Language Teaching Research Quarterly

2024, Vol. 39, 256–280

A DLF Case Study: The Dynamics of Writing Development in Adulthood[1]

Thanh T.G. Trinh[1], Kees de Bot[2], Marjolijn Verspoor[2]*

[1]Ho Chi Minh City University of Technology and Education, Vietnam
[2]University of Pannonia, Hungary

Received *3 July 2023* ***Accepted*** *22 October 2023*

Abstract

This longitudinal case study from a Complex Dynamic Systems Theory (CDST) perspective touches upon an under-researched issue: L1 development over the lifespan. Levinson (1978) predicts three stages in adulthood: early, mid and late, with a decline in late adulthood. We examine Diane Larsen-Freeman's publications over a period of 50 years (from age 27 to 77) and trace seven complexity measures--three lexical (density, sophistication and diversity) and four syntactic (mean length of sentence, finite verb ratio, dependent clause per T-unit and complex nominals per clause)--to investigate whether early, middle, and late stages in adulthood occur as predicted. After employing common CDST methods to find out if there are significant peaks or interactions among the variables over time, we used a Hidden Markov time-series analysis to locate moments of self-organization, suggesting a new stage of development. The HMM shows a clear phase shift between middle and late adulthood when the writer was 63. Her vocabulary became more diversified, but her sentences were shorter, but not less complex. Therefore, we argue that this shift should not be seen as a decline in complexity but a shift in style as more precise words may lessen the need for more words.

Keywords: *CDST, L1 Development, Hidden Markov Model*

Introduction

In 2006, Diane Larsen-Freeman (DLF) argued that second language development research should not adhere to a "developmental ladder" metaphor (Fischer et al., 2003) but should be guided by CDST in seeing development as a dynamic process with 'make-do' solutions in which development is not discrete and stage-like but more like the waxing and waning of patterns. In her paper, she discussed in detail five case studies of Chinese learners of English and traced their linguistic development over six months. She showed that complexity, fluency, and accuracy did not unfold according to some prearranged plan, but as individually owned systems adapting to a changing context through use. In this paper, we will use a very similar

[1] This article is partly based on Thanh (2011).

* Corresponding author.
E-mail address: m.h.verspoor@rug.nl
https://doi.org/10.32038/ltrq.2024.39.17

case study design, not on L2 development, but on L1 development throughout the lifespan. In our case, it is a case study on DLF.

In the biological-maturational field it has been often assumed that development no longer occurs when maturity comes (Birren, 1964). Children's language has been assumed to develop and progress until adulthood and the language of older adults to decline from middle adulthood (Cheshire, 2005). These assumptions are closely connected with the "critical period" proposed by Lenneberg (1967) and are somewhat in line with Labov (2001), who suggested a linear model of incrementation predicting that language becomes stable at the age of 17, the starting point of the threshold period.

Recently, a great number of publications have rejected such a static view on language. It is universally accepted that language should be regarded as a complex, dynamic system (van Geert, 1991, 1994, 2008; Larsen-Freeman, 1997; de Bot et al., 2005, 2007; Ellis & Larsen-Freeman, 2009), which is "a set of variables that mutually affect each other's changes over time" (van Geert, 1994, p. 50). A linguistic system is subject to constant change and the state it is set in (attractor state) is temporary and that order appears through fluctuations (Prigogine, 1978). Change can be caused by either external interactions or internal reorganization of sub-systems. That is to say, a linguistic system self-organizes through interactions. As they proceed, developmental phase shifts occur, triggered by the interactional patterns of different linguistic sub-systems, such as phonology, morphology, and syntax. Phase shifts thus refer to "the coming-into-existence of new forms or properties through ongoing processes intrinsic to the system itself" (Lewis, 2000, p. 38). And the "coming-into-existence" is known as emergence, the key concept invoked as "a general principle for explaining developmental change" (Lewis, 2000, p. 38) throughout the system with no terminal point.

Within a CDST paradigm, L1 intra-individual variability among 1 to 8-year-old children was investigated by van Geert (1991, 1994), van Geert and van Dijk (2002). This age bracket was labeled as the "childhood" stage by Levinson (1978). However, not many CDST studies have investigated the development of language during early adulthood (17-45), middle adulthood (45-65) and late adulthood (65 onwards) (Levinson, 1978). In his review of language development research across the lifespan, de Bot (2007) mentions that "there are hardly any investigations on language development in the age range from 18 till 55" (p. 55), probably because of the erroneous assumption that the fully acquired linguistic system of an adult individual stabilizes (de Bot, 2007). To fill this research gap, we will trace L1 development over three different phases of adulthood, from early to late adulthood. [2]

Based on CDST principles, we have conducted a 50-year longitudinal case study of Diane Larsen-Freeman (DLF) by investigating 60 carefully selected 200-word excerpts from her peer-reviewed publications. We began by coding them for several common lexical and syntactic variables. These variables were then traced over time in detail to see if there were meaningful developmental jumps or interactions between the variables, over time and per adulthood stage (early, middle, and late). In addition, to identify possible phases operationalized as moments of re-organization of the linguistic system, a Hidden Markov time-series analysis was used to

[2] The terms *stage* (the preferred term when dealing with cyclical or oscillatory changes) and *phase* (the preferred term to indicate some form of progression towards an expected end state) are not used consistently in the literature; therefore, we will use the terms as they were used in the literature.

detect the hidden variables out of the dataset of observable variables and locate moments of self-organization where the interactions among sub-systems changes, triggering a new state of development. Thus the aim was to test whether an adult's language shows different phases across adulthood, and to determine whether there was any decline. We did so first by means of traditional CDST analyses and then by means of the Hidden Markov Model (HMM).

Background Literature

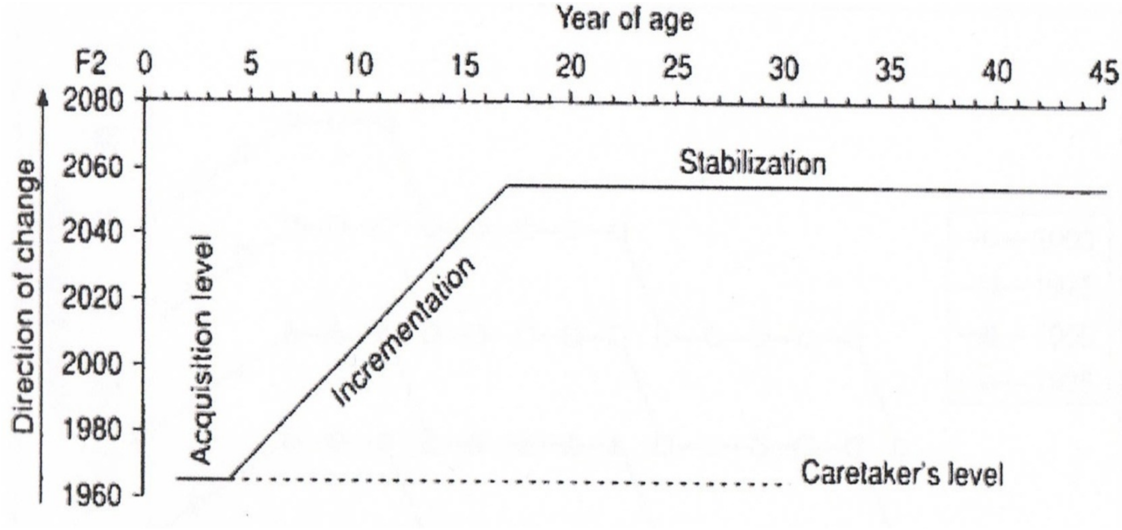
Changes across the lifespan have been of great interest to developmental psychologists. According to Levinson (1978), one is supposed to go through six stages of development accompanied with maturational and psychological changes. These include: (1) Stage 1: Early childhood with the age from 0 to 3; (2) Stage 2: Late childhood from age 3 to age 12; (3) Stage 3: Adolescence between age 12 and17; (4) Stage 4: Early adulthood (age 17-45); (5) Stage 5: Middle adulthood lasting 20 years from age 45 to 65; (6) Stage 6: Late adulthood (age 65 onwards).

These six stages generally cover such dimensions of life as physical health, cognition and social development. As Nevid and Rathus (2005) reported, a typical individual in young adulthood has the best physical condition in terms of strength, health, sharp senses and stamina and the best condition in terms of memory and thinking (cognitive) abilities. In middle adulthood, physical development is marked with a loss of strength, vision and flexibility and some signs of memory loss and reduced ability to perform speed-requiring tasks, and a so-called “mid-life crisis” may enhance either positive or negative health changes. Finally, in late adulthood (65 years old or older), there is a deterioration in terms of physical health, memory and mobility, verbal skills and knowledge, and a readjustment to biological changes and social life (e.g. retirement).

In the field of sociolinguistics, Labov (2007) also assumes that language change or incrementation (“change from below”) is staged as a function of age, from early acquisition to language stabilization. While the starting age at which language changes is supported empirically, it is not yet clear when such stabilization begins to occur (Tagliamonte & D’Arcy, 2009).

Labov (2001) developed a so-called linear incrementation model (see Figure 1), illustrating that the acquisition period takes place from age 0 to age 3, then language grows continuously and steadily from age 4 to 17 when it comes to the threshold period. Labov (2001)’s model suggests that language becomes stable at the outset of adulthood (based on the data of a woman whose language development was traced until the age of 45). However, his earlier work in 1994 states that individuals’ language may still be subject to change during their lifetime, but this may not be so for the whole community.

Figure 1
Linear Model of Incrementation for a Single Female Speaker from 1 to 45 Years of Age



Challenging Labov’s lifespan model, Baltes et al. (1980) characterized lifespan development as a non-linear, multi-dimensional and multi-directional process. Indeed, there is evidence that major life events in both adults and children may play an important role in language development (Cheshire, 2005). According to Hankin and Abramson (2001), substantial negative life events causing stress and depression may affect a person's attitude and social relations (as cited in de Bot & Schrauf, 2009). Likewise, other major life events in education, work, emigration, marriage, divorce, retirement, etc. may affect language development (de Bot, 2007; Lowie et al., 2009), which is in line with Baltes et al. (1980), who argue that development is not only a matter of ontogeny but is also influenced by historical factors. De Bot (2007) suggests that such a complex lifespan perspective provides a framework that may be useful for the study of language development.

In terms of language changes across the lifespan, there is some empirical evidence of decline in late adulthood. In their meta study, Kemper et al. (2001) found that there is a growth in vocabulary throughout middle adulthood and a decline in late adulthood (cf. Botwinick & Siegler, 1980; Zelinski & Burnight, 1997). According to Bowles et al. (2005), basic vocabulary becomes optimal around the age of 30 and declines in late adulthood whereas there is no dependence of advanced vocabulary on age when one is between 35 and 70 years old. According to Kavé (2022), ageing “may lead to an accumulation of vocabulary because people are continuously exposed to language through education, incremental reading or life experience” (p. 2). Meylan and Gahl (2014) examined the interaction between age, intra-speaker lexical diversity and inter-speaker lexical overlap using a corpus of conversations (Switchboard I). Based on Ramscar et al. (2014), they hypothesized that language production changes as a consequence of lifelong learning and found that “a speaker’s lexical diversity is conditioned on the properties of his or her interlocutor, but age and higher levels of education predict increased lexical diversity for individual speakers” (p. 1010).

Nippold et al. (2005) explored the development of syntax by observing mean length of T-unit and relative clause production and concluded that the growth in syntax is found throughout childhood and adolescence and continues in early adulthood. Not until the middle age does syntax become stable. Their findings imply that complexity in thinking drives the development of linguistic complexity and that "individual variability can exist at all points along the age continuum" despite the developmental trend of syntax growth (p. 1).

Longitudinal CDST studies have taken place in the field of second language development and have aimed at investigating the trajectory of separate measures (cf. Larsen-Freeman, 2006; Verspoor et al., 2008, 2012; Penris & Verspoor, 2017). These data analyses have revealed individual differences, the role of variability in progress, and trade-off effects among measures. In our study, we will trace the measures in a similar manner to begin with and focus on the relations that these measures reveal, such as precursor, competitive, or supportive ones (Caspi & Lowie, 2013; Chan et al., 2015) by means of correlation analyses.

Another set of complex dynamic systems studies have looked at transition phases, which is quite relevant to our own study, testing the age-hypothesis. For example, Bassano and van Geert (2007) examined phase shifts of grammatical development in early language. Baba and Nitta (2014) investigated phase transitions in the development of writing fluency of two students of English based on the four phase transition criteria including sudden jumps, anomalous variance, divergence, and qualitative change. These are four out of the eight criteria proposed by catastrophe theory, which Zeeman (1976) claimed to be possibly applicable to sciences such as biology and social sciences where discontinuous and divergent situations are prevalent.

As the number of variables and observed measures increases, it becomes more difficult to observe any changes in these interactions and model them (cf. Lowie et al., 2011). The current study will therefore follow Chan et al. (2015) in using an unsupervised HMM, which detects discontinuity patterns based on the input of data. First of all, a string of data (or a value) is fed into the model for an analysis of change patterns. Then an optimal sequence of changes will be detected, and the data point or the indicator of phase boundaries where there is a shift in the complex system is identified.

Assuming that adult L1 may not be as stable as commonly thought, the current study centers on the language development over 50 years of DLF, an expert writer, whose native language is English. The focus is on changes in lexical and syntactic complexity. The overall research question is: Can stages be identified in adulthood and if so, is there a decline? To answer the overall question, we will focus on the following specific research questions.

RQ$_1$: Are there changes in lexical and syntactic complexity measures over time?
RQ$_2$: Are there significant peaks in lexical and syntactic complexity measures over time?
RQ$_3$: Are there significant changes in lexical and syntactic complexity measures between the assumed three stages of adulthood (early, mid and late)?
RQ$_4$: Are there any significant correlations between lexical and syntactic complexity measures over time?
RQ$_5$: What states are identified by HMM modeling?
RQ$_6$: Are there differences in the measures between HMM's states?

Method

Assessing the linguistic development of the written language of an individual involves three general aspects, i.e., accuracy, complexity and fluency (Larsen-Freeman, 2006). Since the subject of this study is a highly prestigious L1 scholar, her language accuracy and fluency are assumed to be at ceiling level and will be ignored. This study will focus on the third aspect: complexity. It will involve two sub-systems: the lexicon and syntax.

Participant

The subject (DLF) is an American researcher in applied linguistics. The current study traces the linguistic complexity in texts written from 1974 right before she finished her PhD (when she was 28 years old) up to 2022 (when she was 76 years old).

Materials

The materials include 32 articles and 28 book chapters (all peer reviewed) (See Appendix A) on several main themes: SLA and DST research, grammar, language and teacher education. The texts, all single authored, were all written for academic purposes over the course of almost 50 years. To control for the potential differences in book chapters and journal articles, the measures were compared. An independent sample T-test conducted on the lexical and syntactic complexity measures showed no significant differences between them. (See Appendix B.)

From each publication, a 200-word sample was selected randomly from either the introduction or the background. To make sure the analyses reflected the language used by the subject herself, direct quotations were removed, and proper names and numbers were replaced with NAME and NUMB, respectively.

In the data set, the samples were chronologically sequenced according to publication year. If more than one writing from one certain year was qualified for selection, they were arranged as suggested in the Curriculum Vitae provided by the subject herself. However, it should be noted that many of these writings may have been produced at different and overlapping times because of time differences in publication times, so the dates can only be considered approximations of when they were written.

Measures

Complexity at the lexical level is conceptualized as a cluster of three different dimensions, namely lexical density, lexical sophistication and lexical diversity (Read, 2000).

Lexical density is the ratio of the number of lexical or content words over the total number of words in a text (Ure, 1971). High lexical density means greater complexity since the text contains more information-carrying words (Lankshear & Knobel, 2004).

The second dimension is lexical sophistication, the percentage of sophisticated words used in a text. Average word length (AWL) was chosen for this study because this index can reflect sophistication in both terms of length and frequency as Grant and Ginther (2000) showed that proficient English writers tend to use longer words. Moreover, academic or less frequent words are generally longer than frequently used words (Verspoor et al., 2011). Since function words (e.g. prepositions) are commonly found with high frequency in writings, especially in nominalizations, we felt that the inclusion of functions words might even out average word length. Therefore, in this study, AWL concerns only content words.

Measures of lexical diversity, an indicator of vocabulary range, are numerous. However, since all texts were approximately the same length and all function words were excluded in the counts of word types and tokens, a simple Guiraud was used for this study. The advantage of a rooted rather than a raw Type-Token Ratio is that it tends to flatten the curve somewhat so that more subtle differences can be observed in the diversity of lexical item used.

$$\boxed{GI = {}^{types} / \sqrt{\,_{tokens}}}$$

Complexity at the syntactic level was measured by four different but overlapping dimensions. The measure of mean length of sentence (MLS) is a broad measure proven to be a good indicator of increasing complexity by previous studies (Cheung & Kemper, 1992 for L1 assessment of old-aged people, Wolfe-Quitero et al., 1998 for L2). MLS indicates the average number of words per sentence/utterance; however, it does not reflect sentence internal complexifications with nominalizations and non-finite constructions. According to Verspoor et al. (2008), the finite verb token ratio (FVR) is considered a useful measure for syntactic sophistication at a more advanced level, and it correlated significantly with almost all other complexity measures in the Penris and Verspoor study (2017). The FVR is calculated by dividing the number of word tokens by the number of finite verbs in each text. A high FVR indicates more language complexity. The FVR is a general length measure that includes all types of non-finite constructions including those functioning as adverbials but does not indicate the relative frequency of nominalizations, which may contain non-finite verb constructions; thus, we also calculated the number of complex nominals per clause (CN_C). Lu (2011) describes complex nominals per clause as "syntactic constructions such as nominal clauses, infinitives or gerunds in the subject position, nouns modified by adjectives, adjective clauses, appositives, prepositional phrases and/or possessives" (as cited in Kyle, 2016, p.13). Finally, the index of dependent clauses per T-unit (DC_T), which is well described and developed as one of the 14 indices of syntactic complexity (Lu, 2010) was chosen. According to Biber and Gray (2010), such finite dependent clauses actually occur less in advanced academic writing than in speech, and are not per definition a sign of more complex language, and in the Penris and Verspoor study (2017), they decreased over time.

Procedures

Lexical density index was computed on Lextutor, (www.lextutor.ca). Word Frequency Counter software (Writewords website) was used to obtain the list of vocabulary items together with their frequency. Then, the list was fed into Excel where functional words were found and removed. The Guiraud and AWL were computed with the remaining content words in the list. Syntactic complexity variables, i.e. MLS, DC_T, CN_C, were measured through the Tool for the Automatic Analysis of Syntactic Complexity (TAASC) (Kyle, 2016). The FVR was done manually.

Design and Analyses

We adopted a longitudinal design. To test differences between the three assumed stages of adulthood, the dataset was divided into three general sets, reflecting stages of life depending on the age factor: early adulthood (16 texts; from age 28 to age 44), middle adulthood (24 texts; from age 45 to age 64) and late adulthood (20 texts; from age 65 to 76) (Levinson, 1978; Labov, 2010). However, our dataset is not complete as we miss about 10 years of texts from the early adulthood stage, which is supposed to span from age 17 to 45. Because of this missing data, we assumed that the dataset might not show a clear difference between early and middle adulthood and only two stages (middle and late adulthood) might be observable.

We employed the visual inspection and analyses for longitudinal data which were originally developed by van Geert and van Dijk (2002) to visualize the dynamic developmental trajectories of language of children at early ages. In addition, we also used the adapted techniques and methods to trace L2 development as guided in Verspoor et al. (2011).

To sketch the general developmental pattern of the developmental variables, line graphs and trendlines (linear) are used. When developmental jumps are seen in a single variable, the "peak" is tested for significance. Significant peaks are considered as markers for transitions between developmental stages. Poptools (Hood, 2004), an add-in tool of Excel was employed to run a Monte Carlo simulation, which is set to run 5000 iterations to calculate how many times peak values of the simulated sets are higher than those of the observed set. P-value is then obtained by dividing this number with the number of iterations.

The data is normalized (Verspoor et al., 2011) and graphed with moving average trendlines for comparison between various measurement units. These kinds of visualizations provide well-grounded assumptions for detecting the change points throughout the dataset as determined by subsequent modelling (van Dijk et al., 2011).

To test whether time series data significantly increases or decreases over time, the Mann-Kendall Test for Trend of XLSTAT (an add-in tool for Excel) is used. In the Mann-Kendall statistical test for trend, the null hypothesis is made that there is no trend over time for the dataset, which means the data are independent and randomly ordered and the three alternative hypotheses are that the trends are upward, downward or no-null (i.e. the data is time-ordered). The Mann-Kendall test differs from other types of statistical tests in that the magnitude of change is not assessed.

In order to discover the relationship between variables, we turn to SPSS 26's Pearson correlation tests. Independent sample t-tests and one-way MANOVAs are run to test the differences in the measures' mean scores of different stages of adulthood.

Finally, to detect the moments at which the linguistic complexity system self-organizes under the interactions of subsystems' growers, the dataset is modeled with an unsupervised Hidden Markov Model (HMM). HMMs, often used in speech recognition, are statistical models that capture hidden information from observable sequential symbols. In a HMM, the system to be modelled is assumed to have Markov properties, i.e. a transition from one state to another and observational parameters (in this study, these are our measures of lexical and syntactic complexity). The aim is to determine the hidden parameters from the observable ones. (Stamp, 2017)

In the current study, we used normalized data for modelling in order to facilitate the comparison between measures of different measurement units. The coding process was conducted on Python with the step-by-step guidance from the HMM tutorial (see at website of Hmmlearn developers). Based on the literature, we hypothesized that there would be two or three developmental states based on age-based stages of adulthood. The HMM was to figure out the transitional matrix based on the assumed number of states (2 or 3), the observational parameters and the initial matrix (we assumed the developmental state always started from

State 1). The output of the modeling includes: (a) the change point at which the phase shift took place; (b) mean score of each variable for each state; and (c) standard deviation for each measure at each state. (Stamp, 2018)

Results

In the following sections, first all measures are described to determine change over time in terms of developmental jumps and interactions among variables. The last section will discuss the HMM analysis.

Lexical Measures

With respect to lexical development, Figure 2 shows that the Guiraud increases, AWL goes slightly down, and lexical density remains stable throughout the dataset. As only Guiraud indicates an increase, this data is tested with a Monte Carlo analysis for a significant peak. The results reveal that Guiraud does not have any significant peaks over time (p= 0.9184) despite its steady increase.

Figure 2

Guiraud, Lexical Density and AWL with Linear Trendlines

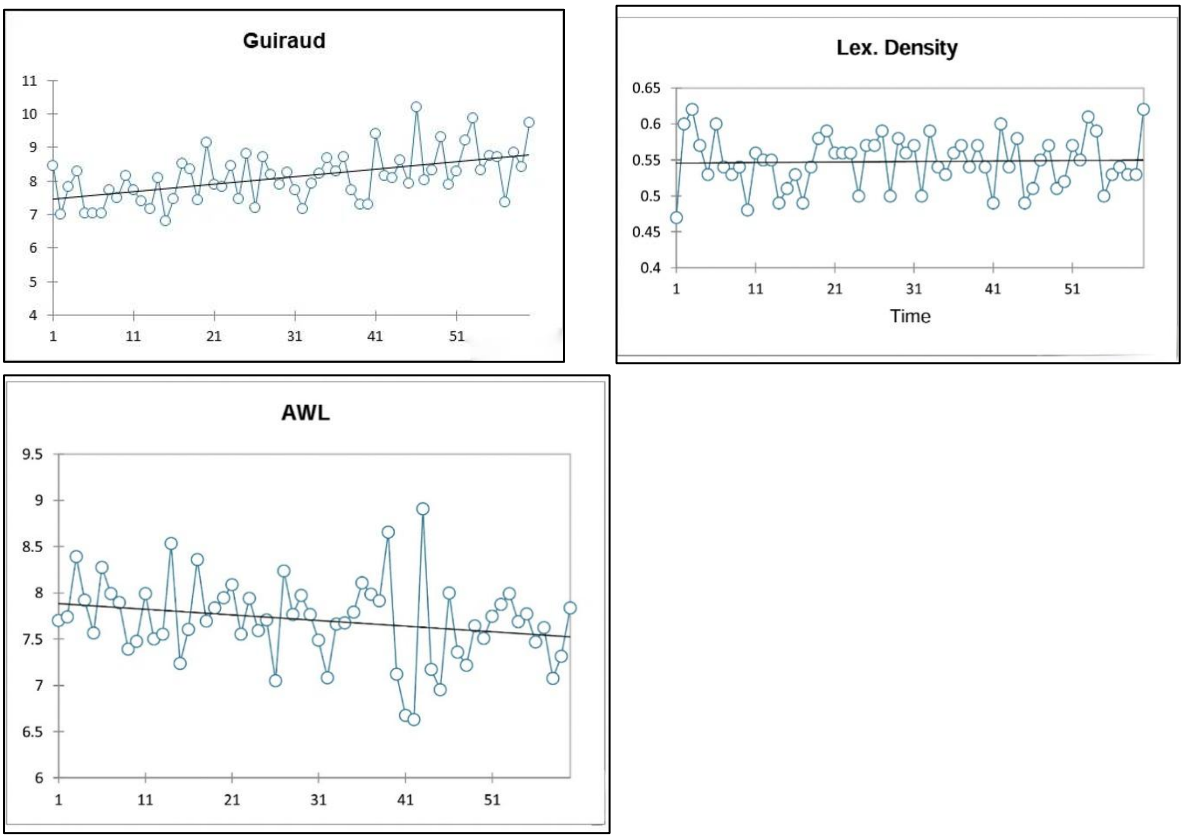


Table 1 presents the results of Mann-Kendall trend tests for the three measures of lexical complexity. It is clear that the Guiraud increases over time significantly (Kendall's tau=0.359, p<0.0001), whereas the other two measures obtain no significant trends.

Table 1

Results of Mann-Kendall Trend Tests

	Guiraud	AWL	Lexical density
Kendall's tau	0.359	-0.145	0.016
S	635	-256	28
Var(S)	24582.333	24583.333	24333.333
p-value (Two-tailed)	**<0.0001**	0.104	0,863
alpha	0,05	0,05	0,05

Figure 3 shows the interactions with moving averages (5 data points) among the three measures of lexical complexity over time. If we look at where the constellation of the lines changes, the data shows roughly three stages. The moving average of lexical density declines from Text 1 to Text 18 (Stage 1) and becomes stable from Text 18 to Text 42 (Stage 2). After that, it starts to decline again before recovering at Text 51. The moving average of Guiraud increases from Text 1 to Text 18 (Stage 1), remains relatively stable from Text 18 to Text 42 (Stage 2), and increases again from Text 43 till the end of the dataset (Stage 3). The moving trendline AWL does not fluctuate much until Text 42 (end of Stage 2); then it suddenly decreases and reaches a new equilibrium.

Figure 3

AWL, Guiraud and Lexical Density with Moving Average of 5 Datapoints

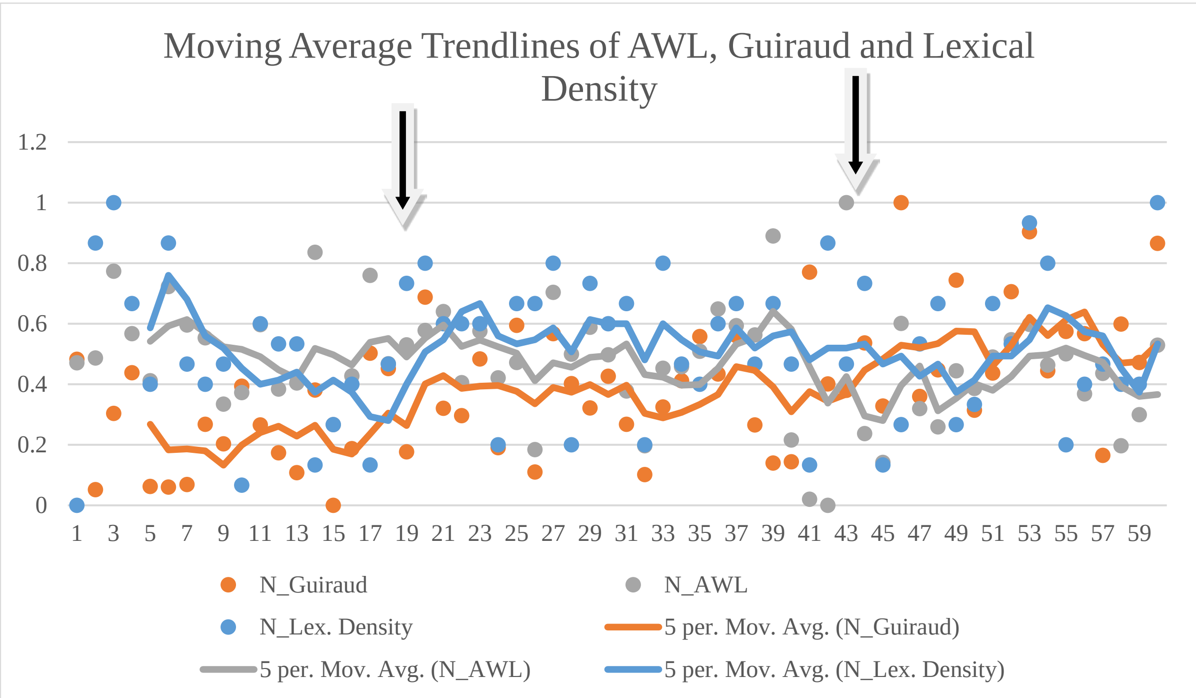


Table 2 shows the results of Pearson correlation test to discover the relationship between the lexical complexity variables. The test reveals no significant relationship among these three measures across the period of 50 years (Pearson r < 2.0 and p >0.05).

Table 2

Correlations of Guiraud, AWL and Lexical Density

		Guiraud	AWL
AWL	Pearson Correlation	.063	
	Sig. (2-tailed)	.633	
Lexical density	Pearson Correlation	.015	.197
	Sig. (2-tailed)	.911	.131

In order to test the differences of the three measures of lexical complexity based on three different stages of life (with age being a factor for life stage divisions), the data were fed into SPSS 26 and one-way MANOVA was run. The statistical result reveals that there are significant differences between three measures of lexical complexity at three different stages of adulthood (Wilks' λ =0.542, p=0). The measure that contributes to the mean differences is Guiraud (F=15.084, p= 0). Table 3 shows the mean scores of the three variables and it is evident that Guiraud shows striking differences between the three stages. Post-hoc tests confirm that the subject used significantly more different types of words in her late adulthood than in her early and middle adulthood with the Guiraud mean differences of 0.1295 and 0.6283 respectively, sig. value =0 and 0.004 <0.05 (p-value).

Table 3

Mean scores of Guiraud, AWL and Lexical Density of Three Stages of Life

Mean score	Early adulthood (Text 1-16)	Middle adulthood (Text 17-40)	Late adulthood (Text 41-60)
Guiraud	7.546	8.047	8.676
AWL	7.799	7.792	7.523
Lexical density	0.542	0.552	0.547

Syntactic Measures

Figure 4 shows that MLS reveals two outstanding peaks at Text 6 and Text 34. However, it has a downward trendline. Finite verb token ratio (FVR) has two peaks at Text 13 and 37, but also seems to decrease over time. Monte Carlos simulations were run to test the peaks. They were not significant.

Figure 4

Mean Length of Sentences and Finite Verb Token Ration with Linear Trendlines

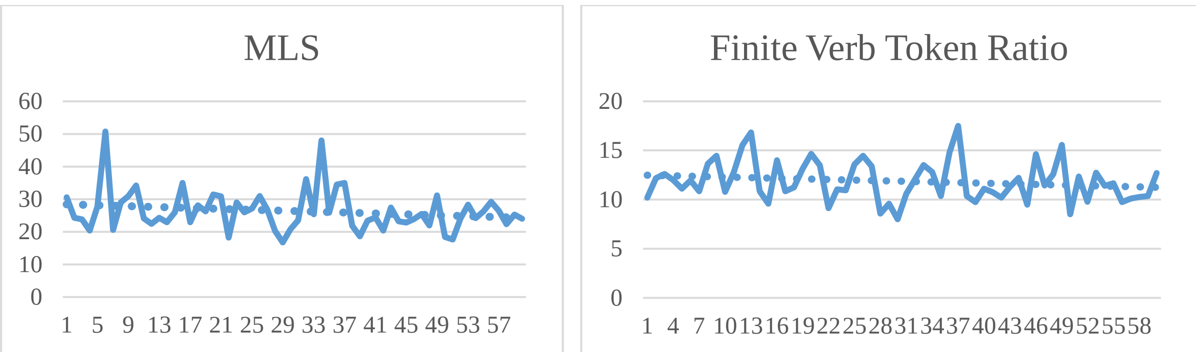


Figure 5 shows that dependent clauses per T-unit (DC_T) remain rather stable, and complex nominals per clause seem to decrease very lightly. Three peaks can be clearly seen in the DC_T and CN_C but the heights of the subsequent peaks seem to be lower than those of the previous peaks. Monte Carlos simulations showed no significant peaks.

Figure 5

Dependent Clauses per T-unit and Complex Nominals per Clause with Linear Trendlines

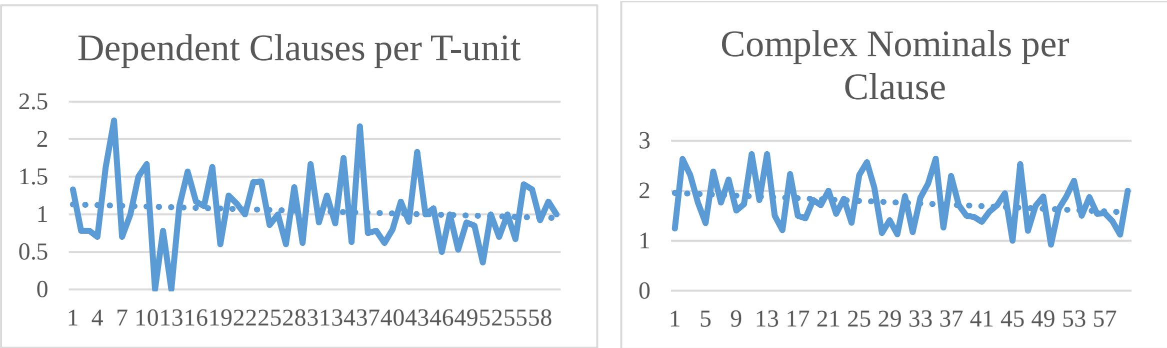


Table 6 presents the results of Mann-Kendall trend tests for syntactic complexity measures. All of the measures have negative trends but none of the trends are quantitatively significant.

Table 6

Results of Mann-Kendall Trend Tests for Syntactic Complexity Measures

	MLS	FVR	DC_T	CN_C
Kendall's tau	-0.104	-0.114	-0.077	-0.145
S	-183	-202	-135	-256
Var(S)	24574.333	24578.667	24493.0	24564.0
p-value (Two-tailed)	0.246	0.2	0.392	0.104
alpha	0.05	0.05	0.05	0.05

Figure 6 shows the moving average trend lines for all four syntactic measures. Even though all four reveal downward trends, there seem to be four stages for syntactic complexity. From Text 1 to Text 10, MLS shows a steady growth followed by a decline and a dip until Text 15. It then picks up slightly and starts to stabilize from Text 17 onwards. Similarly, DC_T goes down at first (Text 1 to Text 15), but then moves up more strongly than MLS. FVR shows an increasing pattern until Text 17 and then stabilizes. CN_C (Complex nominals per clause) starts off relatively higher and seems to have a reversed pattern from Text 1 to Text 16. However, it makes a dip at Text 17 and then stabilizes. The four measures are relatively stable between Text 17 until Text 31 when three measures decrease substantially but remain somewhat more stable. Around Texts 35-37, new peaks occur and seem to signal a new stage of development. The whole dataset is rather stable again from Text 41-59.

Figure 6

Moving Averages (window of 5) of Mean Length of Sentences, Finite Verb Token Ration, Dependent Clauses per T-unit and Complex Nominals per Clause

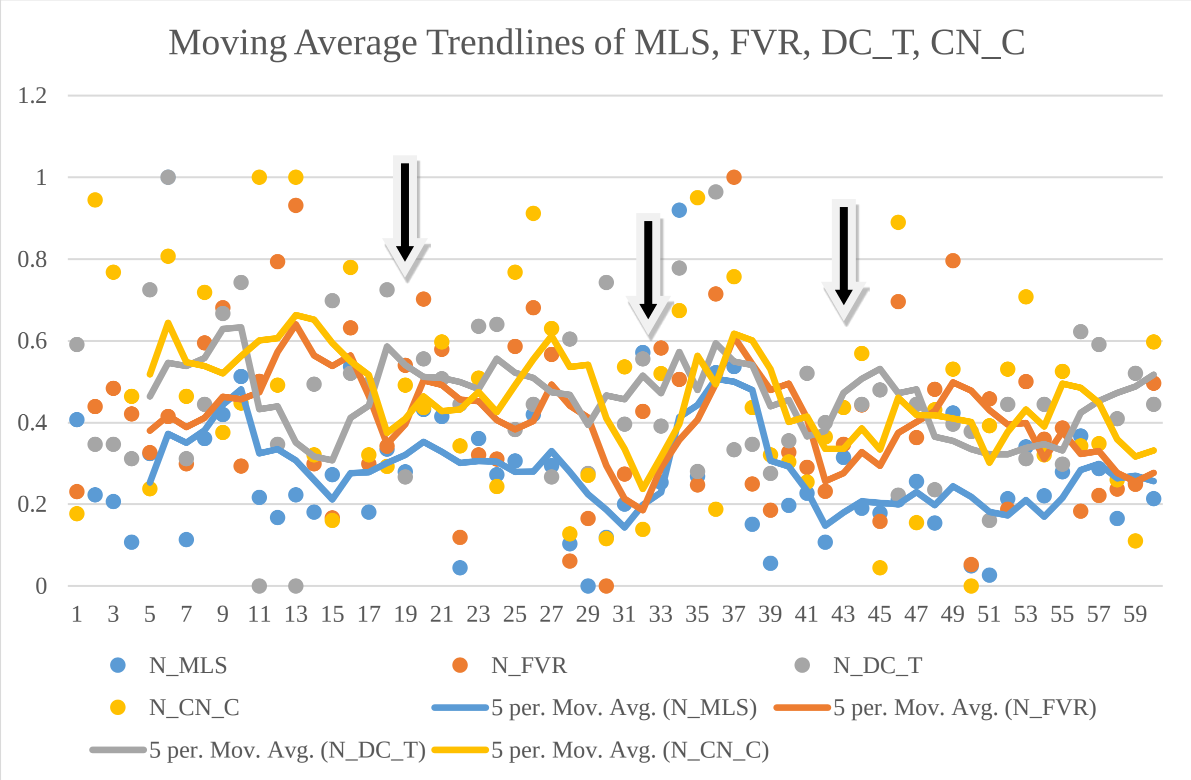


Table 7 shows the results of a Pearson correlation test to test the significance of the interactions with the dimension of syntactic complexity. MLS has a moderate, supportive relationship with DC_T, which means that the number of dependent clauses contributes to average sentence length (Pearson r = 0.569, p=0). Similarly, FVR and CN_C are positively related to MLS. (Pearson r = 0.398 and 0.281, p=0.002 and 0.029 respectively). Together, it can be inferred that the changes in MLS are dependent on the changes in DC_T, FVR and CN_C. A significant competitive relationship is statistically observed for DC_T and CN_C (Pearson r= -0.393, p=0.002), which entails that when the subject uses more finite dependent clauses, she tends to use fewer elaborated noun phrases. Another finding is that there is a significant strong correlation between FVR and CN_C (Pearson r= 0.603, p=0), which makes sense as the FVR measure subsumes complex nominals.

Table 7

Correlations of MLS, FVR, DC_T, CN_C

Measures		MLS	FVR	DC_T
FVR	Pearson Correlation	.398**		
	Sig. (2-tailed)	.002		
DC_T	Pearson Correlation	.569**	-.229	
	Sig. (2-tailed)	.000	.079	
CN_C	Pearson Correlation	.281*	.603**	-.393**
	Sig. (2-tailed)	.029	.000	.002

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Finally, to test the differences of the four measures of syntactic complexity during the three different stages of life, a one-way MANOVA (SPSS 26) was run. The statistical result reveals the Wilks' λ of 0.865 and significance value of 0.431, suggesting that there are no significant differences between MLS, FVR, DC_T and CN_C when the three different stages of life (early, middle and late adulthood) are compared. Table 8 provides information about the mean scores of the syntactic measures in three different stages.

Table 8

Mean Scores of MLS, FVR, DC_T an CN_C in Three Stages of Adulthood

Mean score	Early adulthood (Text 1-16)	Middle adulthood (Text 17-40)	Late adulthood (Text 41-60)
MLS	27.95	26.99	24.35
FVR	12.46	11.87	11.39
DC_T	1.06	1.09	0.96
CN_C	1.96	1.74	1.63

In order to discover how the lexical and syntactic measures interact over time, several Pearson correlation tests are run. (See Appendix C for detail). Results show that MLS, the general index for syntactic complexity does not have any significant relationships with the variables of lexical complexity. Furthermore, Guiraud is not significantly correlated with any syntactic complexity measure. However, lexical density, the general index for lexical complexity, is found to have significant moderate interactions with FVR (Pearson r=0.277, p=0.032) and CN_C (Pearson r=0.413, p=0.001) respectively. Therefore, it can be concluded that the uses of more lexical words in a sample can be attributable to the uses of complex nominals with greater length and the uses of more non-finite constructions. Another interesting interaction is between CN_C and AWL (Pearson r=0.262, p=0.043), suggesting that when the subject uses more complex nominal constructions per clause, she tends to use more sophisticated words.

Hidden Markov Modelling Output

In this study, we obtained seven measures, 3 for lexical complexity and 4 for syntactic complexity. According to the Pearson correlation tests presented above, Guiraud, AWL and lexical density are independent of one another. In contrast, CN_C and FVR show a supportive relationship and the former is also included in the latter. Therefore, we removed CN_C from the observational parameters when we ran the HMMs. Based on the modelling outputs for separate measures of Guiraud and MLS given below (as well as the required number of texts in one state in proportion to the number of observational parameters), we predicted that the interactions of lexical complexity measures might produce 2 or 3 hidden states, the interactions for syntactic complexity might produce 2 states and the interactions of the six measures of complexity would produce 2 hidden states.

To test for the number of possible states, we first conducted the modelling separately for Guiraud, representative of the lexical system and MLS, representative of the syntactic system, with two different assumptions on the number of state occurrences: 2 and 3. With the 2-state assumption, HMM detected Text 40 as the change point for Guiraud. With the 3-state assumption, the HMM output was that the transition of Guiraud from State 1 to State 2 took place at Text 17 and from State 2 to State 3 at Text 40. Not surprisingly, this modelling output is similar to our observations from Figure 3 where Text 18 and Text 42 were seen as transitional moments. Moreover, these modelling outputs showed a balanced distribution of texts among the two or three developmental states. Turning to the modelling of MLS, with the 3-state assumption, the model detected Text 6 and Text 7 as the two change points, suggesting that State 2 was composed of only 1 text, which does not show an equal distribution of texts among the states. With the two-state assumption, the phase shift was found to occur at Text 37 for the 2-state assumption. This is in line with our earlier observation that around Texts 35-37, new peaks occur, signaling a new stage of development.

When the six measures were put together for the configuration in HMM, the exact point of the phase shift was also determined at datapoint 37 as illustrated in Figure 9 with the average trendlines and the dotted line representing State 1 and State 2 of the whole system. At datapoint 37, the majority of measures seem to converge. This change point coincides with the change point determined by HMM for syntactic complexity measures.

Figure 9

Moving Average Trendlines and Phase Shift of Six Complexity Measures

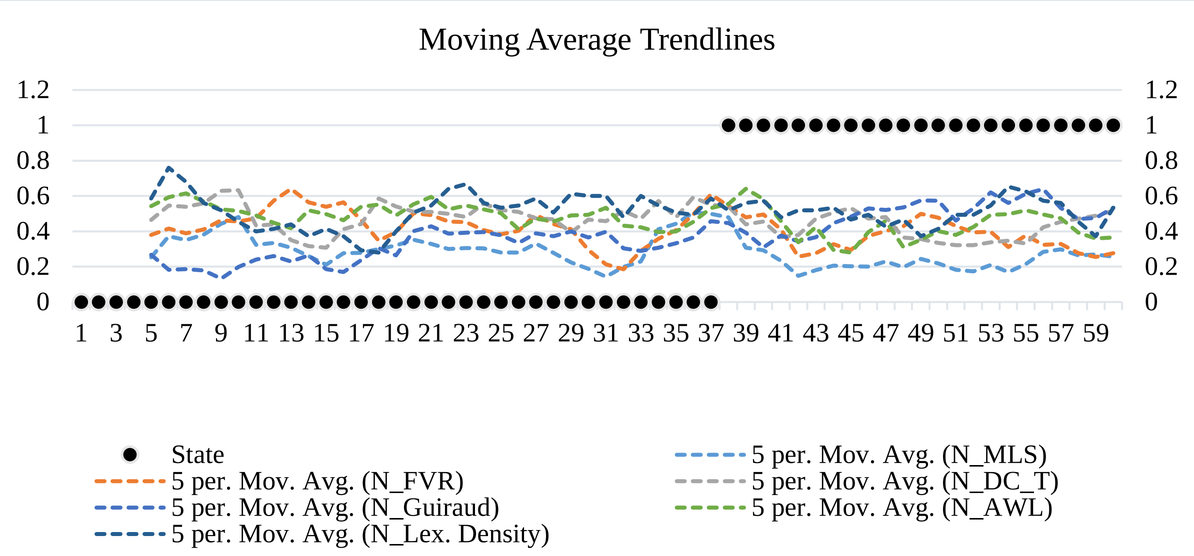


Table 9 presents the mean scores and standard deviations of six complexity measures. Independent samples t-test was employed to compare the differences in the complexity measures between the two HMM's states. It turned out that there was a significant difference in mean values of Guiraud between the 2 newly detected states (p=0.001); the other lexical measures do not show any significant differences. Also, MLS in State 2 was significantly lower than in State 1 (p=0.017). The mean scores of FVR and DC_T did not significantly differ between States 1 and 2.

Table 9

Mean of Lexical and Syntactic Complexity Measures

Mean	Lexical complexity			Syntactic complexity		
State	N_Guiraud	N_AWL	N_Lex. D	N_MLS	N_FVR	N_DC_T
State 1 (Text 1 – Text 37)	0,31	0,51	0,52	0,32	0,44	0,49
State 2 (Text 38 – Text 60)	0,51	0,41	0,51	0,21	0,35	0,42
Difference (State 2-State 1)	0.2 (p=0.001)	-0.1 (p=0.067)	-0.01 (p=0.907)	-0.11 (p=0.017)	-0.09 (p=0.071)	-0.07 (p=0.148)

Discussion

In this paper, we traced the complexity development in 60 texts written by DLF over 50 years to see if there are clear stages in adulthood as ascertained by Levinson (1978). The first text was written when she was 27 and the last when she was 76. Unfortunately, we missed the first 10 years of the early adult stage, probably a time at which she developed her academic writing style the most, so we speculated that this stage would not be clearly visible in our data and thus we might only see two clear stages, if there were any. We carefully selected 200-word samples and coded them for lexical complexity (diversity, sophistication and density) and for syntactic complexity (mean length of sentence, finite verb token ratio, complex nominals per clause, and dependent clauses per T-unit).

We traced these variables over time using traditional CDST techniques and statistics and analyzed them according to the three assumed stages of adulthood. Then we fed the data into an HMM model to see if it could detect meaningful transition phases and compared outcomes of the stages. The overall research question is if meaningful developmental states can be identified in adulthood. We will first summarize the findings according to our research questions before commenting on them and relating them to the literature and discussing the implications.

RQ1: Are there changes in the lexical and syntactic complexity measures over time?

Over time, there are no significant changes in six of the seven complexity measures. The only significant change is found in lexical diversity as calculated with Guiraud on content words.

RQ2: Are there significant peaks in lexical and syntactic complexity measures over time?

Over time, each of the measures fluctuates (waxes and wanes), some with seemingly high peaks and valleys; however, when tested in Monte Carlo simulations, none were found to be significant.

RQ3: Are there significant changes in lexical and syntactic complexity measures between the assumed three stages of adulthood (early, mid and late)?

DLF used significantly more different types of words in her late adulthood than in her early and middle adulthood. There were no differences in any of the syntactic measures.

RQ4: Are there any significant correlations between lexical and syntactic complexity measures over time?

As far as lexical measures are concerned, the Guiraud does not interact with the other two measures across the three stages as it increases over time. As far as syntactic measures are concerned, there are more interactions as they clearly overlap and are interdependent. MLS has a moderate, supportive relationship with DC_T, CN_C and FVR, which means that the number of dependent clauses, the number of complex noun phrases and the measure of overall sentence complexity all contribute to average sentence length. Another expected significant relationship is observed for DC_T and CN_C, which entails that when the subject uses more finite dependent clauses, she tends to use fewer complex noun phrases. Finally, there is a significant strong correlation between FVR and CN_C, which makes sense as the FVR measure subsumes complex nominals.

However, lexical density, the general index for lexical complexity, is found to have significant moderate interactions with FVR and CN_C. Therefore, the use of more lexical words in a sample can be related to the uses of complex nominals with greater length and the uses of more non-finite constructions. Another interesting interaction is between CN_C and AWL, suggesting that when the subject uses more nominal constructions, she tends to use more sophisticated words.

RQ5: What developmental states are identified by HMM modeling?

When running an HMM, the number of states one wants to find needs to be indicated by the researcher. To discover what the best number of states was (2, 3 or 4), we tested different variables in different combinations over time. For the lexical measures, there were clearly 2 states (as the hypothesized 2nd state only covered four texts). For the syntax, there were also clearly 2 states. Thus, we indicated two states for the whole model. When the six measures were put together for the configuration in HMM, the exact point of the phase shift was determined at datapoint 37, when DLF was 63 years old.

RQ6: Are there differences in the measures between HMM's states?

When comparing the complexity measures from State 1 to State 2, we found two significant differences, with the Guiraud increasing significantly and MLS decreasing significantly.

To answer our overall research question, our data definitely rejects Labov's (2001) assumption that language stabilizes in adulthood. In contrast, we should conclude that in our data there are clear stages in adulthood, not exactly as hypothesized by Levinson (1978) with a distinction between early, mid and late adulthood but between mid and late adulthood at age 63 as hypothesized. As mentioned earlier, not finding a shift between early and middle adulthood could very much have been due to the fact that our data set was lacking texts between the age of 17 and 27. We must also confirm that the language of those in late adulthood changes from middle adulthood (Cheshire, 2005), but only in the case of MLS, which is a very broad measure of sentence complexity and includes finite dependent clauses, which contain more words and actually make a sentence less complex. Our data is not in line with earlier findings concerning the lexicon (Kemper et al., 2001) with a growth in vocabulary throughout middle adulthood and a decline in late adulthood (cf. Botwinick & Siegler, 1980; Zelinski & Burnight, 1997) nor with Bowles et al. (2005) who suggest that the optimal level of basic vocabulary is attained at around the age of 30. But in line with CDST thought, learning is iterative and "may lead to an accumulation of vocabulary because people are continuously exposed to language through education, incremental reading or life experience" (Kavé, 2022, p. 2) and "a speaker's lexical diversity is conditioned on the properties of his or her interlocutor, but age and higher levels of education predict increased lexical diversity for individual speakers" (Meylan & Gahl, 2014, p. 1010), which seems to be the case with DLF.

The findings are also in line with Nevid and Rathus (2005) and de Bot (2007), who predicted that major life events may contribute to change. Age 63 was around DLF's partial retirement from teaching, but not from publishing. Considering her publication record, she published as much if not more after age 63, and one might consider the fact that more time may result in time spent on refining her writing. It seems that DLF started avoiding overly long sentences but used a more diversified vocabulary instead, so rather than calling the shift a decline in language skills, the shift is more in style to a more readable and sophisticated one.

As far as research findings are concerned, the detailed tracing of the different complexity measures over time confirms general CDST assumptions. There was fluctuation even in the L1 of an adult, with strong waxing and waning of patterns in all the variables examined. Still, no developmental peaks were found in any of these measures as they were found in advanced L2 writers (cf. Penris & Verspoor, 2017). However, this is not surprising as the first text we measured was right after her graduate schooling in which she probably had developed in her academic writing style early on before she finished her dissertation.

As far as the interaction of the measures is concerned, it was interesting to see that the three lexical measures (density, sophistication and diversity) are quite independent from each other and clearly measure different aspects of lexical development. On the other hand, the syntactic measures overlap to a degree and are more interdependent, but they can clearly reveal the manner in which the language becomes more or less complex.

Finally, when we compare our more typical CDST methods with HMM, we must say they both have their advantages. The typical CDST methods with their clear visualizations can show how different variables develop and interact over time, in what manner sentences become more complex and indicate where approximately possible phase transitions occur, but the HMM is useful in showing empirically when a phase shift emerges, especially in the case of a group of variables. The disadvantage of an HMM is that quite a few datapoints are needed over time, in our case 60.

The limitations of this study are clear. One limitation is the fact that 200-word samples may not reflect the text as a whole, although we did take them from similar parts of the texts (introduction and literature review). Another limitation is that we conducted an individual case study, and we cannot generalize beyond the individual. Repeating this analysis with a few additional L1 scholars would be helpful in detecting any common developmental patterns that might exist. Still, we feel that we did adduce evidence from this case study

Conclusion

Several conclusions can be drawn from the study. First of all, our findings support Baltes et al. (1980) on lifespan development: developmental processes are non-linear but multi-dimensional and multi-directional. We have found that language development is not characterized by a series of discrete stages, but rather the ebbing and flowing of patterns. Also, we have adduced evidence that developmental decline is not inevitable, as the increase in lexical diversity in this writer demonstrates. Actually, even though there was statistical evidence for phases in the development of her writing, we might infer that DLF's academic language use did not decline in complexity but was relatively stable in that respect. However, the stability was the product of fluctuating variables. As Prigogine (1978) demonstrated, order appears through fluctuations.

ORCID

https://orcid.org/0009-0007-1802-7964

https://orcid.org/0000-0001-7442-2693

https://orcid.org/0000-0001-8828-8405

Acknowledgements

Not applicable.

Funding

Not applicable.

Ethics Declarations

Competing Interests

No, there are no conflicting interests.

Rights and Permissions

Open Access

This article is licensed under a Creative Commons Attribution 4.0 International License, which grants permission to use, share, adapt, distribute and reproduce in any medium or format provided that proper credit is given to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if any changes were made.

References

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

Baltes, P. B., Reese, H., & Lipsert, L. (1980). Lifespan developmental psychology. *Annual Review of Psychology, 31*, 65-110.

Bassano, D., & van Geert, P. (2007). Modeling continuity and discontinuity in utterance length: A quantitative approach to changes, transitions and intra-individual variability in early grammatical development. *Developmental Science*, *10*(5), 588-612. https://doi.org/10.1111/j.1467-7687.2007.00629.x

Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes*, *9*(1), 2-20. https://doi.org/10.1016/j.jeap.2010.01.001

Birren, J.E. (1964). *The psychology of aging*. Prentice-Hall.

Botwinick, J., & Siegler, I. C. (1980). Intellectual ability among the elderly simultaneous cross-sectional and longitudinal comparisons. *Developmental Psychology*, *16*(1), 49-53. https://doi.org/10.1037//0012-1649.16.1.49

Bowles, R. P., Grimm, K. J., & McArdle, J. J. (2005). A structural factor analysis of vocabulary knowledge and relations to age. *Journals of Gerontology- Series B-Psychological Sciences and Social Sciences*, *60*(5), 234-241. https://doi.org/10.1093/geronb/60.5.P234

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. https://doi.org/10.1590/s1984-63982013005000002

Chan, H., Verspoor, M., & Vahtrick, L. (2015). Dynamic development in speaking versus writing in identical twins. *Language Learning*, *65*(2), 298-325. https://doi.org/10.1111/lang.12107

Cheshire, J. (2005) Age and generation-specific use of language. In U. Ammon, N. Dittmar, K. Mattheier & P. Trudgill (Eds.), *Sociolinguistics: An introductory handbook of the science of language and society (*pp. 1552-1563). Mouton de Gruyter.

Cheung, H., & Kemper, S. (1992). Competing complexity metrics and adults' production of complex sentences. *Applied Psycholinguistics*, *13*(1), 53-76. https://doi.org/10.1017/s0142716400005427

de Bot, K. (2007). Dynamic systems theory, lifespan development and language attrition. In B. Kopke, M. S. Schmid, M. Keijzer & S. Dostert (Eds.), *Language attrition: Theoretical perspectives* (pp. 53-68). John Benjamins.

de Bot, K., & Schrauf, W. R. (Eds.). (2009). *Language development over the lifespan*. Routledge.

de Bot, K., Lowie, W., & Verspoor, M. (2005). *Second language acquisition: An advanced resource book.* Routledge.

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

Ellis, N. C., & Larsen-Freeman, D. (2009). Constructing a second language: Analyses and computational simulations of the emergence of linguistic constructions from usage. *Language Learning, 59, 90–125.* https://doi.org/10.1111/j.1467-9922.2009.00537.x

Fischer, K., Yan, Z., & Steward, J. (2003). Adult cognitive development: Dynamics in the developmental web. In J. Valsiner, & K. Connolly (Eds.), *Handbook of developmental psychology* (pp. 491-516). Sage.

Grant, L., & Ginther, A. (2000). Using computer-tagged linguistics features to describe L2 writing differences. *Journal of Second Language Writing*, *9*, 123-145. https://doi.org/10.1016/S1060-3743(00)00019-9

Hankin, B. L., & Abramson, L. Y. (2001). Development of gender differences in depression: An elaborated cognitive vulnerability-transactional stress theory. *Psychological Bulletin, 127*(6), 773-796. https://doi.org/10.1037//0033-2909.127.6.773

Hmmlearn Developers. (n.d.). Tutorial. https://hmmlearn.readthedocs.io/en/latest/tutorial.html

Hood, G. M. (2004). PopTools Version 2.6.2 [Software]. CSIRO.
http://www.cse.csiro.au/poptools.

Kavé, G. (2022). Vocabulary changes in adulthood: Main findings and methodological considerations. *International Journal of Language and Communication Disorders*, *October*, 1–10. https://doi.org/10.1111/1460-6984.12820

Kemper, S., Thompson, M., & Marquis, J. (2001). Longitudinal change in language production: Effects of aging and dementia on grammatical complexity and propositional content. *Psychology and Aging, 16*(4), 600-614. https://doi.org/10.1037//0882-7974.16.4.600

Kyle, K. (2016). *Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication* (Doctoral Dissertation). http://scholarworks.gsu.edu/alesl_diss/35.

Labov, W. (1994). *Principles of linguistic change: Internal factors*. Wiley.

Labov, W. (2001). *Principles of linguistic change: Social factors.* Blackwell Publishers.

Labov, W. (2007). Transmission and diffusion. *Language*, *83*(2), 344-387. https://doi.org/10.1353/lan.2007.0082

Lankshear, C., & Knobel, M. (2004). *A handbook for teacher research: From design to implementation*. Open University Press.

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics, 18*(2), 141-165. 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*(4), 590-619. https://doi.org/10.1093/applin/aml029

Lenneberg, E.H. (1967). The biological foundations of language. *Hospital Practice*, *2*(12), 59-67, https://doi.org/10.1080/21548331.1967.11707799

Levinson, D. (1978). *The seasons of a man's life*. Knorf.

Lewis, M. D. (2000). The promise of Dynamic Systems Approaches for an integrated account of human development. *Child Development, 71,* 36–43. https://doi.org/10.1111/1467-8624.00116

Lowie, W., Verspoor, M., & De Bot, K. (2009). A dynamic view of second language development across the lifespan. In K. De Bot, & W. R. Schrauf (Eds.), *Language development over the lifespan* (pp. 125-145). Routledge.

Lowie, W. M., Caspi, T., van Geert, P., & Steenbeek, H. (2011). Modeling development and change. In M. H. Verspoor, K. de Bot, & W. M. Lowie (Eds.)*, A dynamic approach to second language development: methods and techniques* (pp. 99 - 121). John Benjamins.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, *15*(4), 474-496. https://doi.org/10.1075/ijcl.15.4.02lu

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, *45*(1), 36–62. https://doi.org/10.5054/tq.2011.240859

Meylan, S., & Gahl, S. (2014). The divergent lexicon: Lexical overlap decreases with age in a large corpus of conversational speech. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36. l1006-1011. https://escholarship.org/uc/item/4s3757m5

Nevid, J. S., & Rathus, S. A. (2005). *Psychology and the challenges of life: Adjustment in the new millennium* (9th ed.). John Wiley & Sons.

Nippold, M. A., Ward-Lonergan, J. M., & Fanning, J. L. (2005). Persuasive writing in children, adolescents, and adults: A study of syntactic, semantic and pragmatic development. *Language Speech and Hearing Services in Schools*, *36*(2), 125-138. https://doi.org/10.1044/0161-1461(2005/012)

Penris, W. & Verspoor, M. (2017). Academic writing development: A complex, dynamic process. In S. Pfenninger & J. Navracsics (Ed.), *Future research directions for applied linguistics* (pp. 215-242). Multilingual Matters. https://doi.org/10.21832/9781783097135-012

Prigogine, I. (1978). Time, structure, and fluctuations. *Science, 201*(4358), 777-785.

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Harald, B. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in Cognitive Science*, *6*, 5–42. https://doi.org/10.1111/tops.12078

Stamp, M. (2017). *Introduction to machine learning with applications in information security* (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315213262

Tagliamonte, S. A., & D'Arcy, A. (2009). Peaks beyond phonology: Adolescence, incrementation, and language change. *Language, 85*(1), 58-108. https://doi.org/10.1353/lan.0.0084

Ure, J. (1971). Lexical density and register differentiation. In G. E. Perren & J. L. M. Trim (Eds.), *Applications of linguistics* (pp. 443-452). Cambridge University Press.

van Dijk, M., Verspoor, M., & Lowie, W. (2011). Variability and DST. In M. Verspoor, K. de Bot, & W. Lowie (Eds.), *A dynamic approach to second language development* (pp. 55–84). John Benjamins.

van Geert, P. (1991). A dynamic systems theory model of cognitive and language growth. *Psychological Review, 98*, 3-53. https://doi.org/10.1037/0033-295X.98.1.3

van Geert, P. (1994). *Dynamic systems of development: Change between complexity and chaos*. Harvester Wheatsheaf.

van Geert, P. (2008). The dynamic systems approach in the study of L1 and L2 acquisition: An introduction. *The Modern Language Journal*, *92*(2), 179-199. https://doi.org/10.1111/j.1540-4781.2008.00713.x

van Geert, P., & van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior & Development, 25*(4), 340-374. https://doi.org/10.1016/S0163-6383(02)00140-6

Verspoor, M., Lowie, W., & de Bot, K. (2011). *A dynamic approach to second language development: Methods and techniques*. Benjamins. https://doi.org/10.1075/lllt.29

Verspoor, M., Lowie, W., & Van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. *The Modern Language Journal, 92*(2), 214-231. https://doi.org/10.1111/j.1540-4781.2008.00715.x

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

Wolf-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity*. University of Hawai'i, Second Language Teaching & Curriculum Center.

Writewords. (n.d.). *Word Frequency Counter* [computer software]. http://www.writewords.org.uk/word_count.asp

Zeeman, E. C. (1976). Catastrophe theory. *Scientific American*, *234*(4), 65–83.

Zelinski, E. M., & Burnight, K. P. (1997). Sixteen-year longitudinal and time lag changes in memory and cognition in older adults. *Psychology and Aging, 12*(3), 503-513. https://doi.org/10.1037/0882-7974.12.3.503

Appendix A

List of Articles and Book Chapters by Diane Larsen-Freeman used for Data Analysis

Larsen-Freeman, D. (1974). A re-examination of grammatical structure sequencing. In R. Crymes & W. Norris (eds.). *On TESOL '74* (pp.83-93). TESOL.

Larsen-Freeman, D. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly*, *9*(4), 409. https://doi.org/10.2307/3585625.

Larsen-Freeman, D. E. (1976a). An explanation for the morpheme acquisition order of second language learners. *Language Learning*, *26*(1), 125–134. https://doi.org/10.1111/j.1467-1770.1976.tb00264.x

Larsen-Freeman, D. E. (1976b). *Evidence of the need for a second language acquisition index of development*. ERIC Clearinghouse.

Larsen-Freeman, D. (1977). A rationale for discourse analysis in second language acquisition research. In H. Brown, C. Yorio& R. Crymes (Eds.), *On TESOL '77: Teaching and learning English as a second language: Trends in research and practice* (pp.172-177). TESOL.

Larsen-Freeman, D. (1978a). An ESL index of development. *TESOL Quarterly*, *12*(4), 439. https://doi.org/10.2307/3586142

Larsen-Freeman, D. (1978b). Implications of the morpheme studies for second language acquisition. *International Journal of Applied Linguistics, 39*(1), 93–102. https://doi.org/10.1075/itl.39-40.06lar.

Larsen-Freeman, D. (1981). The WHAT of SLA. In A. Hines & W. E. Rutherford (Eds.), *TESOL '81: Selected papers from the fifteenth annual conference of TESOL.* TESOL.

Larsen-Freeman, D. (1983a). Training teachers or educating a teacher. In J.E. Alatis, H.H. Stern & P. Strevens (Eds.), *Georgetown University roundtable on languages and linguistics 1983* (pp. 264-274). Georgetown University Press.

Larsen-Freeman, D. (1983b). Assessing global second language proficiency. In H. Seliger & M. Long (Eds.), *Classroom oriented research in second language acquisition*. Newbury House.

Larsen-Freeman, D. (1985a). State of the art on input in second language acquisition. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 433-444). Newbury House.

Larsen-Freeman, D. (1985b). Considerations in research design in second language acquisition. In M. Celce-Murcia (Ed.), *Beyond basics: Issues and research in TESOL* (pp.125-136) Newbury House.

Larsen-Freeman, D. (1987a). Recent innovations in language teaching methodology. *The ANNALS of the American Academy of Political and Social Science*, *490*(1), 51–69. https://doi.org/10.1177/0002716287490001005

Larsen-Freeman, D. (1987b). Comments on Leo A.W. van Lier's review of "Techniques and principles in language teaching": The author reacts. *TESOL Quarterly*, *21*(4), 769. https://doi.org/10.2307/3586995

Larsen-Freeman, D. (1989). Pedagogical descriptions of language: grammar. *Annual Review of Applied Linguistics*, *10*, 187-195. https://doi.org/10.1017/S026719050000129X.

Larsen-Freeman, D. (1990). On the need for a theory of language teaching. In J.E. Alatis (Ed.), *Georgetown University roundtable on languages and linguistics 1990: Linguistics, language teaching and language acquisition: The interdependence of theory, practice and research* (pp. 261-270). Georgetown University Press.

Larsen-Freeman, D. (1991a). Research on language teaching methodologies: A review of the past and an agenda for the future. In K. de Bot, R. Ginsberg & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp.119-132). John Benjamins.

Larsen-Freeman, D. (1991b).Consensus and divergence on the content, role, and process of teaching grammar. In J.E. Alatis (Ed.), *Georgetown University round table on languages and linguistics 1991: Linguistics and language pedagogy: The state of the art* (pp. 260-273). Georgetown University Press.

Larsen-Freeman, D. (1991c). Second language acquisition research: Staking out the territory. *TESOL Quarterly*, *25*(2), 315. https://doi.org/10.2307/3587466

Larsen-Freeman, D. (1992). A nonhierarchical relationship between grammar and communication. In J.E. Alatis (Ed.), *Georgetown University roundtable on languages and linguistics (GURT) 1992: Language, communication, and social meaning* (pp. 158-173). Georgetown University Press.

Larsen-Freeman, D. (1997b). *Grammar and its teaching: Challenging the myths*. ERIC Clearinghouse. -.

Larsen-Freeman, D. (1997b). Chaos/complexity science and second language acquisition. *Applied Linguistics*, *18*(2), 140–165. https://doi.org/10.1093/applin/18.2.141

Larsen-Freeman, D. (1998). On the scope of second language acquisition research: "The learner variety" perspective and beyond -A response to Klein. *Language Learning*, *48*(4), 551–556. https://doi.org/10.1111/0023-8333.00058

Larsen-Freeman, D. (2000). Second language acquisition and applied linguistics. *Annual Review of Applied Linguistics*, *20*, 165–181. https://doi.org/10.1017/s026719050020010x

Larsen-Freeman, D. (2001). Individual cognitive/affective learner contributions and differential success in second language acquisition. In M. Breen (Ed.), *Learner contributions to language learning: New directions in research* (1st ed., pp. 12-43). Pearson Education.
Larsen-Freeman, D. (2002a). Making sense of frequency. *Studies in Second Language Acquisition*, *24*(2), 275–285. http://www.jstor.org/stable/44486619
Larsen-Freeman, D. (2002b). Language acquisition and language use from a chaos/complexity theory perspective. In C. Kramsch (Ed.), *Language acquisition and language socialization* (pp. 33–46). Continuum.
Larsen-Freeman, D. (2004a). Reflections on form-meaning connection research in second language acquisition. In J. Williams, B. VanPatten, S. Rott, & M. Overstreet (Eds.), *Form-meaning connections in second language acquisition* (pp.251-260). Lawrence Erlbaum.
Larsen-Freeman, D. (2004b). CA for SLA? It all depends. *The Modern Language Journal*, *88*(4), 603–607. http://www.jstor.org/stable/3588590
Larsen-Freeman, D. (2005). SLA and issue of fossilization: There is no end, and there is no state. In Z-H. Han & T. Odlin (Eds.), *Studies of fossilization in second language acquisition* (pp. 189-200). Multilingual Matters.
Larsen-Freeman, D. (2006a). Functional grammar: On the value and limitations of dependability, inference, and generalizability. In M. Chalhoub-Deville, C. Chapelle, & P.A. Duff (Eds.), *Inference and generalizability in applied linguistics: Multiple perspectives* (pp. 115-133). John Benjamins.
Larsen-Freeman, D. (2006b). On the need for a new understanding of language and its development. *Journal of Applied Linguistics and Professional Practice*, *3*(3). https://doi.org/10.1558/japl.v3.i3.281
Larsen-Freeman, D. (2006c). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, *27*(4), 590–619. https://doi.org/10.1093/applin/aml029
Larsen-Freeman, D. (2007a). Reflecting on the cognitive-social debate in second language acquisition. *The Modern Language Journal*, *91*, 773–787. http://www.jstor.org/stable/4626131
Larsen-Freeman, D. (2007b). On the complementarity of chaos/complexity theory and dynamic systems theory in understanding the second language acquisition process. *Bilingualism*, *10*(1), 35–37. https://doi.org/10.1017/S136672890600277X
Larsen-Freeman, D. (2009a). Teaching and testing grammar. *The Handbook of Language Teaching*, *3*, 518–542. https://doi.org/10.1002/9781444315783.ch27
Larsen-Freeman, D. (2009b). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, *30*(4), 579–589. https://doi.org/10.1093/applin/amp043
Larsen-Freeman, D. (2010a). Not so fast: A discussion of l2 morpheme processing and acquisition. *Language Learning*, *60*(1), 221–230. https://doi.org/10.1111/j.1467-9922.2009.00556.x
Larsen-Freeman, D. (2010b). Teaching and learning English: From ideology to empowerment. *Journal of NELTA*, *12*(1). https://doi.org/10.3126/nelta.v12i1.3430
Larsen-Freeman, D. (2012a). Complex, dynamic systems: A new transdisciplinary theme for applied linguistics? *Language Teaching*, *45*(2), 202–214. https://doi.org/10.1017/S0261444811000061
Larsen-Freeman, D. (2012b). On the roles of repetition in language teaching and learning. *Applied Linguistics Review*, *3*(2), 195–210. https://doi.org/10.1515/applirev-2012-0009
Larsen-Freeman, D. (2012c). The emancipation of the language. *Studies in Second Language Learning and Teaching*, *2*(3), 297–309.
Larsen-Freeman, D. (2013). Transfer of Learning Transformed. *Language Learning*, *63*(SUPPL. 1), 107–129. https://doi.org/10.1111/j.1467-9922.2012.00740.x
Larsen-Freeman, D. (2014a). Saying what we mean: Making a case for “language acquisition” to become “language development.” *Language Teaching*, *48*(4), 491–505. https://doi.org/10.1017/S0261444814000019
Larsen-Freeman, D. (2014b). Ten ‘lessons’ from complex dynamic systems theory: What is on offer. In *Teaching English as a second of foreign language* (pp. 256–270).
Larsen-Freeman, D. (2015). Research into practice: Grammar learning and teaching. *Language Teaching*, *48*(2), 263–280. https://doi.org/10.1017/S0261444814000408.
Larsen-Freeman, D. (2016a). Complexity theory and ELF: A matter of nonteleology. *English as a Lingua Franca: Perspectives and Prospects*. https://doi.org/10.1515/9781501503177-021.
Larsen-Freeman, D. (2016b). Classroom-oriented research from a complex systems perspective. *Studies in Second Language Learning and Teaching*, *6*(3), 377–393. https://doi.org/10.14746/ssllt.2016.6.3.2
Larsen-Freeman, D. (2017a). Just learning. *Language Teaching*, *50*(3), 425–437. https://doi.org/10.1017/S0261444817000106.
Larsen-Freeman, D. (2017b). Complexity theory: The lessons continue. In L. Ortega & Z. Han (Eds.). *Complexity theory and language development: In celebration of Diane Larsen-Freeman* (pp. 11-50). John Benjamins. https://doi.org/10.1075/lllt.48.02lar.
Larsen-Freeman, D. (2018a). Looking ahead: Future directions in, and future research into, second language acquisition. *Foreign Language Annals*, *51*(1), 55–72. https://doi.org/10.1111/flan.12314

Larsen-Freeman, D. (2018b). Second language acquisition, WE, and language as a complex adaptive system (CAS). *World Englishes*, *37*(1), 80–92. https://doi.org/10.1111/weng.12304.

Larsen-Freeman, D. (2018c). Task repetition or task iteration? It does make a difference. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 311-329), John Benjamins. https://doi.org/10.1075/tblt.11.12lar

Larsen–Freeman, D. (2019a). On language learner agency: A complex dynamic systems theory perspective. *Modern Language Journal*, *103*, 61–79. https://doi.org/10.1111/modl.12536

Larsen-Freeman, D. (2019b). Second language development in its time: Expanding our scope of inquiry. *Chinese Journal of Applied Linguistics*, *42*(3), 267–284. https://doi.org/10.1515/CJAL-2019-0017.

Larsen-Freeman, D. (2019c). Teaching and researching grammar skills: Theory- and research- based practices. In N. Polat, T. Gregersen & P.D. MacIntyre (Eds.), *Research-driven pedagogy: Implications of L2A theory and research for the teaching of language skills* (PP. 97-124.). Routledge. https://doi.org/10.4324/9781351043281.

Larsen-Freeman, D. (2020). Chaos/complexity theory for second language acquisition/development. *The Encyclopedia of Applied Linguistics*, 1–8. https://doi.org/10.1002/9781405198431.wbeal0125.pub2

Larsen–Freeman, D. (2022). Combinations and connections: Reaching across disciplinary boundaries. *The Modern Language Journal*, *106*, 132–140. https://doi.org/10.1111/modl.12753

Thanh T.G. Trinh. (2011). *An adult's language variability and development*. Master's thesis, University of Groningen.

Appendix B

Comparisons between Book Chapters and Journals

Table 1

Differences in Lexical and Sentence Complexity Measures between Book Chapters and Journals

Measures	p-value	Mean score difference
Lexical sophistication (AWL)	0.611 (>0.05)	-0.512
Lexical density (D)	0.462 (>0.05)	0.740
Lexical diversity (Guiraud)	0.168 (>0.05)	1.397
Mean length of sentence (MLS)	0.372 (>0.05)	-0.900
Sentence sophistication (FVR)	0.190 (>0.05)	-1.325
Dependent clauses per T_unit (DC_T)	0.932 (>0.05)	0.085
Complex nominals per clause (CN_C)	0.804 (>0.05)	-0.249

APPENDIX C

Pearson R Correlations for Complexity Variables

Correlations

		MLS	FVR	DC_T	CN_C	GUIRAUD	AWL	DENSITY
MLS	Pearson Correlation	1						
	Sig. (2-tailed)							
FVR	Pearson Correlation	.398**	1					
	Sig. (2-tailed)	.002						
DC_T	Pearson Correlation	.569**	-.229	1				
	Sig. (2-tailed)	.000	.079					
CN_C	Pearson Correlation	.281*	.603**	-.393**	1			
	Sig. (2-tailed)	.029	.000	.002				
GUIRAUD	Pearson Correlation	-.056	.101	-.113	.074	1		
	Sig. (2-tailed)	.669	.442	.391	.576			
AWL	Pearson Correlation	.069	.113	.027	.262*	.063	1	
	Sig. (2-tailed)	.603	.391	.837	.043	.633		
DENSITY	Pearson Correlation	-.038	.277*	-.228	.413**	.015	.197	1
	Sig. (2-tailed)	.771	.032	.080	.001	.911	.131	

***. Correlation is significant at the 0.01 level (2-tailed).*
**. Correlation is significant at the 0.05 level (2-tailed).*