# Dynamic Relationships Between Lexical Frequency Levels in L2 English Writing at Secondary School: A Learner Corpus Analysis


**Katarzyna Rokoszewska**

**Abstract** According to Complex Dynamic Systems Theory (CDST), language subsystems develop in an interdependent and non-linear way forming supportive, competitive, and pre-conditional relationships. One such language subsystem is lexis whose development is typically investigated in terms of lexical density, sophistication, and variation. The present paper focuses on the development of lexis in terms of lexical sophistication operationalized as four basic frequency levels. More specifically, the aim of the paper is to examine dynamic relationships between lexical frequency levels in L2 English writing at secondary school. The study was based on The Written English Developmental Corpus of Polish Learners (WEDCPL), which consists of 1924 essays. The texts were produced by 100 learners during 21 repeated measurements conducted over the period of 3 years. The results indicated that the learners predominantly relied on the use of words from the first frequency level to the disadvantage of words from higher frequency levels. The relationships between the frequency levels revealed some competition between the first and higher frequency levels and some pre-conditioning between the second and third frequency levels. Thus, developing learners' lexis beyond the first level is necessary to discourage the production of lexically unsophisticated texts and to foster the use of more advanced words.

**Keywords** Complex Dynamic Systems Theory (CDST) · Lexical sophistication · Lexical frequency · L2 writing · Learner corpus

K. Rokoszewska (✉)
Jan Dlugosz University in Czestochowa, Czestochowa, Poland
e-mail: k.rokoszewska@ujd.edu.pl



M. Baran-Łucarz et al. (eds.), *Contemporary Issues in Foreign Language Education*, English Language Education 32,
https://doi.org/10.1007/978-3-031-28655-1_11

## 1 Introduction

Complex Dynamic Systems Theory (CDST) is the term coined by de Bot (2017) to refer to both Complexity Theory (CT) (Larsen-Freeman & Cameron, 2008), which originated from the natural sciences, and Dynamic Systems Theory (DST) (Verspoor et al., 2011), which is a branch of mathematics, on the grounds that the two theories, applied to SLA in different academic centres, share similar theoretical and methodological principles. As an alternative approach to SLA, CDST focuses on language development, as opposed to acquisition, tracing changes in the dynamics of this process (Larsen-Freeman, 2011). According to CDST, language is a complex dynamic system which consists of internally complex subsystems which are said to co-develop non-linearly at different rates and form multiple relationships (Van Geert & Van Dijk, 2002). Language development is an emergent, variable, and self-organising process (de Bot et al., 2007; Larsen-Freeman, 1997).

CDST offers a new perspective on various constructs in SLA, including the CAF triad which consists of complexity, accuracy, and fluency. In this framework, the components of the CAF triad are viewed as multidimensional subsystems whose diachronic development, as opposed to synchronic manifestation, should be examined (Housen et al., 2012). Lexical complexity, which is a counterpart of syntactic complexity, consists of lexical density, variation, sophistication, and compositionality (Bulté & Housen, 2012). So far studies on lexical complexity have been conducted either on the basis of cross-sectional corpora of many subjects (Laufer & Nation, 1995; Lu, 2012; Verspoor et al., 2012) or longitudinal mini-corpora of single subjects (Spoelman & Verspoor, 2010; Caspi & Lowie, 2013; Chan et al., 2014). The present study intends to contribute to lexical research within the CDST framework by investigating the development of one of the components of lexical complexity, namely lexical sophistication operationalised in terms of word frequency. Thus, the present study will examine the dynamics of the relationships between different frequency levels in the development of lexical sophistication in L2 English writing at secondary school on the basis of the learner corpus which provides dense, cross-sectional, and longitudinal data.

## 2 Complex Dynamic Systems Theory

Complex Dynamic Systems Theory (CDST) represents non-symbolic psychological theories which contrast sharply with symbolic linguistic theories (Altman, 2006; Huljstin, 2002). Firstly, the non-symbolic theories construe knowledge in terms of connectionist as opposed to symbolic architectures postulating that knowledge is not represented as a set of symbols and rules which specify relationships between them but as different patterns of activation and connectivity which carry meaning in the neural network (Elman, 2001; MacWhinney, 2005). Secondly, the psychological theories question the existence of an innate human endowment posited by nativists in Chomskyan tradition, and reject inheritance and predeterminism as solutions to

the logical problem of language acquisition arguing that language emergence is driven by domain-general learning mechanisms which are applied to linguistic data and not by domain-specific learning mechanisms which evolved for language acquisition (Ellis, 2006; Saffran & Thiessen, 2007). Thirdly, non-symbolic theories posit that language emerges from use and experience through the process of grammaticalization (Hopper, 1998; Tomasello, 2015). It is learnt from input in a probabilistic way thanks to the ability of the human brain to unconsciously register and store all occurrences of language items in memory and to compute frequency statistics based on the availability, contingency, and reliability of linguistic cues to form and meaning (Ellis, 2002; MacWhinney, 2005).

In the CDST framework, language is construed as a complex, dynamic system. Language development is an emergent and highly variable process which takes place through soft-assembly and co-adaptation to changing demands of the communicative context (Larsen-Freeman, 2011). Like other dynamic systems, language is characterised with nonlinearity, unpredictability, sensitivity to initial conditions, openness, self-organisation, feedback sensitivity, and adaptability (de Bot et al., 2007; Larsen-Freeman, 1997). It consists of internally complex subsystems which develop simultaneously not only at different timescales and levels but presumably also at different rates (Van Geert & Van Dijk, 2002). The subsystems enter multiple interactions which may be of supportive, competitive, conditional, or dual character (Van Geert & Van Dijk, 2002). The trade-offs between and within these subsystems reflect language learners' limited cognitive and linguistic resources which they soft-assemble to co-adapt to other language users in communication (Verspoor et al., 2011).

Research informed by CDST may implement qualitative and quantitative methods, preferably in a mixed methods approach (Dörney, 2009). The former include qualitative comparative analysis, process tracing, concept mapping, agent-based modelling, retrodictive qualitative modelling, social network analysis, and design-based research methods, whereas the latter comprise panel designs, latent growth curve modelling, multilevel modelling, time series analysis, experience sampling method, single-case designs, and idiodynamic method (Hiver & Al-Hoorie, 2020). Studies conducted from the CDST perspective focus on the development of different language subsystems (Baba, 2020; Baba & Nita, 2014; de Bot et al., 2007), inter- and intra-individual variability (Larsen-Freeman, 2006; Verspoor et al., 2008) as well as the competition and coordination among CAF variables (Hou et al., 2020; Van Geert & Verspoor, 2015). The present study will focus on the development of one such subsystem, namely lexical sophistication.

## 3 Lexical Sophistication in CDST

Lexical complexity, which is construed as the breadth and depth of L2 lexical repertoire, is said to consist of lexical density, variation, sophistication, and compositionality which are quantified by various measures (Bulté & Housen, 2012). Lexical density, which indicates the amount of information in a text, is typically measured by the ratio of lexical tokens to all tokens (Ure, 1971). Lexical sophistication, which

denotes the depth of lexis, is expressed either as the proportion of advanced words to all words in a text (Read, 2000) or as the proportion of word types from different frequency levels in the Lexical Frequency Profile (LFP) (Laufer & Nation, 1995). Lexical variation or diversity, which refers to the range of vocabulary displayed in a text, is calculated by means of different corrected and randomised type-token ratios (Malvern et al., 2004). Lexical compositionality, which concerns formal and semantic components of lexical items, is calculated as the ratio of morphemes or syllables per words. In addition, lexical accuracy involves examining the quantity and quality of lexical errors (Read, 2000). According to James (1998), lexical errors are divided into (1) formal errors, which include mis-selections, misformations, and distortions, and (2) semantic errors, which involve problems with sense relations, collocations, connotations, and style.

Word frequency, which is the most common measure of lexical sophistication, provides information on how frequently a given lexical item is used with reference to such language corpora as the British National Corpus (BNC) (BNC Consortium, 2007) and the Corpus of Contemporary American English (COCA) (Davies, 2012). Word frequency lists which are commonly used in research on lexical sophistication are the General Service List (West, 1953), the CELEX lists (Baayen et al., 1995), BNC2000 (Nation, 2006), the COCA lists (Davies & Gardner, 2010), BNC/COCA2000 (Nation, 2016), the SUBTLEX lists (Brysbaert & New, 2009), and the New General Service List (Browne, 2014; Brezina & Gablasova, 2015). It is recommended to implement such lists in the development of the curriculum, teaching materials, and language tests.

As already mentioned, lexical sophistication is typically calculated as the ratio of sophisticated lexical words to all lexical words in a text (Linnarud, 1986; Hyltenstam, 1988). Sophisticated words are often defined as words which go beyond the first 2000 most frequently used words (Wolfe-Quintero et al., 1998; Lu, 2012). Alternatively, the Lexical Frequency Profile (LFP) (Laufer & Nation, 1995) shows the percentage of words from different frequency levels. However, Kyle et al. (2018) point out a number of different measures of lexical sophistication, such as word range, the use of academic language, psycholinguistic properties of words, such as concreteness, familiarity, meaningfulness, and imageability, word recognition norms, contextual distinctiveness, word neighborhood, and semantic network, pointing out the need to investigate multi-word units by means of n-gram indices. Lexical sophistication may be calculated by various computer programmes, such as CLAN (McWhinney, 2000), Range (Nation & Heatley, 2002), Lexical Complexity Analyzer (Lu, 2012), Lextutor (Cobb, 2014), AntWord Profiler (Anthony, 2022), Text Inspector (2018), P-lex (Meara & Bell, 2001), or the Tool for the Automatic Analysis of Lexical Sophistication (TAALES) (Kyle & Crossley, 2015).

Research on lexical sophistication indicated that word frequency affects word recognition and production (Balota et al., 2004), reading comprehension (Crossley et al., 2007; Nation, 2006), writing quality (Laufer & Nation, 1995; McNamara et al., 2015), and speaking proficiency (Kyle & Crossley, 2015). Studies on multi-word units highlighted the importance of knowing word combinations for language development (Bestgen & Granger, 2014; Kyle & Crossley, 2015). CDST studies

examined the development of lexical sophistication, density, and variation over time providing mixed results. Some studies reported an increase in the development of sophistication and variation as opposed to density (Duran et al., 2004; Storch & Tapper, 2009; Zheng, 2016). However, other studies did not find any statistically significant gains in these measures (Bulté & Housen, 2014; Knoch et al., 2014). In contrast to these studies, the present study investigated the development of one subsystem of lexical complexity, namely lexical sophistication, focusing on dynamic relationships between different frequency levels within this subsystem.

# 4 Method

## *4.1 Research Aim and Questions*

The aim of the present study was to examine the development of lexical sophistication in terms of supportive, competitive, pre-conditional, and dual relationships between different frequency levels in L2 English writing at secondary school. With respect to the aim of the study, the following research questions were formulated:

1. What results do learners obtain on the use of lexis from different frequency levels in L2 English writing over the whole learning period at secondary school?
2. What relationships take place between different frequency levels in L2 English writing at secondary school?

## *4.2 Research Method*

The present study took the form of panel design (Salkind, 2010; Hiver & Al-Hoorie, 2020) as it involved measuring the same variables on the same individuals over a longer period of time. The study was conducted at one of Polish secondary schools in 2014–2017. It was based on The Written English Developmental Corpus of Polish Learners (WEDCPL), which consists of over 1900 written texts collected during a series of repeated measurements over the period of 3 years (see Sect. 4.4). The corpus provides dense, cross-sectional, and longitudinal data. The study combined focused description, which was used to examine the relationships between different frequency levels, and CDST research procedures, which were implemented to investigate the dynamics of these relationships over time.

In the present study, lexical sophistication was operationalized as the proportion of words from different frequency levels in a written text. Hence, the research variables were four frequency levels, namely the first 1000 (level 1), the second 1000 (level 2), the third 1000 (level 3) most frequently used words in English, and the off-list words, i.e. words which go beyond these levels. The frequency levels were estimated on the basis of the Lexical Frequency Profile (LFP) (Laufer & Nation,

1995) with the use of the computer programme called Lextutor (Cobb, 2014). The profile was built on the basis of the BNC-COCA Core-4 frequency lists which come from the Common Core List (CCL) (Davies & Gardner, 2010) generated from the British National Corpus (BNC) (BNC Consortium, 2007) and the Corpus of Contemporary American English (COCA) (Davies, 2012).

### *4.3 Research Participants*

The research sample included 100 secondary school learners, i.e. 45 males and 55 females, who were at the age of 16–19 in grades 1–3. They were taught in seven language groups on the basis of the same coursebook by five different teachers with an MA degree in English. In the first grade, they were at the intermediate level (B1), in the second grade–at the upper-intermediate (B2) level, and in the third grade, they took a revision course (B2) for the final exam in English. By the time of the study, the learners had been learning English for about 9 years. At secondary school, they were enrolled in an extended English program and had 4–6 lessons per week depending on the grade. Around half of the learners had some extra-curricular English classes outside school. In the EFL context, the learners had some culture-mediated contact with English thanks to the Internet, music, and films. As for the third language, 86.0% of the learners studied German while 14.0% studied French. The learners' grade point average of all school subjects (GPA) and their average grade in English per all three grades were 4.0 on the 1–6 grading scale. Their results on the final written exam in English (B1–91.8%; B2–72.1%) were higher than the average national results (B1–73.0%; B2–63.0%).

### *4.4 Data Collection*

The Written English Developmental Corpus of Polish Learners (WEDCPL) was built on the basis of 21 tests administered over the period of 3 years (Table 1). The corpus consisted of 1924 per total of 2100 texts, the return rate being 91.6%. The size of the whole corpus was over 510,000 words while the size of the corpus analysed on the basis of the research samples was 393,202 words. The average length of the samples was 204 words.

The procedure of building the corpus involved the following steps: writing essays without reference materials during English lessons every month, scanning the original versions of the essays, grading the essays and storing them in regular files, preparing electronic transcripts by means of the speech recognition programme called Dragon Naturally Speaking (Nuance® 2014), verifying the transcripts with the learners' errors preserved, and preparing appropriate text samples. The learners were asked to write essays of the descriptive and argumentative mode on different topics based on the coursebook. The required length was 200 words and the time

**Table 1** Research design in a time series

| | Semester 1 | | | | | Semester 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Data | September | October | November | December | January | February | March | April | May | June |
| Grade 1 | Org. | Test 1 Fashion | Test 2 Internet | Test 3 Music | Test 4 Education | Winter break | Test 5 Ecology | Test 6 Pets | Test 7 Work | Test 8 Holidays |
| Grade 2 | Org. | Test 9 Books and Films | Test 10 Shopping | Test 11 Friendship | Test 12 Christmas | Winter break | Test 13 Family | Test 14 Health | Test 15 Fame | Test 16 Home and Living |
| Grade 3 | Org. | Test 17 Love | Test 18 TV | Test 19 Crime | Winter break | Test 20 Terrorism | Test 21 Tolerance | End of year | Matura exam | – |

limit was 45 min. The accuracy of the audio-transcripts was verified by an inter-rater who checked the transcripts of 100 texts. The inter-rater reliability, calculated in the form of Pearson's correlation coefficient, equalled 1.00.

## *4.5 Data Analysis*

The data were analysed with the use of the programme called Lextutor (Cobb, 2014). The texts were pre-processed to be decoded by the programme. Firstly, pre-processing involved a few formal procedures, such as changing short forms into long forms, replacing numbers with the word *one*, which belongs to the first frequency level, and recategorizing proper nouns as 1 k words (Cobb, 2014). Secondly, it entailed eliminating such items as L1 words, L3 words, direct borrowings, e.g. *circa, inter alia*, marginal words, e.g. *Ups, Oh, Yea*, and acronyms, e.g. *PE, IT*, but accepting informal words, e.g. *wanna*, *gonna,* and clipped forms, e.g. *oft, thru.* Thirdly, it involved dealing with errors in that words with minor spelling errors and morphosyntactic errors, i.e. word-bound grammatical errors connected with tense or plurality, were corrected, whereas words with major spelling errors and morphological errors, i.e. incorrect derivatives, were excluded from the analysis (James, 1998; Hemchua & Schmitt, 2006).

Data analysis involved some standard and CDST procedures (Verspoor et al., 2011). The data on different frequency levels were normalized to be compared and the general trends in the data sets were plotted by the second degree polynomials (see Sect. 5.1, Figs. 1–4). The learners' initial and final results on lexical frequency levels were compared by means of the Wilcoxon signed-rank test for related samples ($\alpha = 0.05$; $N = 100$). This non-parametric test was used as the Kolmogorov–Smirnov test of normality ($\alpha = 0.05$; $N = 100$) did not show the normal distribution in the compared data sets. Since the Wilcoxon test does not take into account the learners' results obtained between the first and the last test, the learners' results on all tests were correlated with time (Caspi & Lowie, 2013) by means of Pearson's linear correlation coefficient.

The co-development of the frequency levels was illustrated by the sixth degree polynomials (see Sect. 5.2, Fig. 5) which capture developmental variability. Linear and moving correlations were expressed as Pearson's correlation coefficients, but the latter were calculated by means of the moving window of correlations in which each measurement takes into account the previous one. Given 21 repeated measurements, a window of five data points was used, which yielded 17 correlation measurements (Verspoor et al., 2011). The correlations were calculated on the basis of detrended data so that the coefficients were not affected by increasing or decreasing trends. The significance of correlations was checked by means of a Monte Carlo method (10,000 iterations, $\alpha = 0.05$) (Spoelman & Verspoor, 2010) in which the null hypothesis of zero correlation is rejected provided that the null hypothesized value of $\rho$ ($\rho = 0$) falls outside the Monte Carlo confidence interval (Preacher & Selig, 2012).

# 5 Results

## *5.1 The Development of Lexical Sophistication*

The results of the present study indicated that the learners, on average, used 89.8% (SD = 2.56) of words from the first frequency level in written production at secondary school (Table 2, Fig. 1). On the first measurement point, they used 86.5% (SD = 2.80), whereas on the last measurement point–90.7% (SD = 2.57) of such lexical items. Furthermore, the group highest score was equal to 93.3% on test 17 (SD = 2.50), whereas the group lowest score was equal to 85.1% on test 20 (SD = 3.32). At the same time, the single highest score in the whole corpus was 99.0% (test 4), whereas the single lowest score was 75.4% (test 20). The general trend line in the development of words from this frequency level indicated a substantial increase in the first half of the observation period (tests 1–11), followed by an equally substantial decrease in the second half of this period (tests 11–21) (Fig. 1).

**Table 2** The development of lexical frequency–level 1

| Tests | T1 | T2 | T3 | T4 | T5 | T6 | T7 | T8 | T9 | T10 | T11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\bar{X}$ | 86.5 | 89.2 | 89.8 | 92.5 | 85.7 | 88.0 | 92.5 | 90.9 | 90.9 | 92.0 | 91.8 |
| Min | 77.0 | 83.6 | 78.7 | 87.4 | 77.7 | 80.3 | 86.9 | 83.1 | 82.9 | 84.2 | 84.7 |
| Max | 92.5 | 93.7 | 97.2 | 99.0 | 94.0 | 95.5 | 97.7 | 96.6 | 96.1 | 97.5 | 96.5 |
| SD | 2.80 | 2.23 | 3.00 | 2.37 | 3.22 | 3.08 | 2.28 | 3.00 | 2.86 | 2.74 | 2.26 |
| Tests | T12 | T13 | T14 | T15 | T16 | T17 | T18 | T19 | T20 | T21 | Total |
| $\bar{X}$ | 90.2 | 92.9 | 87.0 | 87.4 | 91.3 | 93.3 | 91.9 | 86.8 | 85.1 | 90.7 | 89.8 |
| Min | 83.8 | 87.0 | 75.6 | 81.1 | 85.7 | 83.1 | 86.4 | 79.9 | 75.4 | 84.5 | 85.1 |
| Max | 96.8 | 98.2 | 93.5 | 93.5 | 96.6 | 98.0 | 97.4 | 94.9 | 92.5 | 96.4 | 93.3 |
| SD | 2.70 | 2.61 | 3.39 | 2.77 | 2.58 | 2.50 | 2.36 | 2.80 | 3.32 | 2.57 | 2.56 |

Notes: $\bar{X}$ mean, *Min* minimum score, *Max* maximum score, *SD* standard deviation, *T* test

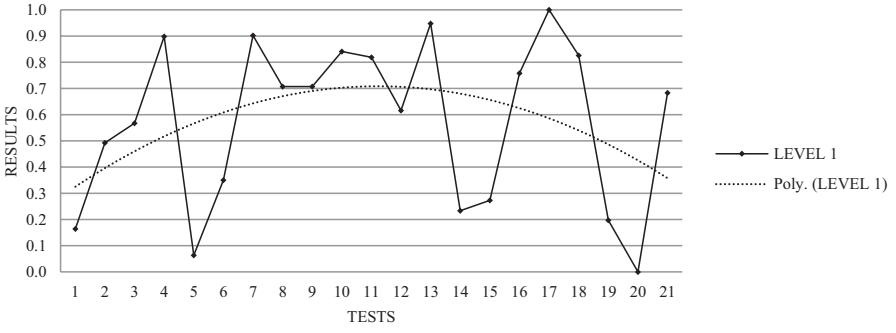


**Fig. 1** The development of lexical frequency–level 1. (Notes: Poly–polynomial trend line of the second degree)

**Table 3** The development of lexical frequency–level 2

| Tests | T1 | T2 | T3 | T4 | T5 | T6 | T7 | T8 | T9 | T10 | T11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\bar{X}$ | 5.6 | 3.0 | 3.3 | 3.3 | 6.3 | 3.0 | 3.9 | 5.7 | 3.1 | 3.0 | 3.6 |
| Min | 2.0 | 1.0 | 0.4 | 1.0 | 0.5 | 0.5 | 0.5 | 1.5 | 0.9 | 0.6 | 1.0 |
| Max | 12.1 | 7.7 | 8.0 | 6.6 | 10.9 | 6.5 | 8.9 | 14.4 | 8.3 | 8.4 | 7.0 |
| SD | 1.69 | 1.37 | 1.69 | 1.33 | 1.97 | 1.34 | 1.62 | 2.53 | 1.67 | 1.63 | 1.27 |
| Tests | T12 | T13 | T14 | T15 | T16 | T17 | T18 | T19 | T20 | T21 | Total |
| $\bar{X}$ | 4.5 | 2.6 | 5.3 | 6.8 | 4.5 | 4.5 | 3.3 | 4.2 | 8.2 | 6.0 | 4.3 |
| Min | 1.4 | 0.5 | 0.5 | 2.8 | 1.9 | 0.9 | 0.8 | 3.7 | 2.6 | 1.4 | 2.6 |
| Max | 9.2 | 8.5 | 11.4 | 12.3 | 9.6 | 11.4 | 10.6 | 13.7 | 11.4 | 9.5 | 8.2 |
| SD | 1.50 | 1.44 | 2.08 | 1.89 | 1.67 | 1.64 | 1.84 | 2.08 | 1.87 | 1.85 | 1.51 |

Notes: $\bar{X}$ mean, *Min* minimum score, *Max* maximum score, *SD* standard deviation, *T* test

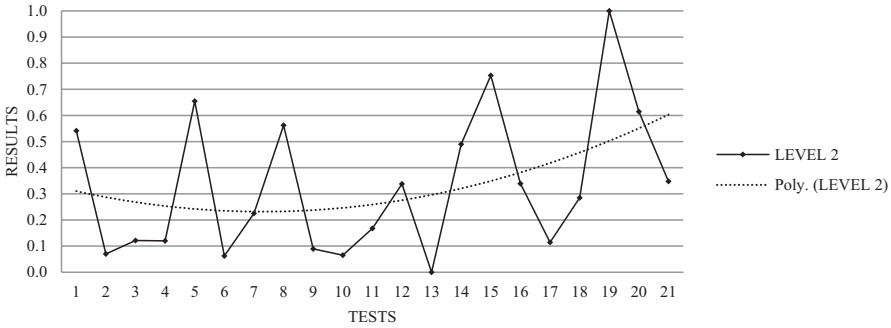


**Fig. 2** The development of lexical frequency–level 2. (Notes: Poly–polynomial trend line of the second degree)

With respect to the second frequency level (Table 3, Fig. 2), the results showed that the learners, on average, used 4.3% of such words (SD =1.51). On the first test, they produced 5.6% (SD = 1.69), whereas on the last test–6.0% (SD = 1.85) of such items. The group highest result equalled 8.2% (SD = 1.87) on test 20, whereas the group lowest result–2.6% (SD = 1.44) on test 13. In comparison, the single highest score in the whole corpus was 13.7% (test 19) while the single lowest score–0.4% (test 3). The general trend line in the development of words from this level illustrated a slight decrease at the beginning of the observation period (tests 1–7) and a rather substantial increase in the remaining part of this period (tests 7–21) (Fig. 2).

With respect to the third frequency level (Table 4, Fig. 3), the results revealed that the learners, on average, used 2.0% of such words (SD = 1.28). On the first test, they used 2.5% (SD = 1.08), whereas on the last test–1.7% (SD = 1.23) of such items. The group highest score amounted to 6.1% (SD = 1.75) on test 20, whereas the group lowest score to–0.9% (SD = 0.68) on test 8. In contrast, the highest individual result in the whole corpus was 12.2% (test 20) while the lowest individual result–0.4% (tests 4, 8, 10, 16–18). The general trend in the development of lexis from this level showed a moderate decrease in the first half of the learning period, followed by a similar increase in the second part of this period (Fig. 3).

**Table 4** The development of lexical frequency–level 3

| Tests | T1 | T2 | T3 | T4 | T5 | T6 | T7 | T8 | T9 | T10 | T11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\bar{X}$ | 2.5 | 4.0 | 2.0 | 1.2 | 2.7 | 3.7 | 1.6 | 0.9 | 2.3 | 1.2 | 1.3 |
| Min | 0.5 | 1.3 | 0.5 | 0.4 | 0.5 | 0.9 | 0.5 | 0.4 | 0.5 | 0.4 | 0.5 |
| Max | 5.3 | 7.3 | 6.4 | 4.4 | 6.2 | 8.3 | 4.8 | 3.7 | 6.1 | 3.4 | 4.9 |
| SD | 1.08 | 1.15 | 1.20 | 0.95 | 1.18 | 1.55 | 0.96 | 0.68 | 1.24 | 0.74 | 0.86 |
| Tests | T12 | T13 | T14 | T15 | T16 | T17 | T18 | T19 | T20 | T21 | Total |
| $\bar{X}$ | 1.3 | 1.5 | 3.8 | 2.7 | 1.2 | 1.3 | 1.4 | 2.5 | 6.1 | 1.7 | 2.0 |
| Min | 0.5 | 0.5 | 1.5 | 0.5 | 0.4 | 0.4 | 0.4 | 1.0 | 2.4 | 0.5 | 0.9 |
| Max | 4.6 | 4.9 | 8.3 | 7.1 | 4.0 | 6.4 | 3.9 | 6.6 | 12.2 | 6.5 | 6.1 |
| SD | 0.85 | 1.12 | 1.49 | 1.55 | 0.70 | 1.12 | 0.85 | 1.16 | 1.75 | 1.23 | 1.28 |

Notes: $\bar{X}$ mean, *Min* minimum score, *Max* maximum score, *SD* standard deviation, *T* test

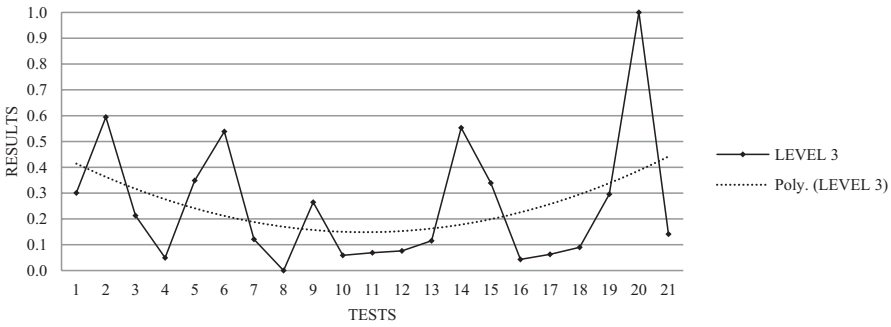


**Fig. 3** The development of lexical frequency–level 3. (Notes: Poly–polynomial trend line of the second degree)

Finally, with respect to the off-list words (Table 5, Fig. 4), the results showed that the learners, on average, used 3.2% of such words (SD = 1.21). On the first test, they produced 6.7% (SD = 2.85), whereas on the last test–3.3% of such items (SD = 2.03). Their highest group result was 6.7% on test 1 (SD = 2.85), whereas their lowest result was 1.7% (SD = 1.52) on test 8. In comparison, the best individual result in the whole corpus was 18.9% (test 1) while the poorest individual result was 0.4% (test 12). The general trend in the development of words from this level was decreasing throughout the whole observation period (Fig. 4).

As far as the learners' progress is concerned, the Wilcoxon signed-rank test for related samples ($\alpha = 0.05$; $N = 100$) indicated statistically significant differences between the learners' initial and final results in the development of the third frequency level ($p = 0.00$*) as opposed to the first ($p = 0.11$) and second ($p = 0.27$) frequency levels as well as words off the list ($p = 0.09$) (Table 6). In contrast to the Wilcoxon test, which takes into account only the first and the last test, the correlations between the learners' results on all tests and time did not show statistically significant relationships in the case of level 1 ($r = 0.03$), level 2 ($r = 0.33$), and level 3 ($r = 0.03$), but it showed a moderate negative relationship in the case of the words off the lists ($r = -0.53$*) (Table 6). The critical value ($r$*) for the significance of the correlation was 0.43 ($N = 21$; $\alpha = 0.05$).

**Table 5** The development of lexical frequency–off-list words

| Tests | T1 | T2 | T3 | T4 | T5 | T6 | T7 | T8 | T9 | T10 | T11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\bar{X}$ | 6.7 | 3.8 | 3.8 | 3.0 | 5.8 | 4.6 | 2.2 | 1.7 | 2.5 | 4.2 | 3.4 |
| Min | 2.0 | 0.8 | 0.5 | 0.5 | 1.0 | 0.5 | 0.5 | 0.5 | 0.5 | 0.8 | 0.8 |
| Max | 18.9 | 10.0 | 9.3 | 10.2 | 13.3 | 11.7 | 10.7 | 6.5 | 8.3 | 9.0 | 14.8 |
| SD | 2.85 | 1.93 | 1.91 | 2.01 | 2.61 | 2.25 | 1.75 | 1.52 | 1.69 | 1.93 | 2.52 |
| Tests | T12 | T13 | T14 | T15 | T16 | T17 | T18 | T19 | T20 | T21 | Total |
| $\bar{X}$ | 3.0 | 2.9 | 3.6 | 3.0 | 3.6 | 2.2 | 2.5 | 2.8 | 2.1 | 3.3 | 3.2 |
| Min | 0.4 | 0.5 | 0.8 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 1.7 |
| Max | 9.5 | 8.8 | 11.8 | 10.2 | 9.1 | 8.2 | 7.1 | 11.4 | 6.5 | 9.2 | 6.7 |
| SD | 1.98 | 1.91 | 2.04 | 1.64 | 1.84 | 1.50 | 1.45 | 1.86 | 1.31 | 2.03 | 1.21 |

Notes: $\bar{X}$ mean, *Min* minimum score, *Max* maximum score, *SD* standard deviation, *T* test

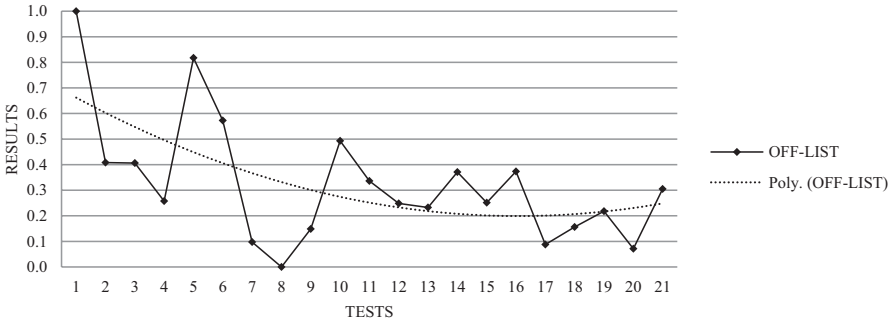


**Fig. 4** The development of lexical frequency–off-list words. (Notes: Poly–polynomial trend line of the second degree)

**Table 6** The learners' progress in lexical frequency over time

| | Level 1 | Level 2 | Level 3 | Other |
|---|---|---|---|---|
| Wilcoxon test (*p*-value) | 0.11 | 0.27 | 0.00* | 0.09 |
| Correlation with time (*r*) | 0.03 | 0.33 | 0.03 | −0.53* |

Notes: An asterisk–statistically significant results ($\alpha$ = 0.05), the Wilcoxon test ($N$ = 100) and Pearson product ($N$ = 21)

## *5.2 The Relationships Between Lexical Frequency Levels*

The correlations between different frequency levels (Table 7) showed that the relationships between the first frequency level and higher frequency levels were negative and statistically significant. This was reflected in the co-development of lexical frequency levels over the whole observation period (Fig. 5) which revealed that an increase in the development of the first frequency level, observed in the middle of this period, took place during a decrease in the development of higher frequency levels. This was in turn confirmed by the moving correlations. Detailed data on the moving correlations are presented in Table 1 in the Appendix.

**Table 7** Correlations between different frequency levels

| Data | *r* | $R^2$ | MC CI LL | MC CI UL | Type |
|---|---|---|---|---|---|
| Level 1 & 2 | −0.76* | 57.76 | −0.90 | −0.49 | Competitive |
| Level 1 & 3 | −0.77* | 59.29 | −0.90 | −0.51 | Competitive |
| Level 1 & Off-list | −0.51* | 26.01 | −0.77 | −0.10 | Competitive |
| Level 2 & 3 | 0.28 | 7.84 | −0.19 | 0.63 | Pre-conditional |
| Level 2 & Off-list | 0.29 | 8.41 | −0.16 | 0.64 | Dual |
| Level 3 & Off-list | 0.19 | 3.61 | −0.26 | 0.57 | Dual |

Notes: *r* correlation coefficient, $R^2$ shared variance, *MC CI LL* Monte Carlo confidence interval lower level, *MC CI UL* Monte Carlo confidence interval upper level, an asterisk–statistically significant results, a Monte Carlo analysis ($\alpha = 0.05$, $N = 21$)

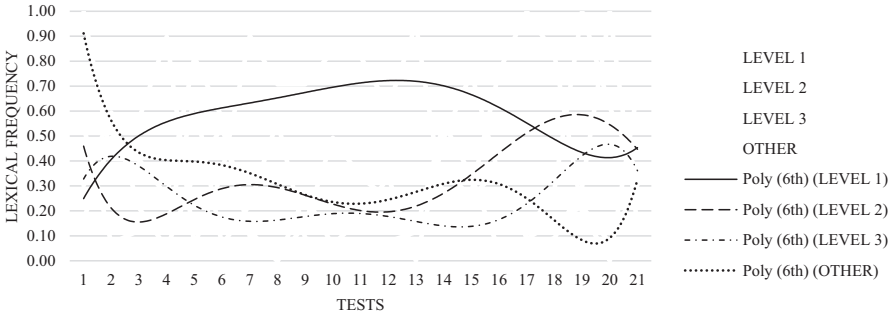


**Fig. 5** The co-development of lexical frequency levels. (Notes: Poly (6th)–polynomial trend line of the sixth degree)

More specifically, the relationships between the first and the second level ($r = -0.76$*) as well as the first and the third level ($r = -0.77$*) were rather strong while the relationship between the first and the off-list level ($r = -0.51$*) was moderate (Table 7). The moving correlations also indicated that words from the first frequency level co-developed with words from higher frequency levels in predominantly competitive relationships during the majority of the observation period (Fig. 6). However, the competition between these variables fluctuated to some extent. In the case of the first two relationships, this competition became weak at data point 8 and at data points 9 to 11, respectively. In the case of the third relationship, it changed into some support between data points 9 and 11. Nevertheless, it may be concluded that the first frequency level and higher frequency levels developed as the so called competitive growers in L2 English writing.

The correlations between higher frequency levels were statistically insignificant (Table 7). However, in contrast to the standard correlation between words from the second and third frequency levels ($r = 0.28$), the moving correlation revealed that the relationship between the two variables was pre-conditional in that they competed with each other in the first part of the observation period (data points 6–11) but supported each other in the second part of this period (data points 12–19) (Fig. 7). Thus, the second and third frequency levels co-developed as pre-conditional growers.

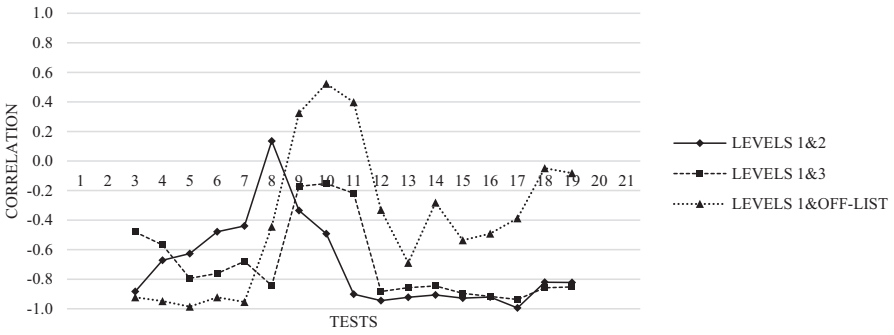


**Fig. 6** The moving correlations between the first and higher frequency levels

The moving correlation between the second frequency level and the off-list level ($r = 0.29$) as well as the moving correlation between the third frequency level and the off-list level ($r = 0.19$) (Table 7) illustrated dual relationships between the variables. In the first case, high and moderate support at the beginning of the observation period (data points 3–7) changed into quite strong competition in the middle of this period (data points 8–10), which in turn changed into moderate support at the end of this period (data points 12–19) (Fig. 7). In the second case, a strong supportive relationship observed at the beginning (data points 3–8) plummeted into a strong negative relationship (data point 11), but then it rose to a temporary strong support (data point 13) only to drop down at the end (data points 17–19) (Fig. 7). Thus, words from the second and third frequency levels developed in relation to the off-list words as dual growers.

Synthesising the results, it is important to point out that support in the relationships between the first frequency level and higher frequency levels (Fig. 6) took place when competition occurred in the relationships between the higher frequency levels (Fig. 7), i.e. around data points 8–12. In addition, Fig. 8 illustrates competitive relationships between the first frequency and higher frequency levels in contrast to the pre-conditional relationship between higher frequency levels, namely levels 2 and 3.

To conclude, the linear regression models revealed that the increase of words from level 1 by 1% would cause a decrease of words from level 2 by 0.45%. In this model, the amount of variance in level 2 was explained by level 1 in 60.8% (Fig. 9). Furthermore, the increase of words from level 1 by 1% would lead to a decrease of words from level 2 by 0.19%, with shared variance between the variables equal to 21.2% (Fig. 10). Such an increase in words from level 1 would result in the decrease of the off-list words by 0.38%, with shared variance of 34.66% (Fig. 11). In contrast, the increase of words from level 2 by 1% would cause an increase of words from level 3 by 0.24%, shared variance being 4.15% (Fig. 12).

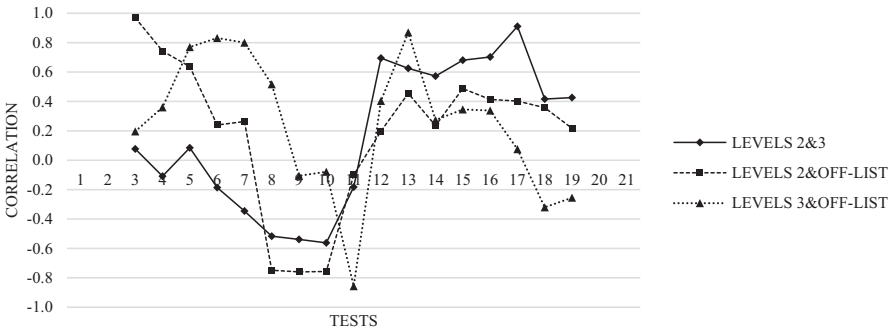


**Fig. 7** The dynamics of correlations between higher frequency levels

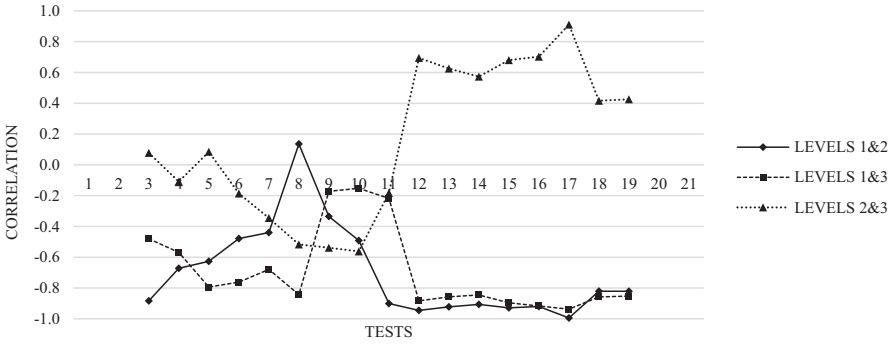


**Fig. 8** The moving correlations between the main frequency levels

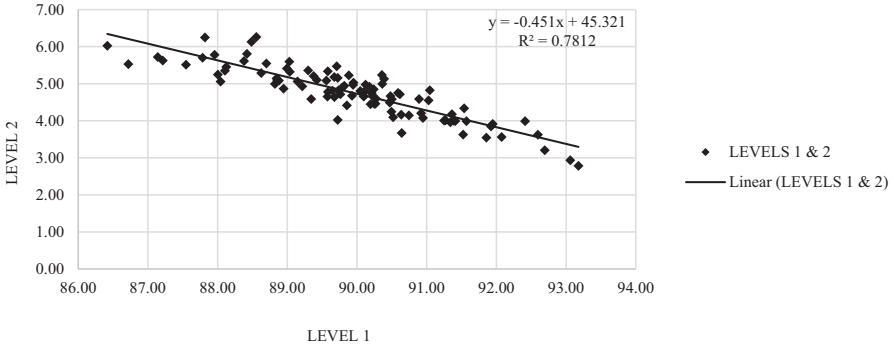


**Fig. 9** The correlation between levels 1 & 2–the linear regression model (*N = 100*)

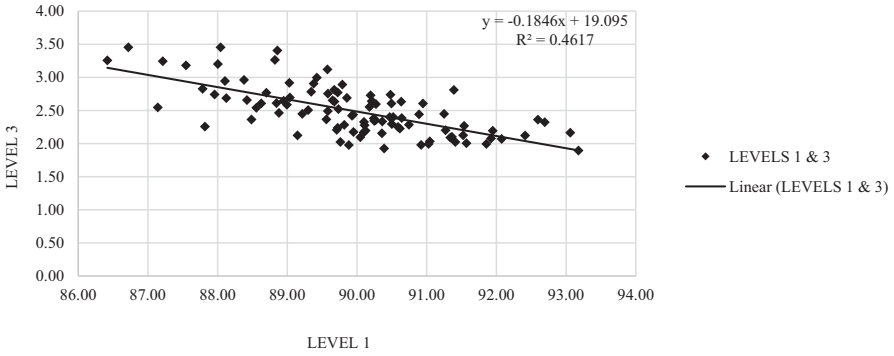


**Fig. 10** The correlation between levels 1 & 3–the linear regression model (*N = 100*)

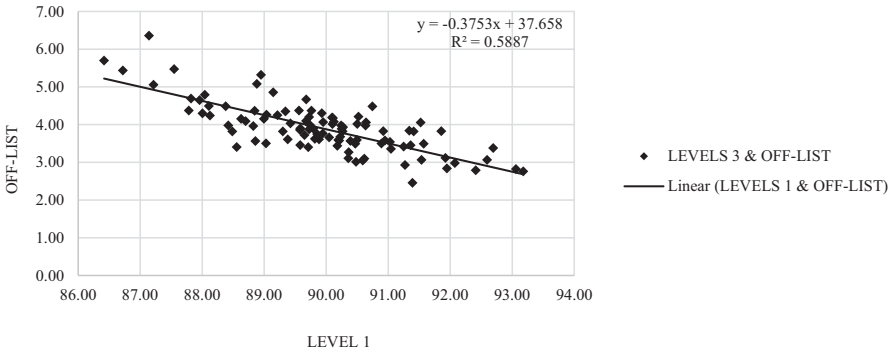


**Fig. 11** The correlation between levels 1 & off-list–the linear regression model (*N = 100*)

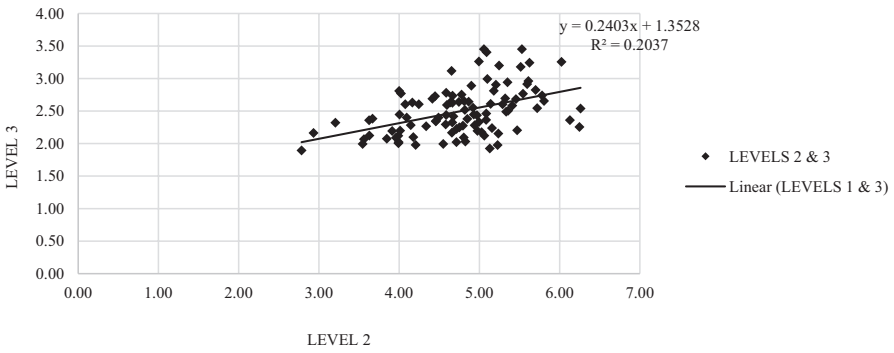


**Fig. 12** The correlation between levels 2 & 3–the linear regression model (*N = 100*)

## 6 Discussion

The aim of the present paper was to examine the development of lexical sophistication and the dynamic relationships between different frequency levels in L2 English writing at secondary school. With respect to the first research question, which concerned the learners' results on the development of lexical sophistication, it was established that the learners mainly used words from the first frequency level to the disadvantage of words from higher frequency levels. As for the learners' progress, it was found that the differences between the learners' initial and final results at secondary school were insignificant in the case of all frequency levels, except the third one. However, the correlation between the learners' results and time was insignificant not only for the first and second but also for the third level. The only exception was the negative relationship between the off-list words and time. Nevertheless, the general trend lines indicated some qualitative increase in the development of words from the second frequency level accompanied by some decrease in the development of words from the first frequency level.

In total, these results indicate that the learners made hardly any progress in lexical sophistication in L2 English writing over the period of 3 years at secondary school. Such results may be caused by a few factors. As already pointed out, some studies conducted within the CDST framework did not find significant progress in lexical sophistication, either (Bulté & Housen, 2014; Knoch et al., 2014). As Meara and Bell (2001) point out, L2 learners may indeed find it difficult to use words beyond the first 1000 productively. The results may also reflect the nature of lexical development. Caspi and Lowie (2013) found that different levels of lexical knowledge were hierarchically ordered in that vocabulary reception pre-conditioned its production while controlled and free production competed with each other. It is reasonable to assume that the learners in the present study must have recognised and practised words from higher frequency levels introduced at the intermediate and upper-intermediate levels in certified coursebooks, but they were unable to use them in free written production. This might also have been due to the type of practice the learners received in the EFL context. The quantity and quality of language practice might have been insufficient to foster the learners' use of more sophisticated lexis, especially if such practice consisted of controlled lexical exercises and entailed little naturalistic practice.

With respect to the second research question, which focused on the relationships between different frequency levels, it was established that the relationships between the use of words from the first and higher frequency levels were rather strong and negative. Indeed, an increase in the use of words from the most basic level caused a decrease in the use of words from more advanced levels. Thus, the first and higher frequency levels co-developed as competitive growers (Van Geert & Van Dijk, 2002). This means that due to limited cognitive resources (Verspoor et al., 2011), the learners were able to focus mainly on the first frequency level at the cost of higher levels. In contrast, the relationships among the higher levels were statistically insignificant. However, the moving correlation between the second and third

frequency levels revealed a clear pre-conditional relationship in which the second level had to be developed prior to the third one. Indeed, an increase in the second frequency level caused some increase in the third frequency level. In other words, the development of words from the second frequency level supported the development of words from the third frequency level. Thus, these two frequency levels functioned as pre-conditional growers (Van Geert & Van Dijk, 2002). In addition, it was observed that a change in one correlation may affect other correlations (Caspi & Lowie, 2013). More specifically, support in the relationships between the first and higher frequency levels was accompanied by competition in the relationships between higher frequency levels. Thus, the study showed that not only individual language variables, like lexical frequency levels, form various dynamic relationships but also the correlations between them contribute to an intricate network of mutual interdependencies in language development. On the whole, the study showed that lexical sophistication is a complex system which consists of different frequency levels which co-develop in dynamic, nonlinear, and inter-dependent ways (Housen et al., 2012; Verspoor et al., 2011).

The present study has some limitations. Firstly, despite the fact that the study was based on the iterative procedure which involved the same type of tasks performed in the same conditions in the language classroom, the topics were different, which was unavoidable in a longitudinal study in the real-life context but might have influenced the learners' results. Secondly, although the learners followed the same coursebook and received the same amount of instruction, they were taught by five teachers whose potentially different teaching styles might have affected the learners' lexical development. Thirdly, the study provided the panel data which yielded some insight into the group language behaviour, but in line with the ergodicity principle (Lowie & Verspoor, 2019), the individual learners' performance may diverge from the group average results. Hence, it is recommended to conduct a critical stance case study which would provide the data on selected learners in comparison to the panel study.

Despite the limitations, the study offers some practical implications for Polish teachers of English. The study revealed that the development of lexical sophistication posed both the teachers and learners with a substantial challenge, and pointed to the necessity to work on this aspect of language more efficiently. The teachers should realize that apart from accuracy and fluency, language development involves both syntactic and lexical complexity. Lexical sophistication, which is one of the main aspects of lexical complexity, should not be neglected. The teachers should focus on lexical sophistication not only in terms of controlled vocabulary practice but also free language production as words presented and practised in mechanical written exercises can hardly be accessed in speech or writing. Being familiarized with word frequency lists, especially with words from the first frequency level, the teachers would be able to monitor learners' level of lexical sophistication and to encourage them to use more sophisticated lexis. It would be also helpful to make learners aware of the lexis they use so that they would pay attention to whether they actually use words encountered at the intermediate and upper-intermediate levels or if they conveniently rely on basic lexis. Accepting the learners' predominant

reliance on the first frequency level, the teachers most unintentionally impede the learners' use of more sophisticated lexis. Creating appropriate affordances for the learners to use words from the second frequency level would support the development of words from further levels.

## 7 Conclusions

The present study has provided some insight into the development of lexical sophistication and the dynamics of the relationships between different frequency levels in L2 English writing at secondary school. The results of the study indicated some qualitative changes in the development of lexical sophistication in that the general trend lines illustrated some decrease in the use of words from the first frequency level to the advantage of some increase in the use of words from the second and third frequency levels in learner writing. The results also showed that relying on basic words from the first frequency level prevents the learners from using more sophisticated words from higher frequency levels. Conversely, developing the learners' lexis from the second frequency level serves as a necessary condition for their use of lexis from higher frequency levels. Since the study revealed that the learners' gains were statistically insignificant over the whole learning period, there arises the need to work on this aspect of lexical complexity more efficiently in order to achieve greater complexification and automatization of lexical knowledge in language production.

## Appendix

See (Table 8).

**Table 8** The moving correlations between different frequency levels

| Tests | Levels 1 & 2 | Levels 1 & 3 | Levels 1&Off-list | Levels 2 & 3 | Levels 2 & Off-list | Levels 3 & Off-list |
|---|---|---|---|---|---|---|
| 1 | – | – | – | – | – | – |
| 2 | – | – | – | – | – | – |
| 3 | −0.88 | −0.48 | −0.92 | 0.08 | 0.97 | 0.20 |
| 4 | −0.67 | −0.57 | −0.95 | −0.11 | 0.74 | 0.36 |
| 5 | −0.63 | −0.79 | −0.99 | 0.08 | 0.64 | 0.77 |
| 6 | −0.48 | −0.76 | −0.92 | −0.19 | 0.24 | 0.83 |
| 7 | −0.44 | −0.68 | −0.95 | −0.35 | 0.26 | 0.80 |
| 8 | 0.14 | −0.84 | −0.44 | −0.52 | −0.75 | 0.52 |
| 9 | −0.34 | −0.17 | 0.33 | −0.54 | −0.76 | −0.11 |
| 10 | −0.49 | −0.15 | 0.52 | −0.56 | −0.76 | −0.08 |

(continued)

**Table 8** (continued)

| Tests | Levels 1 & 2 | Levels 1 & 3 | Levels 1&Off-list | Levels 2 & 3 | Levels 2 & Off-list | Levels 3 & Off-list |
|---|---|---|---|---|---|---|
| 11 | −0.90 | −0.22 | 0.40 | −0.18 | −0.10 | −0.86 |
| 12 | −0.95 | −0.88 | −0.33 | 0.69 | 0.20 | 0.40 |
| 13 | −0.92 | −0.86 | −0.69 | 0.63 | 0.45 | 0.87 |
| 14 | −0.91 | −0.85 | −0.28 | 0.57 | 0.24 | 0.27 |
| 15 | −0.93 | −0.90 | −0.54 | 0.68 | 0.49 | 0.35 |
| 16 | −0.92 | −0.92 | −0.49 | 0.70 | 0.41 | 0.34 |
| 17 | −0.99 | −0.94 | −0.39 | 0.91 | 0.40 | 0.08 |
| 18 | −0.82 | −0.86 | −0.05 | 0.42 | 0.36 | −0.32 |
| 19 | −0.82 | −0.85 | −0.08 | 0.43 | 0.22 | −0.25 |
| 20 | – | – | – | – | – | – |
| 21 | – | – | – | – | – | – |

## References


Altman, G. (2006). History of psycholinguistics. In K. Brown (Ed.), *The encyclopaedia of language and linguistics* (2nd ed., pp. 257–265). Elsevier.

Anthony, L. (2022). AntWordProfiler, Version 2.0.0. [Computer software]. Waseda University.

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1995). *The CELEX lexical database*. Linguistic Data Consortium.

Baba, K. (2020). Exploring dynamic developmental trajectories of writing fluency: Who benefited from the writing task? In G. Fogal & M. Verspoor (Eds.), *Complex dynamic systems theory and L2 writing development* (pp. 3–26). https://doi.org/10.1075/lllt.54.01bab

Baba, K., & Nita, R. (2014). Phase transitions in the development of writing fluency from a complex dynamic systems perspective. *Language Learning, 64*(1), 1–35. https://doi.org/10.1111/lang.12033

Balota, D., Cortese, M., Sergent-Marshall, S., Spieler, D., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology, 133*, 283–316. https://doi.org/10.1037/0096-3445.133.2.283

Bestgen, Y., & Granger, S. (2014). Quantifying the development of phraseological competence in L2 English writing: An automated approach. *Journal of Second Language Writing, 26*, 28–41. https://doi.org/10.1016/j.jslw.2014.09.004

BNC Consortium. (2007). *The British National Corpus, XML Edition.* Oxford Text Archive.

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the New general service list. *Applied Linguistics, 36*(1), 1–22. https://doi.org/10.1093/applin/amt018

Browne, C. (2014). A New general service list: The better mousetrap we've been looking for? *Vocabulary Learning and Instruction, 3*(1), 1–10. https://doi.org/10.7820/vli.v03.1.browne

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*, 977–990. https://doi.org/10.3758/BRM.41.4.977

Bulté, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 23–46). John Benjamins.

Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing, 26*(4), 42–65. https://doi.org/10.1016/j.jslw.2014.09.005

Caspi, T., & Lowie, W. (2013). The dynamics of L2 vocabulary development: A case study of receptive and productive knowledge. *Revista Brasileira de Linguística Aplicada, 13*(2), 437–462.

Chan, H. P., Lowie, W., & de Bot, K. (2014). A case study of lexical development of writing and speaking in identical twins. In J. Gao, M. Guerin, V. Allegranzi, M. Riviere, L. Sauwala, C. Simon, et al. (Eds.), *Actes des 16'emes Rencontres Jeunes Chercheurs en Sciences du Langage: Mod'eles et mod'elisation dans les sciences du langage* (pp. 54–65). HAL-SHS.

Cobb, T. (2014). Compleat Web VP, version 2 [Computer software].

Crossley, S., Dufty, D., McCarthy, P., & McNamara, D. (2007). *Toward a new readability: A mixed model approach*. In Proceedings of the 29th Annual Conference of the Cognitive Science Society (pp. 197–202). Cognitive Science Society.

Davies, M. (2012). Corpus of Contemporary American English (1990–2012) [Dataset]. University of North Texas Libraries, UNT Digital Library.

Davies, M., & Gardner, D. (2010). *A frequency dictionary of contemporary American English: Word sketches, collocates and thematic lists*. Routledge.

de Bot, K. (2017). Complexity theory and dynamic systems theory: Same or different? In L. Ortega & Z. Han (Eds.), *Complexity theory and language development. In celebration of Diane Larsen-Freeman* (pp. 51–58).

de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition, 10*(1), 7–21. https://doi.org/10.1017/S1366728906002732

Dörney, Z. (2009). *The psychology of second language acquisition*. Oxford University Press.

Durán, P., Malvern, D., Richards, B., & Chipere, N. (2004). Developmental trends in lexical diversity. *Applied Linguistics, 25*(2), 220–242. https://doi.org/10.1093/applin/25.2.220

Ellis, N. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143–188. https://doi.org/10.1093/applin/16.3.307

Ellis, N. (2006). Cognitive perspectives on SLA: The associative-cognitive CREED. *AILA Review, 19*, 100–121. https://doi.org/10.1075/aila.19.08ell

Elman, J. (2001). Connectionism and language acquisition. In M. Tomasello & E. Bates (Eds.), *Language development: The essential readings* (pp. 295–306). Blackwell.

Hemchua, S., & Schmitt, N. (2006). An analysis of lexical errors in the English compositions of Thai learners. *Prospect, 21*(3), 3–25.

Hiver, P., & Al-Hoorie, A. (2020). *Research methods for complexity theory in applied linguistics*. Multilingual Matters.

Hopper, P. (1998). Emergent grammar. In M. Tomasello (Ed.), *The new psychology of language* (pp. 215–243). Lawrence Erlbaum Associates.

Hou, J., Loerts, H., & Verspoor, M. (2020). Coordination of linguistic subsystems as a sign of automatization? In G. Fogal & M. Verspoor (Eds.), *Complex dynamic systems theory and L2 writing development*. John Benjamins.

Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 1–20). John Benjamins.

Huljstin, J. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research, 18*(3), 193–223. https://doi.org/10.1191/0267658302sr207oa

Hyltenstam, K. (1988). Lexical characteristics of near-native second-language learners of Swedish. *Journal of Multilingual and Multicultural Development, 9*(1/2), 67–84. https://doi.org/10.1080/01434632.1988.9994320

James, C. (1998). *Errors in language learning and use: Exploring error analysis*. Longman.

Knoch, U., Rouhshad, A., & Storch, N. (2014). Does the writing of undergraduate ESL students develop after one year of study in an English-medium university? *Assessing Writing, 21*(1), 1–17. https://doi.org/10.1016/j.asw.2014.01.001

Kyle, K., & Crossley, S. A. (2015). Automatically assessing lexical sophistication: Indices, tools, findings, and application. *TESOL Quarterly, 49*(4), 757–786. https://doi.org/10.1002/tesq.194

Kyle, K., Crossley, S., & Berger, C. (2018). The tool for the automatic analysis of lexical sophistication (TAALES): Version 2.0. *Behavior Research Methods, 50*(1), 1030–1046. https://doi.org/10.3758/s13428-017-0924-4

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics, 18*(2), 142–165. https://doi.org/10.1093/applin/18.2.141

Larsen-Freeman, D. (2006). The emergence of complexity, fluency and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics, 27*, 590–616. https://doi.org/10.1093/applin/aml029

Larsen-Freeman, D. (2011). A complexity theory approach to second language development/acquisition. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (pp. 48–72). Routledge.

Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford University Press.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 18*(2), 141–165. https://doi.org/10.1093/applin/16.3.307

Linnarud, M. (1986). *Lexis in composition. A performance analysis of Swedish learners' written English*. CWK Gleerup.

Lowie, W., & Verspoor, M. (2019). Individual differences and the ergodicity problem. *Language Learning, 69*, 184–206. https://doi.org/10.1111/lang.12324

Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal, 96*(2), 190–208. https://doi.org/10.1111/j.1540-4781.2011.01232_1.x

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Lawrence Erlbaum Associates.

MacWhinney, B. (2005). A unified model of language acquisition. In J. Kroll & A. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). Oxford University Press.

Malvern, D., Richards, B., Chipere, N., & Duran, P. (2004). *Lexical diversity and language development: Quantification and assessment*. Pelgrave MacMillan.

McNamara, D., Crossley, S., Roscoe, R., Allen, L., Dai, J., et al. (2015). A hierarchical classification approach to automated essay scoring. *Assessing Writing, 23*, 35–59. https://doi.org/10.1016/j.asw2014.09.002

Meara, P., & Bell, H. (2001). P-Lex: A simple and effective way of describing the lexical characteristics of short L2 tests. *Prospect, 16*(3), 5–19.

Nation, P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review, 63*(1), 59–82. https://doi.org/10.3138/cmlr.63.1.59

Nation, P. (2016). *Making and using word lists for language learning and teaching*. John Benjamins.

Nation, P., & Heatley, A. (2002). Range: A program for the analysis of vocabulary in texts [Computer software].

Nuance®. (2014). *Dragon naturally speaking premium*. Speech Recognition Software.

Preacher, K., & Selig, J. (2012). Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures, 6*(2), 77–98. https://doi.org/10.1080/19312458.2012.679848

Read, J. (2000). *Assessing vocabulary*. Oxford University Press.

Saffran, J., & Thiessen, E. (2007). Domain-general learning capacities. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 68–86). Blackwell.

Salkind, N. (2010). *Encyclopaedia of research design*. Sage.

Spoelman, M., & Verspoor, M. (2010). Dynamic patterns in the development of accuracy and complexity: A longitudinal case study on the acquisition of Finnish. *Applied Linguistics, 31*(4), 532–553. https://doi.org/10.1093/applin/amq001

Storch, N., & Tapper, J. (2009). The impact of an EAP course on postgraduate writing. *Journal of English for Academic Purposes, 8*(3), 207–223. https://doi.org/10.1016/j.jeap.2009.03.001

Text Inspector. (2018). Online lexis analysis tool [Computer software].

Tomasello, M. (2015). The usage-based theory of language acquisition. In E. Bavin & L. Naigles (Eds.), *The Cambridge handbook of child language* (pp. 89–106). Cambridge University Press.

Ure, J. (1971). Lexical density: A computational technique and some findings. In M. Coultard (Ed.), *Talking about text* (pp. 24–48). English Language Research.

Van Geert, P., & Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behaviour and Development, 25*, 340–375. https://doi.org/10.1016/S0163-6383(02)00140-6

Van Geert, P., & Verspoor, M. (2015). Dynamic systems and language development. In B. MacWhinney & W. O'Grady (Eds.), *The handbook of language emergence* (pp. 537–556). Wiley.

Verspoor, M., Lowie, W., & Van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. *Modern Language Journal, 92*, 214–231. https://doi.org/10.1111/j.1540-4781.2008.00715.x

Verspoor, M., de Bot, K., & Lowie, W. (2011). *A dynamic approach to second language development*. John Benjamins.

Verspoor, M., Schmid, M. S., & Xu, X. (2012). A dynamic usage based perspective on L2 writing. *Journal of Second Language Writing, 21*(3), 239–263. https://doi.org/10.1016/j.jslw.2012.03.007

West, M. (1953). *A general service list of English words*. Longman.

Wolfe-Quintero, K., Ingaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. University of Hawaii, Second Language Curriculum Centre.

Zheng, Y. (2016). The complex, dynamic development of L2 lexical use: A longitudinal study on Chinese learners of English. *System, 56*, 40–53. https://doi.org/10.1016/j.system.2015.11.00