B. *The development of L2 lexis in the expression of the advanced learner*[J]. *Modern Language Journal*, 1991, 75(4): 440-448.
- [8] Engber C A. *The relationship of lexical proficiency to the quality of ESL compositions*[J]. *Journal of Second Language Writing*, 1995.
- [9] Read J. *Assessing Vocabulary (Cambridge Language Assessment)*[M]. Cambridge: Cambridge University Press, 2000.
- [10] Bao G. *A multi-latitude study on the development of vocabulary richness in compositions of second language learners*[J]. *Foreign Language Audiovisual Education*, 2008(5): 7.
- [11] Laufer B, Nation P. *Vocabulary size and use: lexical richness in L2 written production*[J]. *Applied Linguistics*, 1995(3): 307-322.
- [12] Jarvis S. *Capturing the diversity in lexical diversity*[J]. *Language Learning*, 2013, 63(s1): 87-106.
- [13] Mass H D. *On the relationship between text length and word count*[J]. *Journal of Literature and Linguistics*, 1972, 2(8): 73-79.
- [14] Covington M A, McFall J D. *Cutting the gordian knot: the moving-average type-token ratio (MATTR)*[J]. *Journal of Quantitative Linguistics*, 2010, 17(2): 94-100.
- [15] McCarthy P M, Jarvis S. *Vocd: a theoretical and empirical evaluation*[J]. *Language Testing*, 2007, 24(4): 459-488.
- [16] McCarthy P M, Jarvis S. *MTLD, vocd-D, and HD-D: a validation study of sophisticated approaches to lexical diversity assessment*[J]. *Behavior Research Methods*, 2010, 42(2): 381-392.
- [17] Meara P, Bell H. *P_Lex: a simple and effective way of describing the lexical characteristics of short L2 texts*[J]. *Prospect*, 2001.
- [18] Kojima M, Yamashita J. *Reliability of lexical richness measures based on word lists in short second language productions*[J]. *System*, 2014, 42(1): 23-33.
- [19] Wen Q. *The trend and characteristics of English major students' oral vocabulary change*[J]. *Foreign Language Teaching and Research*, 2006, 38(3): 7.