

# A Corpus-based Investigation of Syntactic Complexity, Fluency, Sentence Variety, and Sentence Development in L2 Genre Writing

Nicolai Struc and Nicholas Wood

## **Abstract**

*Measures of syntactic complexity have been used to evaluate task-related variation and pedagogic interventions in L2 writing production, and to assess differences across proficiency levels and over time. This paper reports on a corpus-based investigation of syntactic complexity and fluency in narrative and argumentative writing, comparing texts produced by 170 L2 learners at the starts of their first and second years at university. Conventional metrics were used to compare syntactic complexity and fluency in the two sets of samples; novel methods were also devised to examine possible changes in sentence variety and the development of sentence construction. Significant differences within the first and second year corpora reflected responses to the genre-specific demands on text production, but length of instruction only significantly impacted on narrative writing. Argumentative texts presented significantly greater syntactic complexity but, in sharp contrast to the findings of*

*previous studies, also evidenced significantly greater fluency. The Sentence Variety Index and sentence reconstruction both offered insights into writing production, suggesting the value of the suite of metrics used in this research to the longitudinal study of L2 genre writing.*

*Acknowledgement:* This research was made possible through a research grant provided by Reitaku Linguistic Research Center (Reitaku University)

## **Overview**

This paper outlines the value of studying syntactic complexity in L2 written texts and briefly describes how the phenomenon has been defined and how it has been measured. Making comparisons with previous studies is problematic, but reviews and research suggest certain patterns of development in syntactic complexity (as measured by ratios) and fluency (as measured by mean length of units). As the present research seeks to investigate possible differences in the construction of written genre, a short description of this facet of language is offered, with a summary of previous studies comparing argumentative and narrative texts. The construct *sentence variety* is introduced, and its role as a tool for analyzing structural complexity explained. The aims of this study are followed by details of the corpus-based research design and our tentative expectations. After the results, we conclude with a discussion of our findings, the limitations of the present study, and a consideration of issues raised and future directions for research.

## **Syntactic Complexity**

Syntactic (or grammatical) complexity has attracted interest from a wide range of disciplines within linguistics and cognitive sciences (Voss, 2005). In the field of second language (L2) writing studies, its development has been considered integral to target language acquisition (Lu, 2010), with complexity metrics utilized to examine the effects of specific pedagogic interventions and task differences, to investigate differences between groups of writers, or to simply describe the texts produced by a group in order to inform instruction (Polio, 2001).

Broadly speaking, *syntactic complexity* refers to the regularized patterns by which a language is sequenced and structured - “the way words are combined to form sentences” (Nunberg, Briscoe, & Huddleston, 2002, p. 1728) - and the intrinsic potential of such patterning to engender textual forms that range from the simple to the complex, multi-componential. As an integral feature of language production, syntactic complexity can, initially, be considered from two perspectives: as an external product of phonological or orthographic text, and the internal process of formulation and comprehension of linguistic forms. Szmrecsányi (2004) regards the external as the “formal”, that is, users must construct and produce material that conforms to the syntactic conventions of the language. Syntactic complexity can, in this sense, be viewed as “the range of forms that surface in language production and the degree of sophistication of such forms” (Ortega, 2003, p. 492). Moreover, it is assumed that with greater proficiency, progressively more elaborate language may be used, as well as a greater variety of syntactic patterning (Foster & Skehan, 1996). The internal construction and comprehension of syntactic complexity appears predicated on knowledge, experience, proficiencies, and the motivation, albeit conditional and variable (Dörnyei

& Skehan, 2005), of learners to restructure language as more complex subsystems are developed (Foster & Skehan, 1996).

## **Measuring Complexity: T-unit or Sentence?**

Word and clause frequencies per unit of text and mean lengths of text unit have been used in many studies of syntactic complexity in both L1 and L2 written texts (e.g., Homburg, 1984; Hirano, 1991; Casanave, 1994; Ishikawa, 1995; Beers & Nagy, 2009). The analysis of clauses within text, however, raises the possibility of two different contextual units to be used: the T-unit and the sentence. The T-unit, or “minimal terminable unit” was defined by Hunt (p. 49, 1965) as consisting of “one main clause plus the subordinate clauses attached to or embedded within it”, and has been widely used in studies of L2 writing to investigate the ability of learners “to exploit the embedding processes available in the target language” (Sharma, 1980, p.320). In addition, the unit offers a means of examining syntactic complexity across the sentence boundaries indicated by L2 writers, so that subordinated clause fragments (and non-clause but semantically related text such as lists and examples) can be included in the syntactic structure of a T-unit in a preceding or following orthographic sentence. It is the “objective” appraisal of text facilitated by T-unit analysis, one that disregards the potential idiosyncrasies of punctuation, that has perhaps led to its wide use as a basis for complexity ratios, including clauses per T-unit (C/TU) (e.g., Hirano, 1991), dependent clauses per T-unit (DC/TU) (e.g., Homburg, 1984), adverbial clauses per T-unit (AC/TU) (Cooper, 1976), T-units per sentence (TU/S) (e.g., Ishikawa, 1995), and passives per T-unit (P/TU) (Kameen, 1979).

In contrast, Bardovi-Harlig and Bofman (1988) argue that the

sentence rather than the T-unit provides a “superior” unit for the analysis of syntactic complexity. They suggest that sentence-based analysis better characterizes learner knowledge; it facilitates comparisons of learners and stages of language development (enabling a better comparison of the shift from the use of coordination by beginner learners to the use of embedding by more advanced learners), and it encourages dialogue between teachers and researchers by providing a common learner-centred unit of analysis and access to data in the same form. The T-unit, they argue, eliminates full-clause coordination from any quantitative description of syntactic development. “T-unit analysis artificially divides sentences which learners see as syntactic units, imposing uniformity of length and complexity on output which is not present in the original production by restricting units to single main clauses” (1988, p. 5).

Both T-unit and sentence have merits for analysis, and one resolution to the issue of which is the optimal unit is to use both.

## **Previous Studies**

Discerning trends among or making comparisons with other studies is problematic. Lack of computational tools in the past limited the number of measures used and the number of samples analyzed (Lu, 2011). Sample sizes vary considerably; the six longitudinal studies analysed by Ortega (2003) range in sample size from four (Casanave, 1994) to 73 (Kern & Schultz, 1992). There is little agreement on the definitions of measures and considerable variety in task type, time allowed, sample size, corpus length and statistical treatments (Ishikawa, 1995; Lu, 2011). Both Polio (1997) and Wolfe-Quintero, Inagaki, and Kim (1998) draw attention to the different ways units of production have been defined in the calculation of syntactic complexity. Little attention appears to have been

paid to the effects of L1 on syntactic complexity, with L2 learners in heterogeneous groups treated as if they were from the same L1 background (Lu, 2011). Ishikawa (1995), Navés, Torras, and Celaya (2003), and Torras, Navés, Celaya, and Pérez-Vidal (2006) indicate that some measures appear better at gauging young and low-proficiency learners, and their use precludes the comparison of results with older or more experienced learners (Navés, 2006).

Furthermore, although Ortega (2003) considers mean length of unit (clause, sentence and T-unit) as a syntactic complexity metric, Wolfe-Quintero, et al. (1998) argue that length of production is more appropriately considered a measure of fluency, fluency being an indication “that more words and more structures are accessed in a limited time” (p. 25).

## **Reported Trends in Syntactic Complexity and Fluency**

From a review of 39 L2 writing studies, Wolfe-Quintero, et al. (1998), while cautioning that some measures failed to differentiate between adjacent levels of proficiency, report that the mean length of T-unit (MLTU), mean length of clause (MLC), mean length of error-free T-unit (MLEFTU), C/TU, dependent clauses per clause (DC/C), and DC/TU, “consistently increased in a linear relationship to proficiency level across studies”. However, Ortega (2003) questions the authors’ practice of taking a vote-count of significant results across studies as it tends to ignore differences in research methodologies at the expense of statistical validity. In Ortega’s own synthesis of results from 21 college L2 studies (2003) it is proposed that C/TU can differentiate between college-level L2 writing groups. For substantial changes in the syntactic complexity of L2 writing as measured by MLTU to be observed, a period

of roughly a year of instruction is required. Acknowledging that the small set of longitudinal studies in the review could only yield exploratory findings, Ortega (2003) does suggest that statistically significant differences in MLTU can be tentatively expected between proficiency levels in cross-sectional studies. Lu (2011) used a set of 14 metrics to assess and compare syntactic complexity in 3,554 texts of L2 college writers at four proficiency (school) levels and found seven (MLC, complex nominals per clause (CN/C), mean length of sentence (MLS), complex nominals per T-unit (CN/TU), MLTU, coordinated phrases per clause (CP/C), and coordinated phrases per T-unit (CP/TU)) showed a linear increase across the four levels. Five measures (MLC, CN/C, MLS, CN/TU, and MLTU) discriminated between adjacent levels, while three measures (clauses per sentence (C/S), dependent clauses per sentence (DC/S), and DC/TU) not only discriminated between adjacent levels but also decreased significantly from lower to higher levels.

## Genre

Language, from an emergentist perspective, is considered social in nature, a cultural artefact that functions to achieve social affiliations and actions within contexts of language use and which is passed on, subject to diachronic change, to succeeding generations (MacWhinney, 1998; Lee & Schumann, 2005; Larsen-Freeman, 2006). At the interface of individual language behaviour and the sociolinguistic terrain, forms of discourse or *genres* have emerged to facilitate communication, to realize interests, and to effect action. As a classificatory construct (Bauman, 1992) genres offer expectations of the type of linguistic engagement (Guenther & Knoblauch, 1995), possess features that include stability and name recognition (Swales, 1990), and in written language present

“differentiated and identifiable written text types” (Purcell-Gates, Duke, & Martineau, 2007, p. 11). As models for communicative action, genres provide routinized, historically-derived conventions that facilitate the interaction of author and audience (Hanks, 1987; Guenther & Knoblauch, 1995). As socially recognized forms of language practice they reflect social norms and expectations, and function to serve specific social purposes (Halliday & Hasan, 1985; Swales, 1990; Hyland, 2003; Purcell-Gates, et al., 2007). Consequently, as users actively engage in the construction of genre with social intent, it is the social, pragmatic function of text that drives the structuring of the written form (Waugh, 1995; Purcell-Gates, et al., 2007).

Studies of L1 written texts suggest a strong relationship between conventionally defined genres and syntactic complexity as measured by mean length of text (MLT), MLTU and C/TU. The latter two measures have been found to be significantly greater in the argumentative essays of younger writers than in their narrative texts (Crowhurst and Piche, 1979; Stomberg and Kurth, 1982; Beers and Nagy, 2009), supporting the suggestion (Crowhurst, 1980; Beers and Nagy, 2009) that the social imperatives of argumentation and persuasion require the writer to more frequently convey complex relationships between ideas (e.g., causality), and thus produce a greater proportion of subordinate clauses, and hence longer and more complex T-units. Beers and Nagy (2009) also found that narrative texts were significantly longer than persuasive essays, but there was no significant difference in the mean length of clause. A study of L2 texts by Yau and Belanger (1984) found similar relationships between genre and complexity: Learners wrote longer narrative compositions, and expository texts were significantly more complex than narrative texts as measured by MLC and MLTU, and approached significance for C/TU. Lu

(2011) found that argumentative essays exhibited significantly greater syntactic complexity than narrative essays (as measured by 14 metrics), and untimed argumentative essays showed greater complexity than timed.

## **Advanced Complexity**

Longitudinal patterns of development suggest the focus of syntactic complexity appears to shift from the clause to the phrase, and fluency is achieved at the expense of subordination. Hunt's study (1965) showed that advanced writers produced significantly longer T-units and this was not due to gains in subordinate clauses but in gains in clause length. Wolfe-Quintero, et al. (1998) note that adverbial, adjective, and nominal clauses can all be reduced to phrases and suggest that more advanced writers tend to use more reduced forms (thus reducing C/TU). De Haan (1987) found that while more formal texts show greater syntactic complexity, this complexity is brought about by embedding relatively simple structures into larger ones and is typically achieved by means of (short) prepositions and subordinators, with the result of a decrease in mean word length [of essays]. Lu (2010) argues that more advanced L2 writers tend to produce longer clauses and T-units, not as a result of an increased use of dependent clauses or complex T-units, but as a result of increased use of complex phrases such as coordinate phrases and complex nominals. Lu (2011) found that three clausal measures (see above) decreased significantly from lower to higher levels, and he suggests that "as students advance to higher levels of proficiency, they learn to capitalize on complexification more at the phrasal level and less at the clausal level" (p. 57). Thus, units of production increase in length as clausal structuring decreases.

## Sentence Variety

The term *sentence variety* is used here to describe the variety of sentence types deployed by a writer within one text. There are four basic sentence types, the typology determined by evident clausal structuring: simple, complex, compound, and compound-complex. A text that presents a limited range of or skewed distribution of types can be said to show less variety, while a text presenting a greater range and a more equal distribution of types can be said to show greater variety.

Polio (2001) suggests that “at an advanced level, too many complex sentences may be a problem and thus at some point, variety may be important to quality.” (p.97). Beers and Nagy (2009) note the importance of sentence variety for readability, for “it is the variety of sentence structure, not complexity of sentence structure, that makes texts flow” (p. 187). The assumption for this aspect of our research is based Foster and Skehan’s proposition (1996): that, as L2 learners develop “more complex subsystems of language” and “more elaborate language”, they will show “a greater variety of syntactic patterning” in language production. While Foster and Skehan investigated complexity in spoken language by analysing of clauses/c-unit ratios and structural variety by focusing on the use of tense, aspect, voice and modality, it is suggested here that sentence variety in L2 written text can also be considered to be a reflection of structural complexity, and a learner’s willingness to attempt more elaborate language as more linguistic resources become available. It is not assumed that there is a development from the use one type of sentence to another, or that the aim of writers should be the equal use of all four sentence types. Sentence variety, it is suggested, is a reflection of syntactic complexity at one explicit level of text production, where

orthographic conventions intersect with linguistic expression.

## **Aims**

The overall aim of this research is to examine, describe and compare evidence of syntactic complexity, fluency, sentence variety, and sentence development in texts written in two genre produced by 170 L2 learners at the start of their first year at university and at the start of their second year. It is hoped that the results of this study may be used to inform curriculum development and pedagogic methodologies, and contribute to current research in this field.

## **Research Design**

### **Corpus**

“All tools are designed to simplify some task” (Miller & Page, 2007, p. 62), and the task of analysing and comparing hundreds of discrete digitized texts is expedited by current software and the essential qualities of a corpus, which include its “machine readability, authenticity and representativeness” (McEnery, Xiao, & Tono, 2006, p. 5). Moreover, quantitative data derived from a corpus-based study can serve to strike a balance between intuitive notions of what is common or typical of certain language varieties and observed actual language use (Oostdijk & de Haan, 1994), while according to Rimmer (2006) “a corpus-informed study of complexity has the potential to reconcile the tension in theoretical linguistics between grammar as being sentence-bound and grammar as discourse” (p. 497).

This study interrogates a corpus initially constructed by Struc and Wood (2009) and extended one year later (Struc & Wood, 2010), with the

methods of data collection remaining constant. The data is comprised of two written texts produced by each L2 learner at the start of their first year in the English Writing programme (EWP) and two written texts produced by the same L2 learners in response to the same prompts at the start of their second year in the programme.

The two texts, one narrative and the other argumentative, were produced in controlled, time-limited conditions. During a period of 40 minutes in the first class of the year, learners were allowed 20 minutes to produce a response to one prompt, and 20 minutes for a response to a second prompt. Texts were written by hand and without use of reference materials, such as a dictionary. Instructions and writing prompts were presented in L1 (Japanese) but are presented here in an English translation:

Prompt 1- Narrative: “Imagine two friends went shopping together last week. One friend returned home happy, the other friend returned home sad. Write a story about what happened. You have 20 minutes.”

Prompt 2- Argumentative: “Studying English abroad. Please write reasons for and against studying English in another country. You have 20 minutes.”

Learners were provided with an explanation of the general goals of the research, a request to participate, and an assurance of anonymity. Each text, later transcribed to electronic text (txt) file, was accordingly coded to hide the learner’s identity, the same code used in the following year in order to match other texts produced by the same learner.

## **Population and Educational Environment**

The population was comprised of all L2 learners at a mid-level private university in Japan enrolled in the EWP, with the exclusion from analysis of learners who were absent on either occasion when data was collected ( $n = 170$ ). The EWP specifically aimed to foster academic writing skills but allowed a range of approaches and variety of texts to be used by instructors delivering the course once a week over two 15-week semesters (amounting to 45 hours of tuition). All the learners also received instruction in English language communication skills in two classes a week, with higher level instruction delivered by English native speakers (NS), lower levels delivered by both NS and Japanese English speakers, and with a similar variety of texts and approaches used by educators. Learners' experience of English included living and studying in English-speaking communities abroad, attendance of private English conversation classes, contact with NS assistant language teachers in secondary education, English language classes at junior and senior high schools, and experience of English language cultural artefacts (films, songs, websites, etc.). The educational environment, and to a lesser degree the language experiences of the population, can, thus, be characterized as one of diversity.

## **Units and Metrics**

**Orthographic and reconstructed sentences.** To investigate complexity and fluency within student-produced texts, and within and across orthographic boundaries, two types of sentence structure were investigated: *orthographic* and *reconstructed* sentences. The orthographic sentence (OS) recognizes and respects the individuality of each writer's text as produced, an OS being "a unit of writing that begins with a capital

letter and ends with a full stop, question mark, or exclamation mark” (Nunberg, et al., 2002, p.1728). The reconstructed sentence (RS) is a unit that combines a preceding and/or a following OS that is syntactically related to a main OS and is included in a T-unit that crosses OS boundaries (e.g., *Because* fragments and *For example* non-clause lists). For instance, [*I can study other things too. For example, culture, food, music, sports, building and so on.*] is comprised of two OS, one T-unit, and one RS. The reconstruction of text into RS and division into T-units provide a means by which text can be “objectively” demarcated within and across learner orthography in order to investigate indicators of complexity “wherever they occur” (Voss, 2005).

**Syntactic complexity.** Five syntactic complexity ratio metrics were employed: clauses per orthographic sentence (C/OS) as a sentence complexity ratio (Ishikawa, 1995), C/TU as a measure of depth of clauses (Wolfe-Quintero, et al., 1998), DC/T and DC/C as measures of subordination (Wolfe-Quintero, et al., 1998), and T-units per reconstructed sentence (TU/RS), as an adaptation of the conventional measure of coordination reported by Wolfe-Quintero, et al. (1998).

**Fluency.** Five fluency metrics were employed, each indicating the mean number of words in a production unit: mean length of orthographic sentence (MLOS), mean length of reconstructed sentence (MLRS), MLC, MLTU, and MLT. MLRS is, however, not a straightforward measure of fluency as it is mediated by the proportion of fragments in a text. A decrease in the number of OS fragments, for example, will result in a decrease in the MLRS relative to MLOS (which is determined only by text length and total number of OS).

**Sentence variety.** A statistically-based Sentence Variety Index (SVI) was devised and used for analysing each text. The index ranges from 0 to 100, 0 indicating no variety (i.e., all sentences of one type) and 100 indicating maximum variety (i.e., all four sentence types equally represented).

## **Analyses**

Four statistical comparisons to be undertaken: two year-in-program comparisons (narrative second year compared to narrative first year, argumentative second year compared to argumentative first year), and two task comparisons (narrative first year compared to argumentative first year, narrative second year compared to argumentative second year). Each comparison examines results for syntactic complexity, fluency, sentence variety, and sentence development.

## **Expectations**

The research has the following four tentative expectations:

1. After a year of instruction, it is expected that the texts of L2 learners will, in general, show evidence of a) increased syntactic complexity as measured by ratios per unit, b) increased fluency in longer units of production and longer text lengths, and c) increased sentence variety.
2. The exposition of genre makes specific demands on the content and construction of text, and this will be reflected in distinct differences in a) syntactic complexity, b) fluency, and c) sentence variety.

3. While the expectation is that most measures will increase, it is possible that advanced complexity might be achieved (consistent with Lu, 2011) by a decrease in some clausal measures (C/OS and DC/TU).
4. After a year of instruction, it is expected that MLOS and MLRS will converge, with MLOS measures increasing as a function of greater fluency, but MLRS measures exhibiting a relatively smaller increase as fewer fragments occur in texts.

## Results

The larger corpus is comprised of four writing samples from each learner which comprise 4 different subcorpora (1. First year narrative; 2. First year argumentative; 3. Second year narrative; 4. Second year argumentative). These were compared using the metrics described previously using paired samples t-tests. For the purpose of clear presentation, the results will be organized into the following four groups: 1) year in program (narrative); 2) year in program (argumentative); 3) first year narrative/argumentative; 4) second year narrative/argumentative.

The results in each group are further subdivided into two types of measures: 1) those that examine change in syntactic complexity (ratios of units) and 2) those that reveal change in fluency (number of words per unit, e.g., sentence, T-unit, clause). Below is a summary of the initials used in the tabular presentation of data:

Complexity	Fluency
TU/S – T-units per sentence	MLOS – Mean length of orthographic sentence
C/TU – Clauses per T-unit	MLTU – Mean length of T-unit
C/OS – Clauses per orthographic sentence	MLRS – Mean length of reconstructed sentence
DC/TU – Dependent clauses per T-unit	MLC – Mean length of clause
DC/C – Dependent clauses per clause	MLT- Mean length of text

### Year-in-program Effects on Narrative Writing

Table 1 shows differences in complexity as measured by the ratios of various syntactic units. Paired t-tests showed significant differences in the mean number of clauses per orthographic sentence and dependent clauses per clause in narrative writing in the second year sample compared with the first. The other measures, while all exhibiting increases between the first and second year, did not reach statistical significance.

Table 1  
*Year-in-program Effects on Complexity (Narrative)*

measure	year	Mean	SD	t	df
C/TU	1	1.16	.195	1.799	169
	2	1.19	.178		
DC/TU	1	0.16	.195	1.799	169
	2	0.19	.178		
DC/C	1	0.12	.119	2.439*	169
	2	0.15	.113		
C/OS	1	1.37	.448	2.199*	169
	2	1.45	.365		
TU/S	1	1.22	.290	.949	169
	2	1.25	.219		

Note. \*= $p \leq .05$ , \*\*= $p \leq .01$ , \*\*\*= $p \leq .001$

Table 2 shows the mean lengths of texts, T-units, orthographic sentences, reconstructed sentences and clauses in first and second year narrative writing. In paired t-tests, significant differences were observed in increases of mean word length in all measures with the exception of mean length of clause. Mean length of clause did show an increase but it failed to reach statistical significance

Table 2

*Year-in-program Effects on Fluency (Narrative)*

measure	year	Mean	SD	t	df
MLOS	1	7.69	2.75	2.999**	169
	2	8.35	2.58		
MLRS	1	8.11	3.00	2.257*	169
	2	8.66	2.59		
MLTU	1	6.71	1.58	2.149*	169
	2	7.00	1.41		
MLC	1	5.79	1.00	.954	169
	2	5.87	.910		
MLT	1	72.69	39	5.672***	169
	2	86.52	38.4		

Note. \*= $p \leq .05$ , \*\*= $p \leq .01$ , \*\*\*= $p \leq .001$

## **Year-in-program Effects on Argumentative Writing**

Table 3 shows differences in complexity as measured by the ratios of various syntactic units. Paired t-tests showed no significant differences in the number of clauses per orthographic sentences, dependent clauses per clause or T-units per sentence, although increases were observed. The mean number of clauses and dependent clauses per T-unit showed declines although they did not reach statistical significance.

Table 3  
*Year-in-program Effects on Complexity (Argumentative)*

measure	year	Mean	SD	t	df
C/TU	1	1.39	.357	-.542	165
	2	1.37	.264		
DC/TU	1	0.39	.357	-.542	165
	2	0.37	.264		
DC/C	1	0.24	1.55	.173	165
	2	0.24	1.35		
C/OS	1	1.46	.416	1.366	165
	2	1.51	.402		
TU/S	1	1.13	.218	.697	165
	2	1.15	.241		

Note. \*= $p \leq .05$ , \*\*= $p \leq .01$ , \*\*\*= $p \leq .001$

Table 4 shows the mean lengths of texts, T-units, orthographic sentences, reconstructed sentences and clauses in first and second year argumentative writing. Only the increase in mean length of texts and orthographic sentences were statistically significant. While the mean length of reconstructed sentences and clauses showed increases, the mean length of T-units declined in the second year sample, although not significantly.

Table 4  
*Year-in-program Effects on Fluency (Argumentative)*

measure	year	Mean	SD	t	df
MLOS	1	9.25	3.15	2.357*	169
	2	9.84	2.85		
MLRS	1	10.04	3.58	-1.006	169
	2	10.33	2.95		
MLTU	1	9.41	2.82	-.160	165
	2	9.37	2.23		
MLC	1	6.82	1.43	.423	165
	2	6.88	1.23		
MLT	1	77.19	43.02	5.846***	169
	2	93.02	43.03		

Note. \*= $p \leq .05$ , \*\*= $p \leq .01$ , \*\*\*= $p \leq .001$

## Task Effects in First Year Writing

Table 5 shows differences in complexity between first year writing in narrative and argumentative genre as measured by the ratios of various syntactic units. Paired t-tests showed consistent significant differences in all the measures. All the ratios showed higher values in argumentative writing with the exception of the mean number of T-units per sentence, which was significantly lower in argumentative writing.

Table 5

### *Task Effects on Complexity (First year)*

measure	genre	Mean	SD	t	df
C/TU	narrative	1.17	.195	7.694***	166
	argumentative	1.39	.356		
DC/TU	narrative	0.17	.195	7.694***	166
	argumentative	0.39	.356		
DC/C	narrative	0.12	.119	8.774***	166
	argumentative	0.24	.155		
C/OS	narrative	1.37	.450	2.224*	166
	argumentative	1.46	.415		
TU/S	narrative	1.23	.290	-3.724***	166
	argumentative	1.13	.218		

Note. \*= $p \leq .05$ , \*\*= $p \leq .01$ , \*\*\*= $p \leq .001$

Table 6 shows the mean lengths of T-units, orthographic sentences, reconstructed sentences and clauses in first year narrative and argumentative writing and the results of the comparison of these means in paired t-tests. All measures showed significant differences between the two genres with argumentative writing exhibiting consistently greater values in all measures.

Table 6  
*Task Effects on Fluency (First year)*

measure	genre	Mean	SD	t	df
MLOS	narrative	7.69	2.75	6.254***	169
	argumentative	9.25	3.15		
MLRS	narrative	8.11	3	6.510***	169
	argumentative	10.04	3.58		
MLTU	narrative	6.74	1.59	11.712***	166
	argumentative	9.40	2.81		
MLC	narrative	5.79	1	8.189***	166
	argumentative	6.81	1.43		
MLT	narrative	72.69	39	1.781	169
	argumentative	77.19	43		

Note. \*= $p \leq .05$ , \*\*= $p \leq .01$ , \*\*\*= $p \leq .001$

## Task Effects in Second Year Writing

Table 7 shows differences in complexity between second year writing in narrative and argumentative genre as measured by the ratios of various syntactic units. Paired t-tests showed significant differences in all the measures except for the number of clauses per orthographic sentence. As in the first year genre/task effect analysis (see table 5), all the ratios showed higher values in argumentative writing with the exception of the mean number of T-units per sentence, which was again significantly lower in argumentative writing. In contrast with the fluency measures compared between first and second year (see tables 6 and 8), the magnitude of the difference appears to diminish slightly in most measures (C/TU, C/OS, DC/TU, DC/C) with the exception of the number of T-units per sentence whose contrast between the genres appears more pronounced in the second year samples.

Table 7

*Task Effects on Complexity (Second year)*

measure	genre	Mean	SD	t	df
C/TU	narrative	1.20	.177	7.454***	167
	argumentative	1.37	.263		
DC/TU	narrative	0.20	.177	7.454***	167
	argumentative	0.37	.263		
DC/C	narrative	0.15	.112	7.552***	167
	argumentative	0.24	.135		
C/OS	narrative	1.45	.364	1.783	167
	argumentative	1.50	.400		
TU/S	narrative	1.25	.218	-5.267***	167
	argumentative	1.14	.240		

Note. \*= $p \leq .05$ , \*\*= $p \leq .01$ , \*\*\*= $p \leq .001$

Table 8 shows the mean lengths of T-units, orthographic sentences, reconstructed sentences and clauses in second year narrative and argumentative writing and the results of the comparison of these means in paired t-tests. All measures showed significant differences

Table 8

*Task Effects on Fluency (Second year)*

measure	genre	Mean	SD	t	df
MLOS	narrative	8.35	2.58	7.828***	169
	argumentative	9.84	2.85		
MLRS	narrative	8.66	2.60	7.604***	169
	argumentative	10.33	2.95		
MLTU	narrative	7.01	1.49	12.751***	167
	argumentative	9.36	2.22		
MLC	narrative	5.86	.90	9.729***	167
	argumentative	6.88	1.22		
MLT	narrative	86.52	38.40	2.799*	169
	argumentative	93.02	43.03		

Note. \*= $p \leq .05$ , \*\*= $p \leq .01$ , \*\*\*= $p \leq .001$

between the two genres with argumentative writing exhibiting consistently greater values in all measures. The observed differences appear to have increased in magnitude consistently from the first year (see Table 4).

### Sentence variety

The mean distribution of sentence types (simple, compound, complex and compound-complex) in each subcorpus are presented in Table 9.

Table 9

*Sentence Variety: Sentence Type Distribution by Year and Genre*

Sentence type	narrative				argumentative			
	1 <sup>st</sup> year		2 <sup>nd</sup> year		1 <sup>st</sup> year		2 <sup>nd</sup> year	
	%(M)	SD	%	SD	%(M)	SD	%(M)	SD
simple	71.18	26.35	65.04	24.68	64.19	25.88	60.08	24.49
complex	10.37	16.41	14.10	15.24	25.11	22.21	28.35	20.51
compound	14.26	17.18	16.45	15.30	6.48	12.38	7.40	12.59
compound-complex	4.18	12.87	4.41	8.58	4.21	11.05	4.17	8.32
total	100		100		100		100	

### Sentence Variety Index

Table 10 shows the mean SVI values for each subcorpus and the results of paired-sample t-tests between years for each task and between tasks for each year. Significant differences in sentence variety were observed between first and second year narrative writing and between first and second year argumentative writing samples. Both comparisons showed increasing sentence variety in second year writing. A significant

difference in sentence variety was observed between first year narrative and argumentative writing with argumentative writing showing greater sentence variety. No significant change was observed between second year narrative and argumentative writing.

Table 10

*Sentence Variety Index Comparison – Year in Program and Task Effects*

	Year in program	genre	M	SD	M (difference)	t	df																																
Year effect narrative	1	N	27.54	22.37	9.53	5.157***	169																																
	2	N	37.07	22.90				Year effect argumentative	1	A	32.48	20.66	5.52	2.900**	165	2	A	38.00	21.20	Task effect First year	1	N	27.92	22.37	4.51	2.465*	166	1	A	32.43	20.66	Task effect Second Year	2	N	37.51	22.90	.37	.190	167
Year effect argumentative	1	A	32.48	20.66	5.52	2.900**	165																																
	2	A	38.00	21.20				Task effect First year	1	N	27.92	22.37	4.51	2.465*	166	1	A	32.43	20.66	Task effect Second Year	2	N	37.51	22.90	.37	.190	167	2	A	37.88	21.20								
Task effect First year	1	N	27.92	22.37	4.51	2.465*	166																																
	1	A	32.43	20.66				Task effect Second Year	2	N	37.51	22.90	.37	.190	167	2	A	37.88	21.20																				
Task effect Second Year	2	N	37.51	22.90	.37	.190	167																																
	2	A	37.88	21.20																																			

Note. \*= $p \leq .05$ , \*\*= $p \leq .01$ , \*\*\*= $p \leq .001$

## MLOS and MLRS Convergence

Table 11 shows the mean lengths of orthographic sentences and reconstructed sentences in both the narrative and the argumentative writing tasks in the first and second years. T-tests comparing the means showed significant differences for the means in all four cases, suggesting that while the mean differences between the first year and second year writing in each task become smaller they still remain significantly different and thus do not completely converge.

Table 11

*Comparison of MLOS and MLRS for Argumentative and Narrative in Year 1 and 2*

Year/task	measure	M	SD	M (difference)	t	df																										
1 <sup>st</sup> year narrative	MLOS	7.69	2.75	.42	5.90***	169																										
	MLRS	8.11	3.00				2 <sup>nd</sup> year narrative	MLOS	8.35	2.58	.31	6.10***	169	MLRS	8.66	2.59	1 <sup>st</sup> year argumentative	MLOS	9.24	3.15	.79	6.86***	169	MLRS	10.04	3.58	2 <sup>nd</sup> year argumentative	MLOS	9.84	2.85	.49	6.51***
2 <sup>nd</sup> year narrative	MLOS	8.35	2.58	.31	6.10***	169																										
	MLRS	8.66	2.59				1 <sup>st</sup> year argumentative	MLOS	9.24	3.15	.79	6.86***	169	MLRS	10.04	3.58	2 <sup>nd</sup> year argumentative	MLOS	9.84	2.85	.49	6.51***	169	MLRS	10.33	2.95						
1 <sup>st</sup> year argumentative	MLOS	9.24	3.15	.79	6.86***	169																										
	MLRS	10.04	3.58				2 <sup>nd</sup> year argumentative	MLOS	9.84	2.85	.49	6.51***	169	MLRS	10.33	2.95																
2 <sup>nd</sup> year argumentative	MLOS	9.84	2.85	.49	6.51***	169																										
	MLRS	10.33	2.95																													

Note. \*= $p \leq .05$ , \*\*= $p \leq .01$ , \*\*\*= $p \leq .001$

## Discussion

A limitation of the present research is that the data was collected at only two points a year apart. Little or nothing can be said of patterns of L2 development during the intervening year of individual learners who contributed to the corpus. Nor can the results be considered the effect of a single, controlled pedagogic intervention in the education of a homogeneous L2 population. Learners were exposed to a range of L2 experiences prior to attending the EWP, whilst during the program, apart from adhering to the aims of a core curriculum, instructors were largely free to adopt their own approaches and texts. Notwithstanding these limitations, two factors favour the value of the study and its results. Firstly, a single year of tuition has been suggested by Ortega (2003) to be sufficient to see evidence of substantial changes in syntactic complexity. And secondly, the corpus is built from the texts from the entire EWP

population (170 L2 learners), and this sample size affords a strong degree of confidence in the validity of any observed trends and/or group differences.

The results themselves present a complicated picture, some in accord with expectations, others in complete contrast. After a year of tuition, it was anticipated that there would be increases in all syntactic complexity measures, with the possible exception of C/OS and DC/TU (following Lu, 2011). However, there was little significant change in the texts of either genre. Of the five measures, only C/OS and DC/C increased significantly in narrative writing. Whilst most measures did indicate gains, in argumentative texts C/TU and DC/TU showed modest decreases. In contrast, fluency measures after a year broadly matched expectations. In narrative writing, there were significant gains in all measures, with the exception of MLC. In argumentative texts, there were small gains in MLRS and MLT, significant gains in MLOS and MLT, and, against the trend, a slight decrease in MLTU. Similarly, sentence variety also presented gains after a year. The second year SVIs for narrative and argumentative (37.07 and 38.00 respectively) indicate significant increases for both genres and evidence of a greater range and balanced distribution of sentence types used. Although apparently similar, these SVIs are problematic as they fail to distinguish the differences in the types of sentence used. This limitation will be explored below.

How are we to interpret these mixed results for the year? In part, they seem to represent genre-specific responses to a move away from learner reliance on simple sentences. Both narrative and argumentative texts show a roughly four percentage decrease in the use of simple sentences, with corresponding increases in the use of complex and compound sentences. In narrative writing, the production of fewer simple

sentences and more complex and compound sentences is reflected in a significant increase in C/OS, gains in C/TU and DC/TU, and significant increases in MLOS, MLRS, and MLTU. In argumentative texts, on the other hand, C/TU, DC/TU and MLTU all decrease, reflecting a reduction in the number of subordinate clauses per complex sentence (possibly to aid an increase in the number of sentences so as to include more opinions). The significant increase in MLOS relative to the modest gain in MLRS suggests the production of relatively fewer fragments (and, thus, greater convergence) in argumentative writing. Both genres exhibit similar increases in the overall proportion of complex sentences, but narrative writing started in the first year as considerably less complex (e.g. C/TU 1.16 as opposed to argumentative C/TU 1.39) and it was therefore perhaps easier for learners to increase complexity (and lexicality) in this genre in the second year in order to develop descriptive reporting (C/TU 1.19 compared to argumentative C/TU 1.37), while the imperatives for argumentative writing demanded both increases in supportive details and the number of supporting sentences.

There are significant differences between texts in the two genres in the first year and in the second year, but not as anticipated. Complexity results accord with expectations, with both first and second year showing a similar pattern: greater complexity in argumentative writing as measured by C/TU, C/OS, DC/TU, and DC/C, and significantly greater complexity in narrative writing as measured by TU/S (reflecting the much higher ratio of compound sentences used for the genre). Contrary to expectations, however, all fluency results for both first and second year are significantly greater for argumentative writing, with the exception of first year MLT. This surprising result may be due to learners receiving more instruction in the genre and/or to writers expediting production by

using a more formulaic construction of text within which to list propositions and support. Whatever the reason, there is little support here for a trade-off between syntactic complexity and fluency.

In terms of sentence variety, argumentative texts show significantly more variety than narratives in the first year, but the measure is only marginally greater in the second year. As mentioned above, while the SVI serves as a good overall indicator of the range and balance in distribution of sentence types in a text, it fails to reveal the relative contribution of the various types sentence used. In the second year, for example, the SVI for narrative texts is 37.51, and for argumentative is 37.88. Roughly five percent more simple sentences are used in narrative writing, but the proportion of compound-complex sentences for both genres is similar. The most substantial differences are in the proportions of complex sentences (argumentative writing using twice as many as narrative) and compound sentences (narrative writing using more than twice as many as argumentative). The statistical outcome is, however, is very similar SVIs. This ambiguity can be overcome by reference to the sentence type distribution table (Results, Table 9), suggesting that a combination of index and table can provide a useful investigative tool for descriptive analysis.

Following Lu (2011), we anticipated a possible decrease in C/OS and DC/TU after one year of instruction. Both measures increased in narrative writing. C/OS increased in argumentative writing, while DC/TU (and C/TU) decreased, but as these were accompanied by a decrease in MLTU, there is no evidence here of complexification at the phrasal level.

Finally, we expected a convergence of results for orthographic and reconstructed sentences as learners produced fewer clausal and non-clause fragments. The results offer no statistical confirmation of our

expectations, but there are clear trends towards convergence in both narrative and argumentative writing.

On the whole, the results of this study have produced expected and surprising results. Significant gains in syntactic complexity and fluency were made after a year's tuition, but these gains are largely limited to narrative texts. Argumentative writing appears to undergo a different pattern of development, with writers perhaps focusing more on overall text construction (number of sentences and text length) rather than developing sentence structure. There are significant differences between texts in the two genres, with argumentative being predictably more complex, but also, unexpectedly, more fluent. No indications of complexification were found, but this is perhaps due to the proficiency level of the learners. Despite its limitations, the SVI proved a useful indicator of developments in the use of a greater range of sentence types, highlighting significant increases in variety for both genres after a year in program. Similarly, the novel metrics MLOS and MLRS were valuable in reflecting sentence development, with the trend toward convergence suggesting a greater proficiency on the part of learners to construct syntactically-conventional sentences.

## **Conclusion**

The suite of measures that we employed in this research has illuminated differences in text production after a year of tuition, differences in genre production, the use of a greater variety of sentences after a year, and developments in sentence construction. Some of the findings conform to those of previous studies, other do not. This suggests that some of the factors affecting the writing production of the L2 learners in this study may be socio-linguistic conventions, some related to the

method of data collection, while others may be unique to the particular educational environment, echoing Lu's conclusion that, "the results suggest that institution, genre, and timing condition have significant effects on the observed mean values of all or most measures" (2011, p. 50).

The expected and unusual trends observed in this research offer two specific directions for further research: first, a longitudinal study to consider if and how these trends develop, and, second, a comparative study with native-speaker texts produced under the same timed conditions. The findings will not and cannot provide a complete picture of language development. That would require not only a comprehensive appraisal of accuracy, vocabulary, morphology, syntax, semantics, pragmatics, and discourse (Lu, 2011), but also a qualitative analysis of individual texts (Rimmer, 2009). These limitations accepted, future corpus-based research can, as with this study, offer valuable insights into the patterns that emerge in L2 written texts as learners strive to develop their skills within specific learning and usage environments.

## References

- Bardovi-Harlig, K., & Bofman, T. (1988, March). *A second look at T-unit analysis: Reconsidering the sentence*. Paper presented at the 22<sup>nd</sup> Annual Meeting of the Teachers of English to Speakers of Other Languages, Chicago, IL. Retrieved from <http://www.eric.ed.gov/>
- Bauman, R. (1992). Contextualization, tradition, and the dialogue of genres: Icelandic legends of the kraftasfald. In A. Duranti & C. Goodwin (Eds.), *Rethinking Context: Language as an Interactive phenomenon* (pp. 125-146). Cambridge: Cambridge University

Press.

- Beers, S. F. & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Reading and Writing*, 22, 185–200. doi:10.1007/s11145-007-9107-5
- Casanave, C. (1994). Language development in students' journals. *Journal of Second Language Writing*, 3, 179-201.
- Cooper, T. C. (1976). Measuring syntactic patterns of second language learners of German. *The Journal of Educational Research*, 69, 176-183.
- Crowhurst, M., & Piche, G. (1979). Audience and mode of discourse effects on syntactic complexity in writing at two grade levels. *Research in the Teaching of English*, 13, 101-109.
- Crowhurst, M. (1980). Syntactic complexity and teachers' quality ratings of narrations and arguments. *Research in the Teaching of English*, 14, 223-232.
- de Haan, P. (1987). Exploring the linguistic database: Noun phrase complexity and language variation. In W. Meijis (Ed.), *Corpus linguistics and beyond* (pp. 185-202). Amsterdam: Rodopi.
- Dörnyei, Z., & Skehan, P. (2005). Individual differences in second language learning. In C. Doughty & M. Long (Eds.), *Handbook of Second Language Acquisition* (pp. 589-630). Oxford: Blackwell.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 299-323.
- Guenther, S. & Knoblauch, H. (1995). Culturally patterned speaking practices: The analysis of communicative genres. *Pragmatics*, 5, 1-32.

- Halliday, M.A.K., & Hasan, R. (1985). *Language, context and text: Aspects of language in social-semiotic perspective*. Geelong, VIC, Australia: Deakin University Press.
- Hanks, W. (1987). Discourse genres in a theory of praxis. *American Ethnologist*, 14, 668-692.
- Hirano, K. (1991). The effect of audience on the efficacy of objective measures of EFL proficiency in Japanese university students. *Annual Review of English Language Education in Japan*, 2, 21-30.
- Homburg, T. J. (1984). Holistic evaluation of ESL compositions: Can it be evaluated objectively? *TESOL Quarterly*, 18, 87-107.
- Hunt, K. W. (1965). *Grammatical structures written at three grade levels*. Champaign, IL: The National Council of Teachers of English, Research Report 3.
- Hyland, K. (2003). Genre-based pedagogies: A social response to process. *Journal of Second Language Writing*, 12, 17-29.
- Ishikawa, S. (1995). Objective measurement of low proficiency EFL narrative writing. *Journal of Second Language Writing*, 4, 51-70.
- Kameen, P. T. (1979). Syntactic skill and ESL writing quality. In C. Yorio, K. Perkins, & J. Schachter (Eds.), *On TESOL '79: The learner in focus* (pp. 343-364). Washington D. C.: TESOL.
- Kern, R. G., & Schultz, J. M. (1992). The effects of composition instruction on intermediate level French students' writing performance: Some preliminary findings. *The Modern Language Journal*, 76, 1-13.
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27, 590-619.
- Lee, N., & Schumann, J. (2005). *The interactional instinct: The evolution*

- and acquisition of language*. Paper presented at the Congress of the International Association for Applied Linguistics, Madison, WI.
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15, 474-496. 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, 36-62.  
doi: 10.5054/tq.2011.240859
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. New York: Routledge.
- MacWhinney, B. (1998). Models of the emergence of language. *Annual Review of Psychology*, 49, 199-227.
- Miller, J. H. & Page, S. E. (2007). *Complex adaptive systems: An introduction to computational models of social life*. Princeton, N. J.: Princeton University Press.
- Navés, T. (2006) The Long-term Effects of an Early Start on EFL Writing. Unpublished PhD dissertation. Universitat de Barcelona. English Department. September 2006. Directed by Carmen Muñoz.
- Navés, T., Torras, M. R., & Celaya, M. L. (2003). Long-term effects of an earlier start. An analysis of EFL written production. In S. Foster-Cohen & S. Pekarek (Eds.), *Eurosla-yearbook. Annual conference of the European second language association* (Vol. 3, pp. 103-130). Amsterdam: John Benjamin.
- Nunberg, G., Briscoe, T., & Huddleston, R. (2002). Punctuation. In R. Huddleston & G. K. Pullum (Eds.), *The Cambridge Grammar of the English Language* (pp. 1724-1764). Cambridge: Cambridge University Press.

- Oostdijk, N., & de Haan, P. (1994). Clause patterns in modern British English: A corpus-based (quantitative) study. *ICAME Journal* 18, 41-79.
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24, 492-518.
- Polio, C. G. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning*, 45, 101-143.
- Polio, C. (2001). Research methodology in second language writing research: The case of text-based studies. In T. Silva and P. K. Matsuda (Eds.), *On second language writing* (pp. 91-115). Mahwah, N J: Lawrence Erlbaum.
- Purcell-Gates, V., Duke, N. K., & Martineau, J.A. (2007). Learning to read and write genre-specific text: Roles of authentic experience and explicit teaching. *Reading Research Quarterly*, 42, 8-45. doi:10.1598/RRQ.42.1.1 Retrieved from: <http://www.reading.org/>
- Rimmer, W. (2006). Measuring grammatical complexity: the Gordian knot. *Language Testing*, 23, 497-519. doi:10.1191/0265532206lt339oa
- Rimmer, W. (2009). Can What Counts in Complexity Be Counted? In L.J. O'Brien & D.S. Giannoni (Eds.), *Language Studies Working Papers*, 1, 25-34. Reading: University of Reading.
- Sharma, A. (1980). Syntactic maturity: Assessing writing proficiency in a second language. *Language Learning*, 37, 469-481.
- Stomberg, L., & Kurth, R. J. (1982). Mode of discourse effects on written syntactic complexity. *American Reading Forum*, 2, 113-115.
- Struc, N., & Wood, N. (2009). Using a learner corpus to develop language profiles. *Reitaku Review*, 15, 53-82.
- Struc, N., & Wood, N. (2010). A corpus-based analysis of Japanese

- university-level learners' L2 writing development over a one-year period. *Reitaku University Journal*, 90, 155-183.
- Swales, J. (1990). *Genre analysis*. Cambridge: Cambridge University Press.
- Szmrecsányi, B. M. (2004). On Operationalizing Syntactic Complexity. *JADT 2004: Journ'ees internationales d'Analyse statistique des Donn'ees Textuelles*, 7, 1031-1038.
- Torras, M. R., Navés, T., Celaya, M. L., & Pérez-Vidal, C. (2006). Age and IL development in writing. In C. Munoz (Ed.), *Age and rate of foreign language learning* (pp. 156-182). Clevedon: Multilingual Matters.
- Voss, M. J. (2005) Determining Syntactic Complexity Using Very Shallow Parsing. *CASPR Research Report 2005-01*, Artificial Intelligence Center, The University of Georgia, Athens, Georgia.
- Waugh, L. R. (1995). Reported speech in journalistic discourse: The relation of function and text. *Interdisciplinary Journal for the Study of Discourse*, 15, 129–173. doi:10.1515/text.1.1995.15.1.129
- Wolfe-Quintero, K., Inagaki, S., & Kim, H-Y. (1998). *Second language development in writing: Measures of fluency, accuracy & complexity*. Honolulu: University of Hawai'i Press.
- Yau, M. S. S. & Belanger, J. (1984). The influence of mode on the syntactic complexity of EFL Students at three grade levels. *TESL Canada Journal/Revue TESL du Canada*, 2, 65-76. Retrieved from <http://www.teslcanadajournal.ca/>

