An exploratory longitudinal case study on the interactions of L2 fluency, lexical, and syntactic development in oral narratives during a study abroad program.

Teruaki Tsushima

ABSTRACT

This study aimed to explore the interactions between fluency, lexical complexity, and syntactic complexity in oral narratives produced by a Japanese adult learner of English during a two-year study abroad program in the Complex Dynamic Systems Theory (CDST) framework. A total of 173 oral narratives were analyzed, focusing on fluency (e.g., speech rate), syntactic complexity (e.g., complex nominals per T-unit), and lexical complexity (e.g., corrected type token ratio). The findings revealed significant improvements in fluency and syntactic complexity from the initial stage to the study abroad period. However, only two out of six measures of lexical complexity showed significant improvement. Notable change points were detected in the fluency and syntactic complexity indices, with several indices displaying phase shifts at the end of intensive speech training, leading to overall improvement. Similar phase shifts were observed in the fluency and lexical complexity indices during the first half of the study abroad program. Additionally, the study found that the detrended correlations between fluency and syntactic complexity were much higher in the program's second year, indicating a supportive relationship as the participant's proficiency improved. Moving correlation analysis between fluency and syntactic complexity revealed an alternating pattern of positive and negative correlations in the first year, shifting to mostly positive correlations during the intensive speech training and study-abroad period. Regarding the interactions between fluency and lexical complexity, a similar alternating pattern was evident in the first year, followed by predominantly positive correlations during the intensive training phase. However, these correlations mostly remained negative throughout the period of study abroad. The implications of these results for teaching practices are discussed within the framework of CDST.

keywords: L2 learning, fluency, lexicon, syntax, study abroad, longitudinal

1. Introduction

1-1. Review of literature

Previous research in L2 (second or foreign language) acquisition has revealed much individual variability in learning a target language. One of the theories that address this variability is the Complex Dynamic System Theory (CDST), which highlights the complexity and non-linearity inherent in second or foreign-language learning (Fogal, 2022; Hiver, Al-Hoorie, & Evans, 2022; Larsen-Freeman, 1994, October, 2009). Several key characteristics of CDST are particularly relevant to the present study (see Hiver et al., 2022 for a review). First, changes in one aspect of the L2 system affect all other aspects, indicating that L2 language components are interconnected throughout the learning process. Second, the developmental process includes shifts between stable states, known as "attractors." These shifts are usually associated with changes across multiple components of the language system. Moreover, various L2 components——such as vocabulary, phonology, syntax, and discourse——interact dynamically, sometimes competing with or supporting one another. Three patterns of interaction have been identified in this context (van Geert, 2008). The first pattern, "competitive relations," describes interactions between components that vie for limited resources. In such cases, the development of one component may hinder the development of another. The second pattern, "supportive relations," occurs when components enhance each other's functioning, contributing to the overall stability and growth of the L2 system. Finally, "precursor relations" refer to the dependencies between components, where the development of one component is essential for the subsequent development of others.

In this theoretical framework, previous research has extensively examined variability in L2 (e.g., Bulté & Housen, 2018; Larsen-Freeman, 1997; Spoelman & Verspoor, 2010). It has been argued that variability is not merely noise but a significant source of information regarding the L2 developmental process. For instance, variability plays a crucial role in the iterative process through which specific components of L2 skills significantly improve. Studies have shown that relatively large variability in a particular ability can be followed by a dramatic increase in L2 learners' performance (Baba & Nitta, 2014). From this perspective, many previous studies have primarily focused on the developmental processes of L2 learning using a longitudinal research design, whereby a group of participants or a few individuals are observed over an extended period (Baba & Nitta, 2014; Spoelman & Verspoor, 2010; M. Verspoor & Behrens, 2011). For example, Baba and Nitta (2014) investigated how two Japanese university students improved their writing fluency over a single academic year. They found that increased variability in writing fluency was associated with sudden phase shifts in performance (e.g., Spoelman & Verspoor, 2010). A number of subsequent studies explored various aspects of writing development (e.g., Kyle, Crossley, & Verspoor, 2021 on syntactic complexity measures; M. Verspoor, 2017 on syntactic and lexical development).

While many studies have focused on second language (L2) writing development within this research framework, there is a lack of research on L2 speech development (Chan, Verspoor, & Vahtrick, 2015; Lu, 2012; Polat & Kim, 2013). Polat and Kim (2013) investigated the accuracy, syntactic complexity (including mean length of utterance, subordination, and phrasal elaboration), and lexical variability (measured by *D*) in the speech of a Turkish learner of English over one year. Their findings indicated that only lexical variability showed significant improvement. In another study, Chan, Verspoor, and Vahtrick (2015) examined the development of syntactic complexity in both speaking and writing among Taiwanese twins learning English in secondary school. They found that, in the initial stages of development, the twins demonstrated more advanced syntactic complexity (measured by mean length of T-unit, dependent clauses per T-unit, and coordinated clauses per T-unit) in their oral speech compared to writing. However, over time, this trend reversed.

To the best of the author's knowledge, little research has been done on how fluency, syntax, and vocabulary interact in the longitudinal development of L2 speech. To address this gap, the present study examined the development of fluency, syntax, and vocabulary in the speech production of an adult Japanese learner of English over a two-year period, which included a five-month study abroad period.

Previous longitudinal research on the effects of studying abroad on speech development found that L2 learners significantly improved their oral fluency during the study abroad period (Freed, 1995; Towell, Hawkins, & Bazergui, 1996; Valls-Ferrer & Mora, 2014). However, longitudinal studies focused on individual differences found variability in the degree and timing of fluency improvement during the study-abroad period (Tsushima, 2019, 2023). Tsushima (2019) examined fluency improvements among two Japanese English learners who participated in a study abroad program. They recorded spontaneously produced narratives from the participants, collecting 272 and 284 samples over 24 and 20 months, respectively. The results indicated that one student showed significant improvement across most fluency measures, including composite measures (i.e., speech rate and mean length of runs), a speed measure (i.e., articulation rate), and breakdown measures (i.e., frequency and duration of pauses). Conversely, the other student did not show much improvement in any fluency measures during the study-abroad period but did improve significantly in most measures after returning from abroad. A recent study examined the holistic proficiency and syntactic/lexical complexity measures in the writings of a group of advanced sojourners (N=26) during a semester-long study abroad, using weekly diary entries as data (Köylü, Eryılmaz, & Pérez-Vidal, 2023). The results revealed considerable individual variability in holistic proficiency and most syntactic and lexical measures. Although participants showed significant improvement in holistic proficiency, there was a notable decrease in verb phrase complexity and lexical diversity. To the author's knowledge, there has been limited research on the development of syntactic and lexical complexity in speech production during study abroad experiences (see Borràs & Llanes, 2021 for a review).

The present study attempted to provide new data on the development of fluency, syntactic complexity, and lexical complexity in a university student learning English in a study-abroad program where students go to study abroad in Australia for approximately 5.5 months in the latter half of the second year of the program. The purposes of the study were the following. First, it attempted to investigate how fluency, syntactic complexity, and lexical complexity changed significantly throughout the program. Second, it aimed to identify at which specific points, if any, these measures exhibited significant changes and how the timing of these changes varied among the different measures. Finally, the study examined the relationships and interactions between fluency, syntactic complexity, and lexical complexity during the program. For the analysis of these interactions, the focus was limited to the relationships between 1) fluency and syntactic complexity and 2) fluency and lexical complexity due to space constraints, although the relationship between lexical and syntactic complexity was of great importance.

1-2. Specific Research Questions

1) How did fluency, lexical complexity, and syntactic complexity indices change

during the study abroad program?

2) At which points in development did the indices for fluency, syntactic complexity, and lexical complexity show significant changes? How did the timing of these changes differ among the various indices?
3) How were the fluency, syntactic complexity, and lexical complexity indices correlated and interacted during the program?
4) How did the composite fluency indices (i.e., SRP (speech rate (pruned) and MLoR (mean length of runs)) interact with the syntactic complexity and lexical complexity indices?

2. Method

2-1. Participant

The participant (P, henceforth) was a female student at a private university in Tokyo, majoring in business administration. She joined the study-abroad program at the start of her university curriculum. P was monolingual in Japanese and had never studied abroad before this program. According to her TOEIC scores, her initial English proficiency level was assessed as Upper A1 on the CEFR scale at the beginning of university, and she achieved Upper B2 by the end of the current study. She was highly motivated to improve her English skills, mainly speaking and pronunciation.

P attended general English communication classes twice a week during the program's first year (Stages 1 and 2, as shown in Table 2; from 0 to 11 months). She also participated in two program-specific English classes twice a week. In the first semester of the second year (Stage 3; from 12 to 15 months), she took a conversation practice class taught by native English instructors every weekday, which provided intensive speech training. Following this, she attended various English classes every weekday morning during the study-abroad period (Stage 4; from 16 to 20 months). During this time, she lived with a homestay family and joined a sports club, offering her ample opportunities to converse in English. After the study-abroad period, she continued with English classes taught by native English instructors and participated in classes that prepared her for presentations in English (Stage 5; from 21 to 23 months).

2-2. Data acquisition

2-2-1. Speech Data

Speech data were collected by requesting P to produce a monologic narrative lasting approximately one minute about a daily event in her life. She was given less than two minutes to prepare for the speech. The topics varied, covering daily events, classroom activities, trip memories, hobbies, and so forth. In total, 186 narratives were recorded over the course of two years. The first set of narratives (N=12) was recorded weekly during the first semester of the first year in an English class for the study-abroad program. The remaining narratives (N=174) were recorded at her home from July 2022 to March 2024. P was instructed to make a recording at least eight times a month, and she averaged 8.3 recordings per month. From these recordings, the initial set of 12 narratives and an additional 161 (about eight narratives per month) were submitted for analysis.

The recordings were made in the university classes using a high-quality microphone (*AT4040*) attached to a PCM recorder (*DR-44WL*). At home, P was asked to make a recording in a quiet environment, using an iPhone with a high-quality microphone (*Zoom IQ7*) with a pop filter attached to it and recording software (*Zoom Handy Recorder*) with a sampling rate of 44,000 Hz and 16 bits of resolution. The recorded file was sent to the author via email. The sound file was denoised, low pass filtered at 8,000 Hz, and normalized for average intensity at 70 *dB* on the sound analysis software *Praat* (Boersma & Weenink, 2014). Using a *Praat* script (de Jong & Wempe, 2009), the sound waves were first automatically segmented into silent (i.e., pause) and non-silent portions. Then, the vowel boundaries were manually segmented, including the nasal and vowel boundaries, the semivowel and vowel boundaries, and the liquid and the vowel boundaries. Following the previous research, the pause was defined as a silent period of 250ms and longer (e.g., Saito, Ilkan, Magne, Tran, & Suzuki, 2018; Tavakoli, Nakatsuhara, & Hunter, 2020). Then, the repair portions (e.g., repetition, false starts) and Japanese portions (where Japanese words were produced) were coded in separate software tiers and excluded from the calculation of the fluency indices. The sentences that started 60 seconds from the beginning were excluded from the analysis to set a specific limit to the amount of manual segmentation work described above.

2-2-2. Text Data

The speech data was initially transcribed using a transcription program written in *Python*. Afterward, the author manually reviewed and corrected the transcribed texts. To ensure consistency between the speech and text data, the sentences that began 60 seconds from the beginning in the oral data were excluded. The texts were submitted to the automatic Lexical Complexity Analyzer (LCA) (Spring & Johnson, 2022), which is a modified version of Lu's original program (Lu, 2012), and the automatic Syntactic Complexity Analyzer (L2SCA) (Lu, 2010, 2011; Lu & Ai, 2015). Following Spring's (2022) advice, the reported LCA results were based on those using SPACY as a natural language processing toolkit.

2-3. Analysis procedure

2-3-1. Fluency Indices

Following the previous research (Tavakoli & Wright, 2020), the following fluency measures were used in the present study.

1) Composite measures

SRP (Speech Rate (pruned)): the total number of syllables produced in a narrative divided by the amount of total time required to produce it (including pause time) expressed in minutes. The portion of repetitions, false starts, and filled pauses were excluded from the total time.

MLoR (Mean Length of Runs): the average number of syllables produced in utterances between pauses of 250ms and above.

2) Speed measures

AR (Articulation Rate): the total number of syllables produced in a narrative divided by the amount of time taken to produce them (excluding pause time) expressed in minutes.

3) Breakdown measures

3-1) Between-clause pauses: Example: I live in Tokyo // and // I study at a university. This includes all the pauses that take place from the end of the clause and the beginning of the following clause that starts with its subject.

◦ **BCPauseFreq** (frequency of the between-clause pauses per 100 syllables)

◦ **BCPauseDur** (the total duration of the between-clause pauses in seconds)

3-2) Within-clause pauses: Example: I // live in Tokyo and study // at a // university. This includes all the pauses from the beginning and the end of the clause.

◦ **WCPauseFreq** (frequency of the within-clause pauses per 100 syllables)
◦ **WCPauseDur** (the total duration of the within-clause pauses in seconds)

2-3-2. Syntactic complexity indices

Among many syntactic complexity indices, the following were chosen for the present analysis following Spring's (2022) recommendation.

1) Length of the production unit
MLT (Mean length of T-unit: # of words/# of T-units)

2) Degree of Subordination
C/T (T-unit complexity ratio: # of clauses/# of T-units)

3) Amount of Coordination
CP/T (Coordinate phrases per T-unit: # of coordinate phrases/# of T-units)

4) Degree of phrasal sophistication
CN/T (Complex nominals per T-unit: # of complex nominals/# of T-units)
VP/T (Verb phrases per T-unit: # of verb phrases/# of T-units)

In the same way, the following lexical complexity indices were chosen for the present analysis.

1) Lexical Sophistication
LS1 (Lexical Sophistication-I: N_{slex}/N_{lex})
CVS1 (Corrected Verb Sophistication-1: $T_{sverb}/\sqrt{2N_{verb}}$)

2) Lexical Variation
NDW-ER50 (Mean T of 10 random 50-word samples)
CTTR (Corrected Type Token Ratio: $T/\sqrt{2N}$)
LV (Lexical Word Variation: Tlex/Nlex)
CVV1 (Corrected Verb Variation-1: $Tverb/\sqrt{2Nverb}$)

Note: T=total types; N=total words; S=sophistication; LEX=lexical items.

2-3-3. Analytical procedure and techniques for the individual longitudinal data (Verspoor, Bot, & Lowie, 2011)

First, the data points of each index were smoothed using LOESS (locally weighted scatter plot smoothing) in *Microsoft Excel*. To examine the interactions among the indices on the same scale, the data points were normalized such that each index has a maximal value of 1 and a minimal value of 0.

1) Mean comparisons

Non-parametric Mann-Whitney tests were employed to compare the means of an index between two stages. As the statistical results were relatively straightforward, *U*, *W*, and *Z*-values were unreported due to space limitations.

2) Min-max graph

Min-max graphs were used to examine how the index and its variability changed over the data period. Each data point represented a moving window of five data points.

3) Change-Point Analysis

This statistical analysis attempts to pinpoint when an index changes significantly along time series data (Horváth & Rice, 2024). It has recently been used in previous studies on writing development (Baba, 2020; Baba & Nitta, 2014; Wang & Tao, 2020). The present study used *Change-Point Analyzer,* Version 2.3 (Taylor, 2000). The analysis settings were as follows: the number of bootstrap samples=5,000 without replacement; the method of estimation=MSE Estimates; the confidence interval around change=95%; the inclusion in the table of changes=90%; identifying candidate changes=50. Five data points were combined to satisfy the assumptions for the statistical analyses, except for a few cases. The number of data points combined will be reported below.

4) Correlation and moving correlation analyses of indices

Correlation and detrended analyses of the indices across all (or half of) the data points were conducted using a bootstrapping method (10,000 times). The detrended correlation coefficients were calculated to eliminate the effects of changes due to the data trend in the correlation coefficients. The constant and slope were calculated using a linear regression for each index. The detrended data were created by subtracting the data for each data point from the raw data. The moving correlation analyses were conducted us-

Table 1. The mean and SD (standard deviation) of fluency indices across the stages of the program. The *p*-value in the rightmost column shows the result of non-parametric tests comparing the means of Stages 1 and 4. SRP=Speech rate (pruned); AR=Articulation rate; MLoR=Mean length of run; BCPauseFreq=Between-clause pause frequency; BCPauseDur=Between-clause pause duration; WCPauseFreq=Within-clause pause frequency; WCPauseDur=Within-clause pause duration; 1yr 1st Half=the 1st half of the 1st year; 1yr 2ndHalf=the 2nd half of the 1st year; 2ndyr IntTrn=the 1st five months of the 2nd year with intensive speaking training; SA=six months of the study abroad period; Post-SA=three months after SA.

		Stage						
		1	2	3	4	5		Stage 1 vs 4
		1yr 1stHalf	1yr 2ndHalf	2ndyr IntTrn	SA	Post-SA	Total	*p*
	Months	$0 \leq M < 6$	$6 \leq M < 12$	$12 \leq M < 16$	$16 \leq M < 21$	$21 \leq M < 24$		
SRP	*Mean*	87.96	93.27	110.70	133.75	121.94	106.99	<.001
	SD	10.12	4.80	9.80	6.20	18.52	20.12	
AR	*Mean*	172.15	178.35	175.92	187.86	186.35	179.48	<.001
	SD	8.77	5.97	3.72	6.27	11.45	9.32	
MLOR	*Mean*	4.14	4.78	5.58	6.94	5.97	5.35	<.001
	SD	0.67	0.21	0.40	0.48	0.81	1.10	
BCPauseFreq	*Mean*	8.48	8.42	8.47	9.16	8.75	8.63	0.002
	SD	1.05	0.47	0.98	0.79	0.88	0.84	
BCPauseDur	*Mean*	1.57	1.79	1.47	1.08	1.32	1.48	<.001
	SD	0.42	0.16	0.25	0.13	0.34	0.37	
WCPauseFreq	*Mean*	14.44	11.70	10.66	9.18	10.92	11.55	<.001
	SD	4.26	0.67	1.08	1.06	1.63	2.76	
WCPauseDur	*Mean*	0.95	0.98	0.80	0.72	0.75	0.86	<.001
	SD	0.13	0.12	0.09	0.06	0.10	0.15	
	N	42	43	29	32	29	173	

ing a window of 15 data points to ensure the reliability of the correlation coefficients (Bulté & Housen, 2020).

3. Results

3-1. Specific Research Question 1: How did fluency, lexical, and syntactic complexity indices change during the study abroad program?

As shown in Table 1, all fluency indices, except for AR, changed significantly in the expected direction. Notably, the most prominent change occurred between Stage 3 and Stage 4, indicating a substantial improvement in P's fluency before and after the start of the study abroad period. Additionally, all indices regressed between Stages 4 and 5, sug-

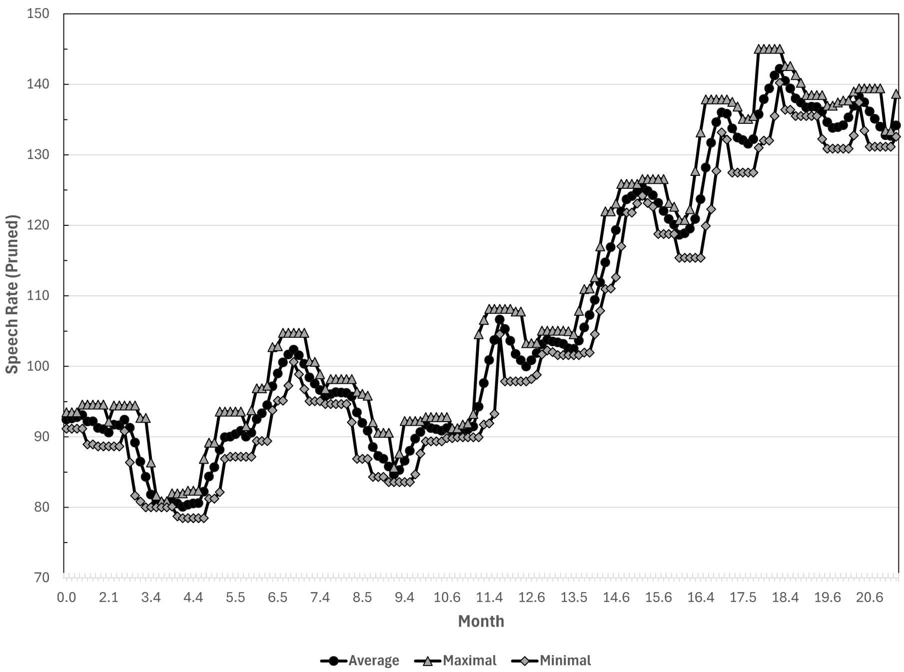

Figure 1. The moving average of Speech rate (pruned) with the window size of five data points and the maximal/minimal value within each window.

gesting a decline in P's fluency to some extent after the conclusion of the study abroad period. Statistical analyses (i.e., non-parametric Mann-Whitney tests) found that SR, AR, and MLoR significantly increased, while BCPauseDur, WCPauseFreq, and WCPauseDur significantly decreased between Stages 1 and 4. Overall, these findings indicated that P significantly improved speech fluency due to the study-abroad program.

Figure 1 illustrates the min-max graph of SRP over time. As is common with the developmental data discussed earlier, SRP exhibited cycles of progression and regression. There was a notable steep increase between the 13th and 15th months, aligning with the intensive training period. This was followed by a short regression period, succeeded by another significant increase between the 15th and 18th months, corresponding to the end of the intensive training and the third month of SA. After this period, SRP reached an asymptotic level and declined following SA. The data indicates a non-linear developmental pattern and suggests possible interactions with other components of overall speaking ability.

As demonstrated in Table 2, the syntactic complexity indices exhibited patterns similar to those of the fluency indices. Overall, there was a notable increase leading up to SA,

Table 2. The mean and SD of the syntactic complexity indices across the stages of the program. The *p*-value in the rightmost column shows the result of non-parametric tests comparing the means of Stages 1 and 4. MLT=Mean length of T-Unit; C/T=T-unit complexity ratio; CP/T=Coordinate phrases per T-unit; CN/T=Complex nominals per T-unit; VP/T=Verb phrases per T-unit. See the description of Table 1 for the abbreviations under Stage numbers.

		Stage						
		1	2	3	4	5		Stage 1 vs 4
		1yr 1stHalf	1yr 2ndHalf	2ndyr IntTrn	SA	Post-SA	Total	*p*
	Months	0 ≤ M < 6	6 ≤ M < 12	12 ≤ M < 16	16 ≤ M < 21	21 ≤ M < 24		
MLT	*Mean*	7.08	7.38	8.35	10.06	8.99	8.23	<.001
	SD	0.76	0.39	0.58	0.81	0.77	1.29	
C/T	*Mean*	1.15	1.18	1.15	1.37	1.26	1.21	<.001
	SD	0.07	0.05	0.06	0.08	0.09	0.11	
CP/T	*Mean*	0.072	0.072	0.085	0.115	0.108	0.088	<.001
	SD	0.052	0.024	0.058	0.040	0.042	0.047	
CN/T	*Mean*	0.327	0.451	0.516	0.729	0.601	0.508	<.001
	SD	0.118	0.074	0.065	0.109	0.144	0.173	
VP/T	*Mean*	1.65	1.61	1.75	2.10	1.86	1.77	<.001
	SD	0.39	0.16	0.25	0.24	0.24	0.32	
	N	42	43	29	32	29	173	

followed by a slight decline after SA. Statistical analyses revealed highly significant differences between Stages 1 and 4 across all indices. The results indicated that P made significant improvements in the length of the production unit (i.e., MLT), the level of subordination (i.e., C/T), the amount of coordination (i.e., CP/T), and the degree of phrasal sophistication (i.e., CN/T and VP/T) during the program.

Table 3 illustrates that the lexical complexity indices exhibited a pattern distinct from the fluency and syntactic complexity indices. Firstly, two indices, CVS1 and NDW-ERZ, did not show significant changes between Stages 1 and 4. Secondly, two other indices, LS1 and LV, significantly decreased between these stages. These findings suggest that P could not enhance lexical sophistication (i.e., LS1 and CVS1) and lexical variation (i.e., NDW-ER50 and LV). Conversely, the lexical variation measures corrected for the text length (i.e., CTTR and CVV1) demonstrated a similar trend as in fluency and syntactic complexity measures.

Table 3. The mean and SD of the lexical complexity indices across the stages of the program. The *p*-value in the rightmost column shows the result of non-parametric tests comparing the means of Stages 1 and 4. LS1=Lexical sophistication-I; CVS1=Corrected verb sophistication-1; NDWER-ER50=Number of different words (random 50); CTTR=Corrected type token ratio; LV=Lexical word variation; CVV1=Corrected verb variation-1. See the description of Table 1 for the abbreviations under Stage numbers.

		Stage						
		1	2	3	4	5		Stage 1 vs 4
		1yr 1stHalf	1yr 2ndHalf	2ndyr IntTrn	SA	Post-SA	Total	*p*
	Months	0 ≤ M < 6	6 ≤ M < 12	12 ≤ M < 16	16 ≤ M < 21	21 ≤ M < 24		
LS1	*Mean*	0.249	0.257	0.226	0.194	0.220	0.233	<.001
	SD	0.044	0.035	0.033	0.040	0.036	0.043	
CVS1	*Mean*	0.059	0.049	0.113	0.051	0.064	0.064	>.05
	SD	0.048	0.041	0.041	0.035	0.074	0.051	
NDWER-ER50	*Mean*	35.72	36.06	35.85	36.16	35.83	35.93	>.05
	SD	0.97	0.68	0.78	0.69	0.73	0.75	
CTTR	*Mean*	3.81	3.90	4.03	4.26	4.07	3.99	<.001
	SD	0.14	0.12	0.16	0.09	0.20	0.21	
LV	*Mean*	0.839	0.830	0.818	0.783	0.798	0.816	<.001
	SD	0.026	0.019	0.027	0.027	0.021	0.031	
CVV1	*Mean*	1.99	1.98	2.23	2.33	2.18	2.12	<.001
	SD	0.15	0.09	0.24	0.06	0.23	0.21	
	N	42	43	29	32	29	173	

3-2. Specific Research Question 2) At which points in development did the fluency, syntactic, and lexical complexity indices significantly change? How did the change points differ among fluency, syntactic, and lexical development?

Table 4 presents the results of the change-point analyses for the syntactic complexity indices. All indices, except for BCPauseFreq, identified at least one change point with a confidence level of 95% or higher. As illustrated in Figure 1, fluency development demonstrated a phase of rapid improvement followed by a period of relative stability or regression. Notably, multiple indices indicated significant (or marginally significant) change points occurring at similar times. For instance, SRP, MLoR, BCPauseDur, and WCPauseDur all showed a change point in the 12th month, which aligns with the onset of the intensive speaking training period in the second year. Furthermore, SRP, MLoR, BCPauseFreq, BCPauseDur, WCPauseFreq, and WCPauseDur showed shifts in the 15th month or at the beginning of the 16th month, coinciding with the conclusion of the intensive training period and the start of the study abroad period. Additionally, the confidence intervals for SRP, MLoR, and BCPauseDur were minimal. It was also found that the

Table 4. The results of the change-point analysis conducted by the *Change-Point Analyzer* on the fluency indices. CI=confidence interval; CL=confidence level; From=the mean before the change point; To=the mean after the change point.

Index	Change #	Month	CI	CL	From	To	# of Combined Data Points	*p*
SRP	1	6.5	(4.6, 8.4)	92%	86.8	92.9	5	<.001
	2	12.9	(10.8, 12.9)	95%	93.4	103.9	5	<.001
	3	15.3	(15.3, 15.3)	92%	102.6	121.7	5	<.001
	4	17.8	(17.8, 17.8)	100%	121.7	135.9	5	<.001
AR	1	6.7	(4.4, 7.4)	92%	171.9	180.4	6	<.001
	2	11.2	(9.7, 14.4)	90%	180.4	175.3	6	<.001
	3	17.9	(17.9, 17.9)	100%	175.3	191.9	6	<.001
MLoR	1	6.5	(6.5, 6.6)	99%	4.06	4.75	8	<.001
	2	12.6	(12.6, 12.6)	92%	4.75	5.52	8	<.001
	3	16.4	(16.4, 16.4)	98%	5.52	6.78	8	<.001
BCPauseFreq	1	15.8	(9.7, 21.4)	91%	8.3	9.1	5	<.001
BCPauseDur	1	3.1	(3.1, 3.1)	96%	1.126	1.827	5	<.001
	2	12.3	(7.7, 12.9)	94%	1.827	1.611	5	<.001
	3	15.3	(15.3, 15.3)	100%	1.611	1.127	5	<.001
WCPauseFreq	1	2.5	(2.5, 2.5)	100%	20.6	11.0	5	<.001
	2	15.6	(15.1, 17.9)	100%	11.2	9.6	5	<.001
WCPauseDur	1	2.5	(2.5, 7.7)	93%	0.805	0.945	5	<.001
	2	9.7	(5.2, 10.8)	99%	0.945	1.085	5	<.001
	3	12.3	(12.3, 12.3)	94%	1.085	0.866	5	<.001
	4	15.3	(14.6, 16.4)	98%	0.866	0.712	5	<.001

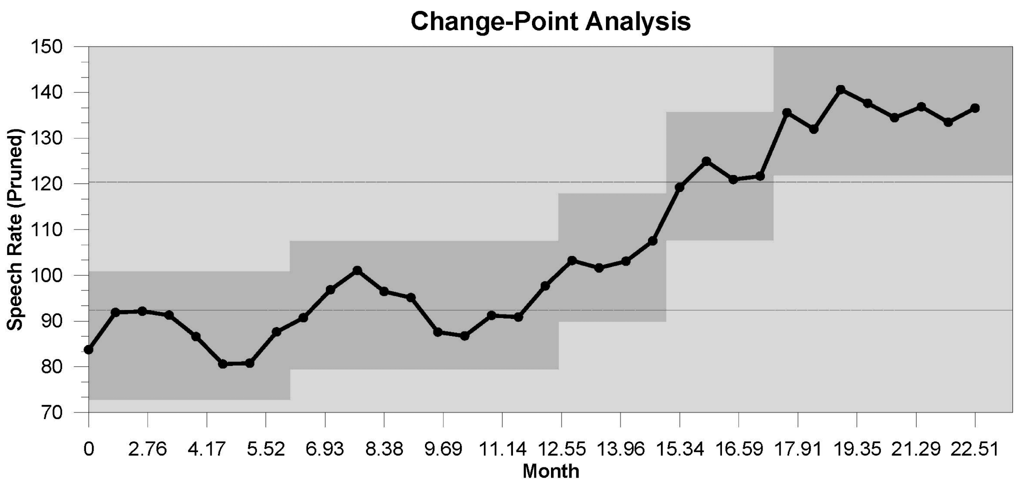

Figure 2. The graph produced by the *Change-Point Analyzer* on the SRP index across the data points during the study abroad program.

Table 5. The results of the change-point analysis conducted by the *Change-Point Analyzer* on the syntactic complexity indices. CI=confidence interval; CL=confidence level; From=the mean before the change point; To=the mean after the change point.

Index	Change #	Month	CI	CL	From	To	# of Combined Data Points	*p*
MLT	1	5.8	(5.8, 9.7)	93%	6.96	7.58	5	<.001
	2	15.3	(15.3, 15.3)	100%	7.58	9.08	5	<.001
	3	18.4	(18.4, 20.3)	99%	9.08	10.13	5	<.001
C/T	1	15.8	(15.3, 15.8)	96%	1.12	1.29	5	<.001
	2	18.9	(18.9, 21.4)	98%	1.29	1.38	5	<.001
CP/T	1	15.1	(9.0, 20.3)	97%	0.07	0.11	6	<.001
CN/T	1	8.4	(7.0, 8.4)	100%	0.33	0.47	5	<.001
	2	14.1	(13.4, 15.3)	99%	0.47	0.59	5	<.001
	3	18.9	(18.9, 21.4)	95%	0.59	0.76	5	<.001
VP/T	1	3.9	(3.9, 4.6)	92%	1.43	2.09	5	<.001
	2	5.2	(4.6, 5.2)	100%	2.09	1.60	5	0.003
	3	14.6	(14.6, 17.1)	100%	1.60	2.01	5	<.001

means for each index before and after the change points differed significantly across all fluency indices.

As exemplified in Figure 2, the *Change-Point Analyzer* detected four significant change points where SRP increased to the next level.

Table 5 illustrates the change points in the indices of syntactic complexity. It is noteworthy that, compared to the fluency indices, the initial significant change points occurred later, around the 14th or 15th month, corresponding to the final two months of the speaking training period. Additionally, significant change points were observed in the 18th month across three indices (i.e., MLT, C/T, and CN/T), which aligns with the third month of the study-abroad period. These results suggest that improvements in syntactic complexity may occur later than enhancements in fluency during L2 speech development.

Table 6 presents the change points in the indices of lexical complexity. Two indices of lexical sophistication (i.e., LS1 and CVS1), along with LV, exhibited significant decreases in the 15th and 16th months. In contrast, the two indices of lexical variation (i.e., CTTR and CVV1), showed a simultaneous change point at approximately 15.3 months, coinciding with the change points identified in the syntactic complexity indices during the last two months of the speaking training period. Additionally, CTTR and CVV1 indicated further significant decreases later, specifically at 21.4 and 18.0 months, respectively.

Table 6. The results of the change-point analysis conducted by the *Change-Point Analyzer* on the lexical complexity indices. CI=confidence interval; CL=confidence level; From=the mean before the change point; To=the mean after the change point.

Index	Change #	Month	CI	CL	From	To	# of Combined Data Points	*p*
LS1	1	15.3	(12.9, 16.4)	100%	0.25	0.20	5	<.001
CVS1	1	10.8	(8.4, 10.8)	100%	0.04	0.10	5	<.001
	2	16.4	(16.4, 16.4)	100%	0.10	0.04	5	<.001
	3	21.4	(21.4, 22.5)	95%	0.04	0.11	5	<.001
CTTR	1	11.4	(9.0, 13.4)	91%	3.82	3.99	5	<.001
	2	15.3	(14.6, 15.3)	95%	3.99	4.24	5	<.001
	3	21.4	(18.4, 22.5)	94%	4.24	4.12	5	<.001
LV	1	16.4	(14.1, 18.4)	99%	0.83	0.79	5	<.001
CVV1	1	15.3	(15.3, 15.3)	100%	1.99	2.42	5	<.001
	2	18.9	(15.8, 20.3)	96%	2.42	2.27	5	<.001

3-3. Specific Research Question 3) How were the fluency, lexical, and syntactic complexity indices correlated during the program?

Table 7 presents the correlation coefficients between two fluency indices, namely SRP and MLoR, and the syntactic and lexical complexity indices that exhibited significant increases and change points. The correlations based on raw scores were significant ($p<0.001$) and ranged from moderate to strong (i.e., 0.57 to 0.90), likely reflecting the observable trend of increase in each index. For the detrended correlation, SRP exhibited a moderate and significant correlation ($p<0.001$) with the syntactic complexity indices (MLT and C/T) and the lexical complexity indices (CTTR and CVV1), with correlation coefficients ranging from 0.4 to 0.6. However, SRP showed a significant but weak correlation ($p<0.001$) with two syntactic complexity indices (CN/T and VP/T), with coefficients of 0.22 and 0.28, respectively. MLoR displayed a strong correlation with MLT (0.62), a moderate correlation with C/T and CN/T (0.52 and 0.48, respectively), and a weak correlation with VP/T (0.21), as well as with the two lexical complexity indices (CTTR and CVV1, both at 0.39). For further details, refer to Appendix 1 for the raw correlation coefficients and Appendix 2 for the detrended correlation coefficients among all the indices.

The previous data analyses revealed that many indices significantly increased in the middle of the observed period, suggesting that the interactions among the indices might differ between the first and second halves of the time frame. The bottom half of Table 7 displays the correlation coefficients for each period. As anticipated, the correlation coeffi-

Table 7. Correlation coefficients between the fluency indices (i.e., SRP=Speech rate (pruned), MLoR=Mean length of runs), the syntactic complexity indices (i.e., MLT=Mean length of T-unit, C/T=T-unit complexity ratio, CN/T=Coordinate phrases per T-unit, VP/T=Verb phrases per T-unit, and the lexical complexity indices (i.e., CTTR=Corrected type token ratio, CVV1=Corrected verb variation-1)).

			Fluency		Syntactic				Lexical	
			SRP	MLoR	MLT	C/T	CN/T	VP/T	CTTR	CVV1
Raw Score	All	SRP	1	0.957**	0.904**	0.735**	0.843**	0.569**	0.840**	0.780**
		MLoR	0.957**	1	0.910**	0.770**	0.890**	0.569**	0.830**	0.769**
Detrended	All	SRP	1	0.748**	0.602**	0.401**	0.280**	0.215**	0.446**	0.435**
		MLoR	0.748**	1	0.622**	0.520**	0.476**	0.210**	0.390**	0.391**
	1st Half	SRP	1	0.851**	0.543**	0.262*	0.088	−0.058	.505**	0.346**
		MLoR	.851**	1	0.297**	0.068	0.107	−0.086	0.414**	0.338**
	2nd Half	SRP	1	0.736**	0.678**	0.555**	0.489**	0.558**	0.410**	0.519**
		MLoR	0.736**	1	0.767**	0.656**	0.68**	0.418**	0.372**	0.409**
		N	173	173	173	173	173	173	173	173

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).
Unless otherwise noted, bootstrap results are based on 10000 bootstrap samples

cients were higher in the second half of the overall period (i.e., the second year) for all indices except for CTTR. This difference was particularly evident in the correlations between SRP and MLoR with the syntactic complexity indices for phrasal sophistication (i.e., CN/T and VP/T). These results imply that the relationship between fluency and syntactic development may be generally competitive in the first half and supportive in the latter half. Consequently, further examination of the interactions between the fluency indices and the syntactic and lexical complexity indices was deemed necessary.

3-4. Specific Research Question 4) How did the composite fluency indices (i.e., SRP (speech rate (pruned) and MLoR (mean length of runs)) interact with the syntactic and lexical complexity indices?

Figure 3 illustrates the interaction between SRP (i.e., speech rate (pruned) and C/T (i.e., Coordinate phrases per T-unit) during the entire study period, using normalized scores that ranged from 0 to 1. The two indices exhibited negative correlations between 4.4 to 5.5 months and 9.4 to 10.7 months; when one index increased, the other decreased. However, from around 13.0 to 17.5 months, the indices displayed a positive correlation, generally moving in the same direction. Notably, there was a rapid increase――a spurt――between 15.0 and 16.5 months, indicating parallel movement. This period coincided

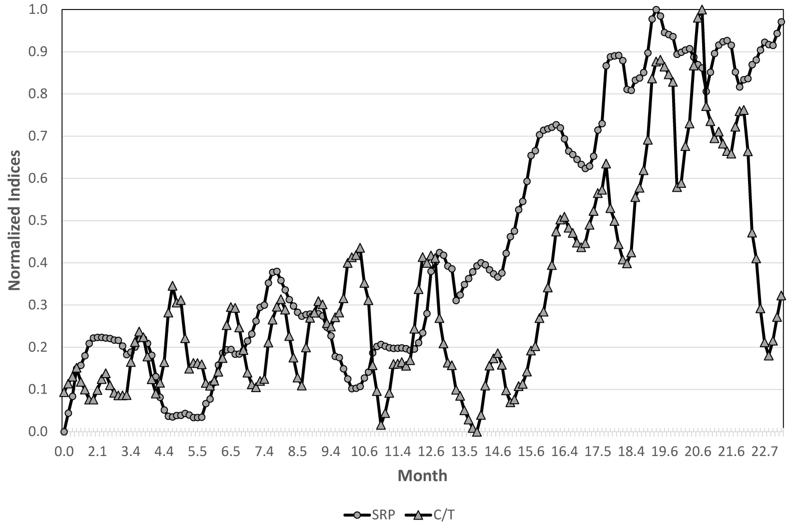

Figure 3. The interaction of SRP (i.e., Speech rate (pruned) and C/T (i.e., Coordinate phrases per T-unit) with normalized scores as a function of data points.

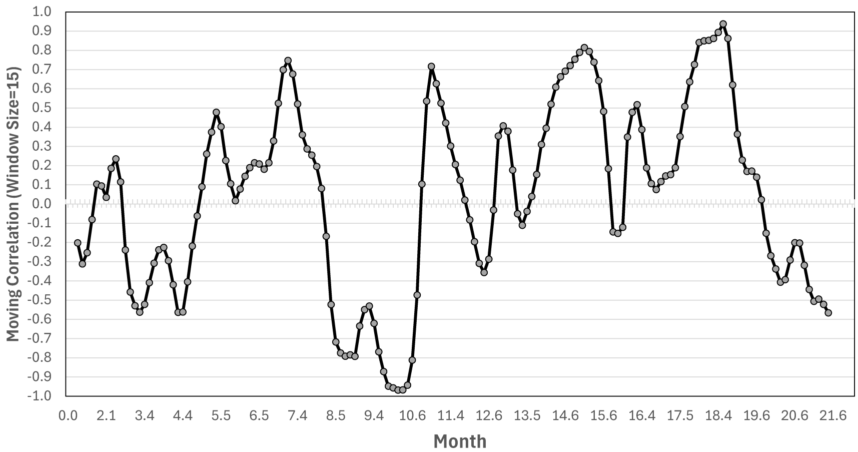

Figure 4. The moving correlation of SRP (i.e., Speech rate (pruned) and C/T (i.e., Coordinate phrases per T-unit) with normalized scores as a function of data points.

with the conclusion of intensive speaking training and the commencement of study abroad. A negative correlation was observed at the program's end, specifically between the 22nd and 23rd months.

Figure 4 illustrates the moving correlation between the two indices using a window size of 15 data points. The graph reveals alternating periods of negative and positive correlations up until around the 10th month. Specifically, there were two notable periods of

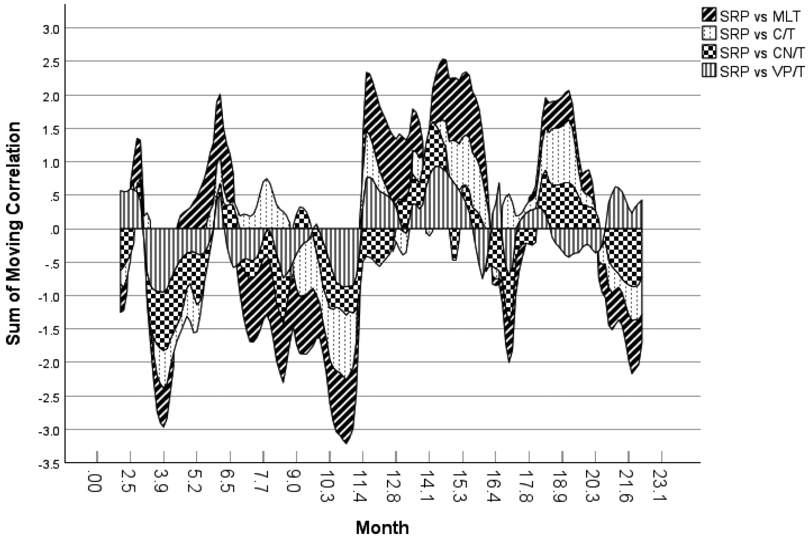

Figure 5. The moving correlation of SRP (i.e., Speech rate (pruned) and the sum of the syntactic complexity indices. MLT=Mean length of T-Unit; C/T=T-unit complexity ratio; CN/T=Complex nominals per T-unit; VP/T=Verb phrases per T-unit.

negative correlation: the first occurred between 3.0 and 4.5 months, and the second, which was considerably more negative, was observed between 8.5 and 10.8 months. Following this latter period, the correlation shifted primarily to positive, with significantly high correlations recorded between 14.6 and 15.8 months, as well as between 18 and 19 months. It is important to note that, due to the relatively large window size of 15 data points, the graph reflects the overall trends in correlation between the two indices. Additionally, it is worth mentioning that the decrease in correlation was noted just before the significant increase in the fluency and syntax complexity indices.

Figure 5 examined whether the interaction pattern was consistent across various indices of syntactic complexity. The moving correlations between SRP (i.e., speech rate (pruned)) and the indices of syntactic complexity were calculated, summed, and then plotted over the data points. The results showed that the moving correlation was predominantly negative during the first year, specifically before the 12th month, after which it shifted to being mainly positive. This indicates that the relationship between fluency and syntactic development was largely competitive during the first year and became supportive in the second year.

The same analysis was conducted to examine the interaction between SRP (i.e., speech rate (pruned)) and the lexical complexity indices. Figure 6 illustrates the relation-

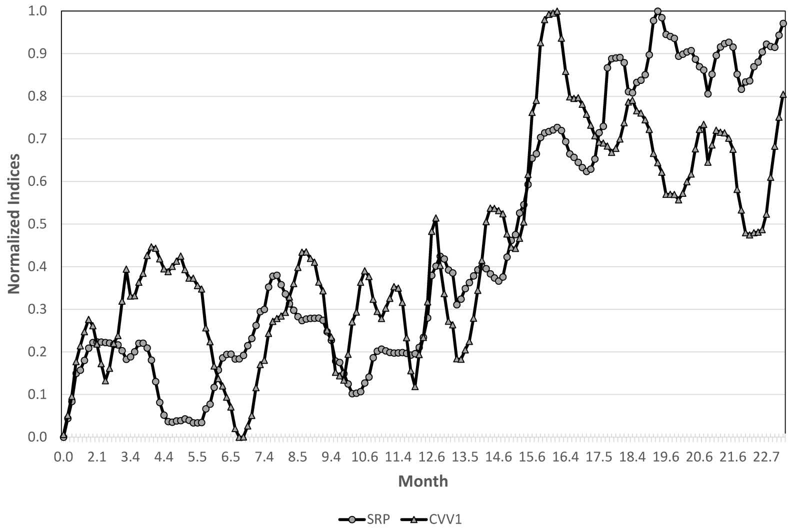

Figure 6. The interaction of SRP (i.e., Speech rate (pruned) and CVV1 (i.e., Corrected verb variation-1) with normalized scores as a function of data points.

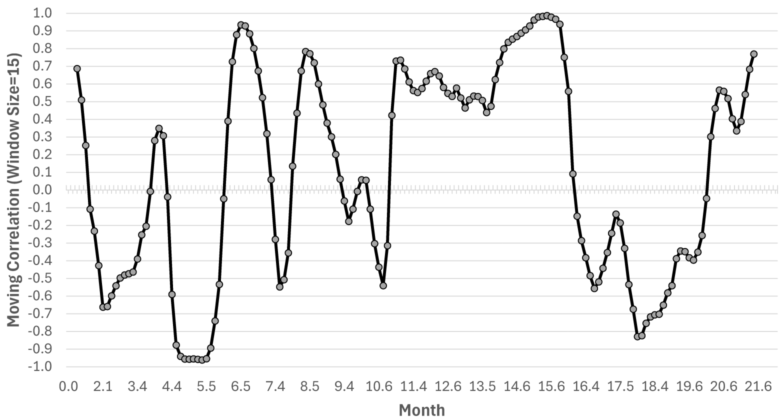

Figure 7. The moving correlation of SRP (i.e., Speech rate (pruned) and C/T (i.e., Coordinate phrases per T-unit) with normalized scores as a function of data points.

ship between SRP and CVV1 (i.e., corrected verb variation). A negative correlation was observed during the relatively early period, specifically between 3.4 and 6.5 months. However, the relationship shifted after this period and remained mainly positive until approximately 17.5 months. Following that, up to 20.6 months, the correlation became negative again, but in the final few months, a positive correlation was observed once more.

Figure 7 shows the moving correlation between the two indices. Although a dip was observed before the prolonged positive correlation started around the 11th month, its

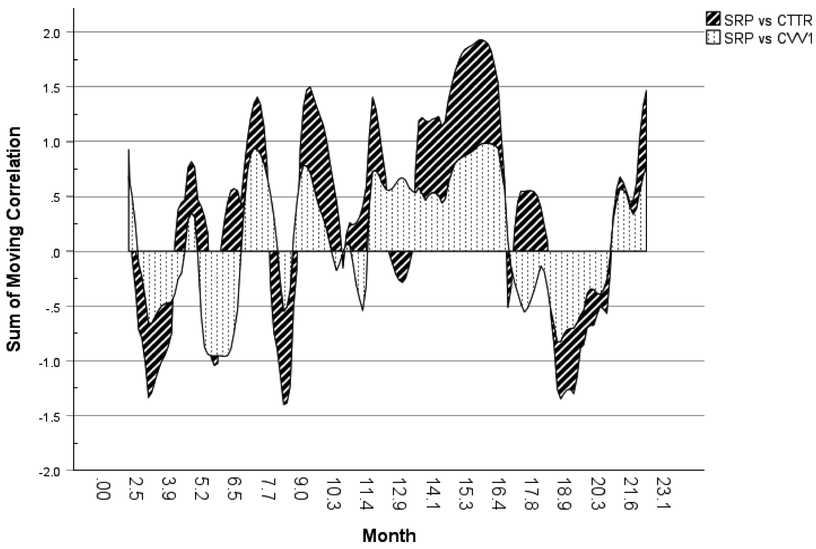

Figure 8. The moving correlation of SRP (i.e., Speech rate (pruned) and the sum of the lexical complexity indices. CTTR=Corrected type token ratio; CVV1=Corrected verb variation-1.

magnitude and length were much smaller.

The interaction between SRP and the lexical complexity indices, specifically CTTR and CVV1, is illustrated in Figure 8. This interaction revealed a pattern of alternating positive and negative correlations. Notably, the timing of these alternations differed between CTTR and CVV1. The correlation was mainly positive between 11.0 and 17.0 months for both indices. However, a negative correlation was observed between 18.3 and 21.4 months for both indices. These results indicate that the relationship between fluency and lexical complexity was primarily positive during the intensive speech training period and the first half of the study abroad experience but primarily negative in the latter half.

4. Discussion and Conclusion

The current study aimed to present new data on the interactions among L2 fluency, lexical development, and syntactic development in oral narratives produced by a Japanese adult learner of English. This research was conducted over a two-year study abroad program, which included approximately 5.5 months spent studying in Australia. The following sections summarize the results and offer possible interpretations for each specific research question.

4-1. Specific Research Question 1: How did fluency, lexical complexity, and syntactic complexity indices change over the two-year period, including the approximately 5.5-month study-abroad period?

All fluency indices showed significant improvement between the first year and the study-abroad period (see Table 1). This included composite measures (i.e., SRP (speech rate (pruned)) and MLoR (mean length of runs)), a speed measure (i.e., AR (articulation rate)), and breakdown measures (frequencies and durations of pauses both between and within clauses). These findings are consistent with previous research indicating that study-abroad experiences significantly enhance the speech fluency of L2 learners (Tsushima, 2019; Valls-Ferrer & Mora, 2014). The indices also exhibited a wave-like change characterized by alternating progression and regression. (see Figure 1). The first noticeable improvement occurred during the intensive speech training period in the first half of the program's second year, while the second improvement took place during the initial half of the study-abroad period. This finding aligns with Baba's (2020) observation that the high-growth group's writing fluency exhibited a wave-like progression over a 30-week study period (see Figure 2 on p. 15).

In terms of syntactic complexity indices, all measures demonstrated significant improvement over the period studied (see Table 2). This includes the length of the production unit (i.e., MLT), the level of subordination (i.e., C/T), the amount of coordination (i.e., CP/T), and the degree of phrasal sophistication (i.e., CN/T and VP/T). The findings align with prior research, showing that both preparation before studying abroad and the experience of studying abroad positively influence the syntactic development of second language learners (Borràs & Llanes, 2021).

The results of the lexical complexity indices indicated that the lexical sophistication measures did not show significant changes (e.g., CVS1) or exhibited a significant decline (e.g., LS1) between the beginning of the program and the study-abroad period (see Table 3). Additionally, one lexical variation index did not show significant change (i.e., NDW-ER50), while another (i.e., LV) showed a significant decline during the same timeframe. Conversely, two lexical variation indices (i.e., CTTR and CVV1) demonstrated significant improvement. The lack of significant increases in the aforementioned indices may be attributed to the relatively small number of words produced in a single narrative, particularly at the beginning of the entire period. In contrast, the corrections made for text length in CTTR and CVV1 may have facilitated the identification of significant differences. To address this issue, a reanalysis of the data using a modified method, such as combining multiple texts, may be warranted. Overall, the present findings are consistent with previous research suggesting that studying abroad has a limited impact on lexical development (Zaytseva, 2016).

4-2. Specific Research Question 2: At which points did the fluency, syntactic complexity, and lexical complexity indices significantly change? How did the change points differ among the indices?

Significant change points were identified in fluency, syntactic complexity, and lexical complexity indices (see Figure 2 and Tables 4, 5, and 6). The result aligned with previous research within the framework of CDST, which has demonstrated such shifts in writing development (e.g., Baba & Nitta, 2014; Kyle et al., 2021; Spoelman & Verspoor, 2010; Verspoor, 2017). It was found that phase shifts were detected among multiple fluency indices in the 12th month, including composite measures (i.e., SRP and MLoR) and breakdown measures (i.e., BCPauseDur and WCPauseDur; see Table 4). This timing corresponded with the start of the intensive speech training period in the 1st month. The results indicated that various indices across all three areas experienced phase shifts roughly simultaneously. Specifically, fluency measures (i.e., SRP, BCPauseDur, and WCPauseDur), syntactic complexity measures (i.e., MLT, C/T, CP/T, CN/T, and VP/T), and lexical complexity measures (i.e., CTTR and CVV1) showed significant change points in either the 14th or 15th month. The timing aligned with the final two months of an intensive speech training period. It was observed that several fluency indices (such as SPR, AR, and MLoR) exhibited noticeable changes in the 16th or 17th month, which corresponded to the first two months of studying abroad (see Table 4). Additionally, various indices of syntactic complexity showed shifts in the 18th month, aligning with the third month of the study abroad period (see Table 5). Previous research on speech production has indicated that L2 fluency may be linked to several factors. These include accessing and retrieving lexical items, creating a concept for the utterance, formulating a syntactic structure, developing a phonological structure, and executing articulation (Kormos, 2006; Lambert, Aubrey, & Leeming, 2021; Lambert, Kormos, & Minn, 2017; Levelt, 1989). The results indicated that, during the overall increase following the change points, many aspects of speech production improved simultaneously.

4-3. How were the fluency, syntactic complexity, and lexical complexity indices correlated?

The results indicated that the detrended correlations between fluency indices——specifically, the composite measures SPR and MLoR——were significantly lower in the first half of the entire period, ranging from approximately 0.0 to 0.3 (except MLT). In contrast, these correlations increased in the second half of the period, reaching around 0.4 to 0.8 (see Table 7). This suggests that, on average, the relationship between fluency and syntactic complexity was competitive during the first half and became supportive after the first change point, which occurred around the 14th and 15th months. During the program's first year, producing more syntactically complex sentences resulted in slower speech and a more significant number or length of pauses. However, after the change point, fluency was no longer adversely affected by producing more syntactically complex sentences. Both fluency and the degree of syntactic complexity improved significantly in tandem. On the other hand, the detrended correlations between fluency indices and measures of lexical complexity——namely CTTR and CVV1——were moderately high, ranging from about 0.4 to 0.5. These values did not show substantial differences between the first and second halves of the period. This implies that, on average, producing a greater variety of vocabulary (as measured by CTTR) or verbs (as measured by CVV1) did not negatively impact the speed, frequency, and duration of pauses in speech.

4-4. How did the composite fluency indices (i.e., SRP (speech rate (pruned) and MLoR (mean length of runs)) interact with the syntactic and lexical complexity indices?

The analysis of the interaction between fluency indices and syntactic complexity indices revealed several key findings. In the first half of the observation period, an alternating pattern of positive and negative moving correlations was noted (see Figures 3 and 4 for the interaction of SRP and C/T). A significant, simultaneous increase followed this in both indices. When the moving correlations of SRP and the syntactic complexity indices were aggregated, the correlations were primarily negative, except for MLT (see Figure 5).

Interestingly, the degree of negative correlation was particularly low for two months leading up to the 12th month, coinciding with a phase shift in multiple fluency indices. After this shift, a positive correlation persisted for about five months during an intensive training period. A brief period of negative correlation was observed for a couple of months after the study-abroad period began, followed by a phase of positive correlation until the conclusion of the study-abroad experience. Subsequently, the correlation shifted mainly to negative as the syntactic complexity indices (excluding VP/T) declined, while fluency levels remained stable.

The interaction between fluency indices and lexical complexity indices exhibited an alternating pattern of positive and negative correlations during the first half of the study period (see Figures 6 and 7 for the interaction of SRP and CVV1). This was followed by a prolonged period of high correlation, which coincided with the intensive speech training phase. Subsequently, the correlation turned primarily negative during the study-abroad period before returning to a positive correlation. This pattern is also evident when examining the summed moving correlations of the two lexical complexity indices (see Figure 8), as both indices reflect lexical variability.

The first year of P's fluency, syntax, and lexical complexity learning can be characterized by alternating subperiods of competition and automatization. Various components may compete for limited resources during this time, while one component becomes automatized. Subsequently, another set of components may compete, leading to the automatization of a different aspect. This cycle may explain the observed pattern of alternating positive and negative correlations. However, a more detailed data analysis is necessary to identify the specific set of indices that competed during this period.

During the period when fluency was improving, both the syntactic and lexical complexity indices also showed significant advancements, particularly during the intensive speech training phase. It is possible that the automatization of lexical retrieval and access contributed to this enhancement in fluency. This improvement may have led to a reduction in the frequency and duration of pauses and an increase in the average length of spoken runs. Similarly, automating syntactic processing may have further enhanced fluency by extending production units, promoting subordination and coordination, and increasing phrasal sophistication. However, a more detailed analysis is necessary to understand how these indices interacted during this phase.

In the latter half of the study-abroad period, a primarily negative correlation was found between fluency and syntactic complexity. In contrast, the relationship between fluency and lexical complexity was primarily positive. During this time, all indices of syntactic complexity improved (see Figure 3), whereas the indices for lexical complexity, including lexical and verb variety (CTTR and CVV1), decreased (see Figure 6). Despite these changes, fluency indices showed improvement. This suggests that while the automation of processing more complex syntactic structures may have enhanced fluency, it also competed with lexical variety.

4-5. Limitations of the study

As an exploratory individual study, this research has several limitations. First, the results cannot be generalized to L2 learners like any individual-based study. The available group data on the long-term interactions between oral fluency and syntactic and lexical complexity is limited. (but see Kanda, 2024). Further group-based research is needed to interpret specific individual data. Additionally, statistical data analysis using modeling techniques is currently underway. For example, Lowei, Capsi, van Geert, & Steenbeek (2014) introduce statistical techniques called the "dynamic growth model" to model individual, longitudinal data. It is particularly interesting to explore whether there is a precursor relationship between the development of fluency and syntactic complexity. For instance, the increase in the speech rate (SRP) seemed to occur before the development of the T-unit complexity ratio (C/T), as illustrated in Figure 3. The proposed statistical technique allows us to test for the existence of this precursor relationship. The final comment concerns a methodology in data acquisition. In the present study, P was instructed to spend at most two minutes before starting the narration. However, the author had little control over P's time on the pre-task planning. Previous research has shown that pre-task planning has beneficial and significant effects on oral narratives' syntactic and lexical complexity (Kanda, 2024). In future studies, this factor should be more carefully treated (e.g., recording the pre-task planning time).

4-6. Teaching implications

The current study highlights that improving oral fluency requires a significant amount of time, particularly for learners with relatively low proficiency. In Kanda's study (2024), a group of low-proficiency Japanese learners of English did not demonstrate significant improvement in their oral fluency after one year of study. In contrast, the present study indicates that participants' fluency improved noticeably during the intensive speech training in the second year of the learning period. This suggests that foreign language instructors should be mindful of the challenges faced by L2 learners. Additionally, the study implies that the iterative narrative tasks used positively affect developing fluency, syntactic complexity, and lexical complexity (Baba, 2020; Bygate, 2018; Kanda, 2024). In the conclusion section of her chapter (Larsen-Freeman, 2018), Larsen-Freeman states, "CDST [Complex Dynamic Systems Theory] proposes that learning is not a process of internalizing an external reality. Rather, through iterated opportunities to make meaning in specific contexts, learners perceive and act on their affordances, which assist them in constructing their language resources. (p. 324)" The present results suggest that iterative narrative tasks offer learners opportunities to enhance their linguistic skills.

4-7. Concluding remarks

The current study aimed to explore the development of L2 fluency and lexical and syntactic complexity during a study abroad program that lasted approximately 5.5 months. The findings were consistent with Complex Dynamic Systems Theory (CDST), revealing non-linear and interconnected fluency, syntactic complexity, and lexical complexity changes. Significant shifts were observed across all three components, and they interacted dynamically——sometimes competing with one another and at other times supporting each other during various stages of development. The author conducted a series of studies focusing on the individual development of speech fluency and rhythm (Tsushima, 2018, 2019, 2020, 2021, 2023), but these studies were not conducted within the framework of the CDST. This exploratory study reveals that the principles, concepts, and methodologies of CDST provide a deeper understanding of the learning process. By reanalyzing data from previous studies using CDST principles, dynamic developmental processes and interactions among relevant variables overlooked in earlier research may be uncovered.

Appendix 1. Correlation coefficients between the fluency indices (i.e., SRP=Speech rate (pruned), MLoR=Mean length of runs) and the Syntactic complexity indices (i.e., MLT=Mean length of T-unit, C/T=T-unit complexity ratio, CN/T=Coordinate phrases per T-unit, VP/T=Verb phrases per T-unit), and the lexical complexity indices (i.e., CT-TR=Corrected type token ratio, CVV1=Corrected verb variation-1).

		SRP	MLoR	MLT	C/T	CN/T	VP/T	CTTR	CVV1
Fluency	SRP	1	.957**	.904**	.735**	.843**	.569**	.840**	.780**
	MLoR	.957**	1	.910**	.770**		.569**	.830**	.769**
Syntax	MLT	.904**	.910**	1	.839**	.890**	.702**	.784**	.710**
	C/T	.735**	.770**	.839**	1		.670**	.624**	.584**
	CN/T	.843**	.890**	.869**	.759**	1	.550**	.749**	.662**
	VP/T	.569**	.569**	.702**	.670**	.550**	1	.572**	.620**
Vocabulary	CTTR	.840**	.830**	.784**	.624**	.749**	.572**	1	.788**
	CVV1	.780**	.769**	.710**	.584**	.662**	.620**	.788**	1
	N	173	173	173	173	173	173	173	173

** Correlation is significant at the 0.01 level (2-tailed).
Unless otherwise noted, bootstrap results are based on 10000 bootstrap samples

Appendix 2. Detrended correlation coefficients between the fluency indices (i.e., SRP=Speech rate (pruned), MLoR=Mean length of runs) and the Syntactic complexity indices (i.e., MLT=Mean length of T-unit, C/T=T-unit complexity ratio, CN/T=Coordinate phrases per T-unit, VP/T=Verb phrases per T-unit), and the lexical complexity indices (i.e., CTTR=Corrected type token ratio, CVV1=Corrected verb variation-1).

		SRP	MLoR	MLT	C/T	CN/T	VP/T	CTTR	CVV1
Fluency	SRP	1	.748**	.602**	.401**	.280**	.215**	.446**	.435**
	MLoR	.748**	1	.622**	.520**	.476**	.210**	.390**	.391**
Syntax	MLT	.602**	.622**	1	.685**	.505**	.542**	.322**	.268**
	C/T	.401**	.520**	.685**	1	.342**	.490**	.190*	.192*
	CN/T	.280**	.476**	.505**	.342**	1	.188*	.177*	0.11
	VP/T	.215**	.210**	.542**	.490**	.188*	1	.270**	.393**
Vocabulary	CTTR	.446**	.390**	.322**	.190*	.177*	.270**	1	.507**
	CVV1	.435**	.391**	.268**	.192*	0.11	.393**	.507**	1
	N	173	173	173	173	173	173	173	173

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).
Unless otherwise noted, bootstrap results are based on 10000 bootstrap samples

Acknowledgment

This paper was supported by a research grant from Tokyo Keizai University in the academic year of 2023 (23-19). The author is grateful to the project participant for her dedication, patience, and eagerness to enhance her English skills.

References

Baba, K. (2020). Exploring dynamic developmental trajectories of writing fluency. In G. G. Fogal & M. H. Verspoor (Eds.), *Complex Dynamic Systems Theory and L2 Writing Development* (Vol. 54, pp. 3–27). Amsterdam John Benjamins Publishing Company.

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

Boersma, P., & Weenink, D. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.11 retrieved 8 April 2022 from http://www.praat.org/.

Borràs, J., & Llanes, À. (2021). Re-examining the impact of study abroad on L2 development: a critical overview. *The Language Learning Journal, 49* (5), 527–540. doi: 10.1080/09571736.2019.1642941

Bulté, B., & Housen, A. (2018). Syntactic complexity in L2 writing: Individual pathways and emerging group trends. *International Journal of Applied Linguistics, 28* (1), 147–164. doi: https://doi.org/10.1111/ijal.12196

Bulté, B., & Housen, A. (2020). A critical appraisal of the CDST approach to investigating linguistic complexity in L2 writing development. In G. G. Fogal & M. H. Verspoor (Eds.), *Complex Dynamic Systems Theory and L2 Writing Development* (pp. 207–238). Amsterdam: John Benjamins Publishing Company.

Bygate, M. (2018). *Learning Language Through Task Repetition*. Amsterdam: John Benjamins Publishing Company.

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

de Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behav Res Methods, 41* (2), 385–390. doi: 10.3758/BRM.41.2.385

Fogal, G. G. (2022). Second language writing from a complex dynamic systems perspective. *Language Teaching, 55* (2), 193–210. doi: 10.1017/S0261444821000410

Freed, B. (1995). What makes us think that students who study abroad become fluent? In B. Freed (Ed.), *Second Language Acquisition in a Study Abroad.* (pp. 123–148). Amsterdam: John Benjamin's.

Hiver, P., Al-Hoorie, A. H., & Evans, R. (2022). Complex dynamic systems theory in language learning: a scoping review of 25 years of research. *Studies in Second Language Acquisition, 44* (4), 913–941. doi: 10.1017/S0272263121000553

Horváth, L., & Rice, G. (2024). *Change Point Analysis for Time Series*. Cham, Switzerland: Springer.

Kanda, M. (2024). A longitudinal study of low-proficiency learners' English oral proficiency development. *JACET Journal, 68*, 27–52.

Kormos, J. (2006). *Speech Production and Second Language Acquisition*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Köylü, Z., Eryılmaz, N., & Pérez-Vidal, C. (2023). A dynamic usage-based analysis of L2 written complexity development of sojourners. *Journal of Second Language Writing, 60*, 101002. doi: https://doi.org/10.1016/j.jslw.2023.101002

Kyle, K., Crossley, S., & Verspoor, M. (2021). Measuring longitudinal writing development using indices of syntactic complexity and sophistication. *Studies in Second Language Acquisition, 43* (4), 781–812. doi: 10.1017/S0272263120000546

Lambert, C., Aubrey, S., & Leeming, P. (2021). Task preparation and second language speech production. *TESOL Quarterly*. doi: 10.1002/tesq.598

Lambert, C., Kormos, J., & Minn, D. (2017). Task repetition and second language processing. *Studies in Second Language Acquisition, 39*, 167–196.

Larsen-Freeman, D. (1994, October). *On the parallels between chaos theory and second language acquisition*. Paper presented at the Second Language Research Forum, McGill University, Montreal, QC, Canada.

Larsen-Freeman, D. (1997). Chaos/Complexity Science and Second Language Acquisition. *Applied Linguistics, 18* (2), 141–165. doi: 10.1093/applin/18.2.141

Larsen-Freeman, D. (2009). Adjusting Expectations: The Study of Complexity, Accuracy, and

Fluency in Second Language Acquisition. *Applied Linguistics, 30* (4), 579-589. doi: 10.1093/applin/amp043

Larsen-Freeman, D. (2018). Task repetition or task iteration? : It does make a difference. In M. Byygate (Ed.), *Learning Language through Task Repetition* (pp. 311-329). Amsterdam: John Benjamin Publishing Company.

Levelt, E. J. M. (1989). *Speaking: from Intention to Articulation*. Cambridge: A Bradford book: the MIT press.

Lowie, W., Caspi, T., van Geert, P., & Steenbeek, H. (2011). Modeling development and change. In M. H. Verspoor, K. de Bot, & W. Lowie (Eds.), *A Dynamic Approach to Second Language Development: Methods and Techniques* (pp. 99-121). Amsterdam: John Benjamins Publishing Company.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics, 15* (4), 474-496. doi: doi 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. doi: https://doi.org/10.5054/tq.2011.240859

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

Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing, 29*, 16-27.

Polat, B., & Kim, Y. (2013). Dynamics of complexity and accuracy: a longitudinal case study of advanced untutored development. *Applied Linguistics, 35* (2), 184-207. doi: 10.1093/applin/amt013

Saito, K., Ilkan, M., Magne, V., Tran, M. N., & Suzuki, S. (2018). Acoustic characteristics and learner profiles of low-, mid- and high-level second language fluency. *Applied Psycholinguistics, 39*, 593-617.

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

Spring, R., & Johnson, M. (2022). The possibility of improving automated calculation of measures of lexical richness for EFL writing: A comparison of the LCA, NLTK and SpaCy tools. *System, 106*, 102770. doi: https://doi.org/10.1016/j.system.2022.102770

Tavakoli, P., Nakatsuhara, F., & Hunter, A. M. (2020). Aspects of fluency across assessed levels of speaking proficiency. *The Modern Language Journal, 104* (1), 169-191. doi: 10.1111/modl.12620

Tavakoli, P., & Wright, C. (2020). *Second Language Speech Fluency*. Cambridge: Cambridge University Press.

Taylor, J. (2000). Change-Point Analysis: A powerful tool for detecting changes. Retrieved from http://www.variation.com/cpa/tech/changepoint

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17* (1), 84–119.

Tsushima, T. (2018). An exploratory case study: improvement of the ability to produce the rhythmic properties of English in a spontaneous speech task through long-term, individual-based speech training. *Journal of Humanities & Natural Sciences, 143*, 15–41.

Tsushima, T. (2019). An exploratory case study on development of fluency and English speech rhythm through individual speech training and a study abroad program. *Journal of Humanities & Natural Sciences, 145*, 29–63.

Tsushima, T. (2020). Development of fluency and English speech rhythm through individual-based, long-term speech training in the formal instruction settings. *Journal of Humanities & Natural Sciences, 147*, 59–93.

Tsushima, T. (2021). A case study: longitudinal development of oral fluency in production of monologic narratives among two adult Japanese learners of English. *Journal of Humanities & Natural Sciences* (149), 69–92.

Tsushima, T. (2023). An interim report of a case study: longitudinal development of L2 oral fluency during a study abroad program. *Journal of Humanities & Natural Sciences, 152*, 61–85.

Valls-Ferrer, M., & Mora, J. C. (2014). L2 fluency development in formal instruction and study abroad: The role of initial fluency level and language contact. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts* (pp. 111–136). Amsterdam: John Benjamins Publishing Company.

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. doi: https://doi.org/10.1111/j.1540-4781.2008.00713.x

Verspoor, M. (2017). Complex Dynamic Systems Theory and L2 pedagogy. In L. Ortega & Z. Han (Eds.), *Complexity Theory and Language Development: In celebration of Diane Larsen-Freeman* (pp. 143–162). Amsterdam: John Benjamins Publishing Company.

Verspoor, M., & Behrens, H. (2011). Dynamic Systems Theory and a usage-based approach to second language development. In M. H. Verspoor, K. de Bot, & W. Lowie (Eds.), *A Dynamic Approach to Second Language Development: Methods and techniques* (pp. 25–38). Amsterdam: John Benjamins Publishing Company.

Verspoor, M. H., Bot, K., & Lowie, W. (2011). *A Dynamic Approach to Second Language Development.* Amsterdam: John Benjamins.

Wang, Y., & Tao, S. (2020). The dynamic co-development of linguistic and discourse-semantic complexity. In G. G. Fogal & M. H. Verspoor (Eds.), *Complex Dynamic Systems Theory and L2 Writing Development* (pp. 49–80). Amsterdam: John Benjamins Publishing Company.

Zaytseva, V. (2016). Vocabulary acquisition in study abroad and formal instruction: an investigation on oral and written lexical development. (PhD). Universitat Pompeu Fabra (UPF), Barcelona. Retrieved from http://hdl.handle.net/10803/387120