# Developing Syntactic Complexity in L2 Writing

J. Paul Marlowe
*Kobe University*

**Reference Data:**
Marlowe, J. P. (2016). Developing syntactic complexity in L2 writing. In P. Clements, A. Krause, & H. Brown (Eds.), *Focus on the learner*. Tokyo: JALT.

Although complexity is considered an important aspect of language proficiency, little is known about the role of different writing activities in developing complexity. Therefore, in this study the effects were investigated of different types of sentence-based writing practice on the writing of 105 Japanese 1st-year university students. The researcher compared three groups who received regular targeted writing practice in either sentence-combining, translation, or timed-writing over the course of 1 academic year. Three writing samples were collected and analyzed for 4 quantitative measures of syntactic complexity. A repeated measures ANOVA was used to ascertain differences over time and between groups for all syntactic measures examined. The results indicate that the sentence-combining and translation groups outperformed the timed-writing group on several measures, but that there were no significant differences between the 2 groups.

この論文でセンテンス・ライティングの実践と統語的に複雑な文章を構成する力との関係を述べる。言語の習熟度を測る際、複雑な文章を構成する力の有無は重要な要素であると考えられているが、その力を養う上で様々なライティング訓練が果たす役割についてはあまり知られていない。本研究では、日本の大学1年生105名を対象に、センテンスレベルにおける異なるライティング訓練を課し、その効果を調査した。学生は、(1) センテンス・コンバイニング、(2) 文法訳読、(3) 時間制限を設けたライティング課題のいずれかを30週間、定期的に課せられた。研究者は統語的に複雑な文章を構成する力を測る4種類の定量的尺度を用い、経時的及びグループ間のばらつきを分析する為に、分散分析を使用した。結果、センテンス・コンバイニング及び文法訳読を行ったグループが、時間制限を設けたライティング課題を行ったグループをいくつかの測定で上回る結果が得られたが、二つのグループの間には有意な差は見られなかった。

Foreign language writing teachers often suffer from a lack of interesting reading. This is not necessarily because students do not have interesting thoughts, ideas, or opinions. Rather, their inability to demonstrate syntactic complexity and variety in writing sometimes leads to a sense of redundancy that poorly reflects some of the complex ideas being expressed. This is especially true of teachers working with adult language learners, who are attempting to express sophisticated ideas despite a limited grasp of the second language. Furthermore, syntactic complexity is an important aspect of teaching in programs that require the expression of complex ideas. These include programs that focus on English for academic purposes, English for specific purposes, content-based instruction, and any programs with a focus on research writing. Therefore, methods of teaching learners the forms of syntactic complexity should be a central focus for writing instructors in these areas.

## L2 Complexity and T-Units

Complexity is considered one of three major traits of language production in SLA research. Together, complexity, accuracy, and fluency (CAF) make up the three constructs of the model of language production originally proposed by Skehan (1989; 1998). Researchers and language practitioners have used CAF measures to evaluate the effectiveness of pedagogical interventions, to investigate the effects of different types of tasks, and to assess differences between proficiency levels by gauging the development, proficiency, or both of language learners through the use of several different metrics applied to oral and written L2 data. This is often achieved by researchers through metrics such as counting errors to measure accuracy, counting the number and length of pauses for fluency, and counting the number of words or clauses within sentences for complexity. Complexity itself can be broken down into multiple areas of focus, including lexical, morphological, and phonological complexity (Bulté & Housen, 2014). For the purposes of this study, complexity refers to syntactic forms, including phrasal, clausal, and sentential. According to Ortega (2003), syntactic complexity is defined by "the range of forms that surface in language production and the degree of sophistication of such forms" (p. 492). Syntactic complexity has been linked to L2 proficiency (Cooper, 1976; Ortega, 2003;

Wolfe-Quintero, Inagaki, & Kim, 1998), as well as to higher ratings of essays (McNamara, Crossley, & McCarthy, 2009).

Studies of syntactic complexity originated in L1 writing research and Hunt's (1965) development of the T-unit as a standard of measure of what he termed syntactic maturity. Essentially, a T-unit is an independent clause with any dependent clauses that are embedded in or attached to it. T-units increase within a sentence when one independent clause is conjoined with another. Hunt used the T-unit measure to show that learners of different age levels gradually progressed through stages of increasing complexity. The T-unit has been adopted as one of the most widely used units of measurement in studies of L2 syntactic complexity (Ortega, 2003).

## Purpose and Background

The purpose of this study was to understand the effect that different types of sentence-writing practice had on the development of syntactic complexity in Japanese university students' writing. Specifically, three different types of sentence-writing practice were examined: translation (Japanese to English), sentence combining, and timed writing.

Translation was chosen because it is part of the grammar-translation method, which has a long history as a form of language instruction in East Asian countries such as Korea and Japan. Although a more eclectic approach to teaching writing has become popular in many writing programs, many Japanese instructors of English in high school and university still tend to favor instructional techniques that are part of the grammar-translation method, as evidenced in teacher practices (Gorsuch, 1998; Hino, 1988; Nishino & Watanabe, 2008) and textbooks approved by the Ministry of Education, Culture, Sports, Science and Technology (MEXT; Kobayakawa, 2011). In fact, students are often required to translate passages from Japanese to English as a part of university entrance examinations. Grammar translation is typically a 3-step process. The first step usually includes an introduction to the target grammar and relevant examples. Next, learners read an exemplary sentence or short text and translate it into a target language. Finally, learners compare their translation either to the original text or to a model answer and analyze the errors and differences.

Timed writing is an activity that can be traced to Elbow's "freewriting" approach, in which writers focus on content without much regard to accuracy (as cited in Casanave, 2004). This is an approach that essentially allows the writer to focus on fluency, sometimes at the expense of accuracy and complexity. In this activity, learners write about familiar topics within a designated time set by the instructor. The amount of time can vary, but Nation (2008) suggested 10 minutes for English language learners. During this time, learners are instructed to write about either a teacher- or self-selected topic without stopping to think, erase, or make corrections. The goal of timed writing is for learners to write as many words as possible and begin to build fluency over the course of several writings.

Sentence combining involves providing a set of "kernel" sentences to learners. This can include several sets of simple sentences, but requires minimally at least a pair. The learners are instructed to join the sets of simple sentences through conjoining, nominalization, or subordination. Finally, the answers are compared and checked with possible solutions. There is evidence in L1 studies that practicing sentence combining leads to significant gains in syntactic fluency as well as improvements in the quality of writing (Mellon, 1969; Morenberg, Daiker, & Kerek, 1978; O'Hare, 1973). Furthermore, studies of foreign language writing have revealed significant differences in syntactic development in learners of French, Spanish, and German (Cooper, 1981), French as a foreign language (Cooper & Morain, 1980; Monroe, 1975), English as a second language (Gaies, 1976), and English as a foreign language (Abdan, 1983). However, there have been no investigations of the effects of sentence combining on Japanese learners of English as a foreign language. This is especially important, because Japanese learners have been characterized by their avoidance of more challenging syntactic structures in favor of accuracy (Schachter, 1974). Moreover, Japanese learners generally write with less syntactic complexity when compared to other learners of English (Lu & Ai, 2015). Therefore, in this study the following research questions were investigated:

RQ1. Does targeted sentence-writing practice lead to significant changes in learners' syntactic complexity?

RQ1a. If targeted sentence-writing practice leads to significant changes in learners' syntactic complexity, what type of changes occur?

RQ2. Does practicing sentence combining lead to greater gains in measures of syntactic complexity when compared to a translation group and a timed-writing group?

## Design, Materials, and Methods

### *Participants*

The participants were 105 first-year students at a national university in Western Japan. They consisted of a majority of male students (85%). The average TOEIC score was 430.

Additionally, the participants' vocabulary size was tested using the Vocabulary Size Test (Nation & Beglar, 2007). The results of this test indicated that the participants had an average vocabulary size within the 4,000-word range ($M$ = 4,487; $SD$ = 554). All the participants shared Japanese as a first language and were between the ages of 18 and 20 years ($M$ = 19.9; $SD$ = .90). At the time of the study, the students were enrolled in three English courses: Intensive Reading, English Communication, and Reading and Writing, with the latter two courses being coordinated. All the courses were taught over two 15-week semesters. Participants from three intact classes were randomly assigned sentence-combining and translation treatments. Another separate intact class that practiced timed writing was used, as this is what is commonly practiced in the regular curriculum in this program.

## Design

The study took place over 1 academic year, which consisted of 30 weeks. The reading and writing course focused primarily on academic paragraph and essay writing in addition to extensive reading practice. Writing samples were collected in the first class of the first semester (Time 1), the last class of the first semester (Time 2), and the final class of the academic year (Time 3). In between, the participants practiced their respective treatments every week for 14 weeks each semester (see Figure 1).

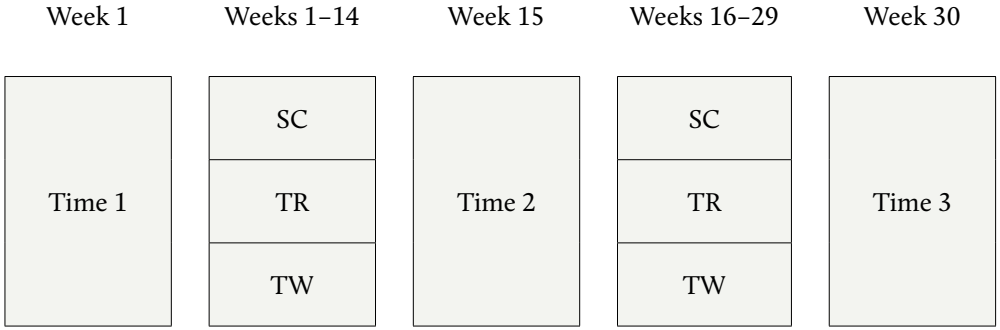


*Figure 1.* Research design overview. SC = Sentence-combining group; TR = Translation group; TW = Timed-writing group.

## Instruments and Materials

The writing samples tests consisted of four comics from Grade Pre-1 of a STEP (Eiken) practice speaking test task and were adapted and utilized for the purpose of collecting narrative writing samples. The comics were chosen in order to provide controlled content for all participants as well as provide stories that take place within a familiar Japanese context. Each comic consisted of four cells depicting a short story. Each story had a distinct setting: one about farming, one about moving to a new apartment, one about behavior on a train, and a fourth about a community cleaning a neighborhood pond.

Two comics were provided during each test time with instructions written in Japanese. The four comics were anchored across the three test times. Therefore, Comic A (farming), was used for Time 1 and Time 2 and Comic C (train) was used for Time 2 and Time 3 (Table 1). This was done in order to have one sample that would serve as a direct comparison, linking each time together. Another reason for this was to have a new sample included at Time 2 and Time 3 that would demonstrate any developmental changes that would not simply be the result of a practice effect (Table 1).

Table 1. Anchoring Plan for the Writing Samples Prompts

| Time 1 | | Time 2 | | Time 3 |
|---|---|---|---|---|
| Comic A | → | Comic A | | Comic D |
| Comic B | | Comic C | → | Comic C |

The researcher developed all of the treatment worksheets following the recommended sequence laid out by Cooper (1973) and Lawlor (1983). This sequence gradually progressed through coordinates, adverbials, noun modifiers, noun substitutes, and free modifiers over the course of one academic year. Each week, two target structures were presented with seven sentence-writing exercises for each structure. A narrative paragraph exercise was included that integrated the two structures and recycled past structures. The vocabulary was restricted to the most frequent 2,000 words using a combination of the British National Corpus and the Corpus of Contemporary American English. These were chosen because they are two of the largest, balanced, and freely available corpora. All exercises were analyzed using a web-based vocabulary profiler (Cobb, 2015). Idiomatic expressions were kept to a minimum and cultural references were avoided. Each semester included an introduction and review at the start, two quizzes, and a semester review. All worksheets had been piloted for a year prior to the study with students in the same faculty and revised accordingly.

## Procedures

The participants in both the sentence-combining and translation groups received weekly worksheets that followed the same procedure. After reading a Japanese explanation of the target structure, participants were provided an example task and then worked through seven practice tasks. This was done for two target structures, for a total of 14 practice tasks. Upon completion, the participants were provided answers and instructed to individually correct any syntactic errors. The last step included a series of sentences that together comprised a 1-paragraph narrative story that focused on the two target structures in addition to previously practiced structures. Finally, participants were provided a model answer to compare their answers to and correct any syntactic errors on their own. The main difference between the two groups was that whereas the sentence-combining group combined short sentences in English into one English sentence, the translation group wrote the same sentence, but translated it from Japanese text. In addition to the practice exercises, participants were encouraged to combine sentences together in their academic writing through the use of a self-check rubric. All of these steps were done in order to follow Kameen's (1978) suggested framework for sentence-combining writing activities in which exercises progress from mechanical to meaningful and finally communicative.

Meanwhile, the group that practiced timed writing was provided a list of topics from which they freely chose each week. Upon selection, students marked their selection on the list and could only choose from the remaining unmarked options. Upon the teacher's instruction to begin, the students wrote as much as they could without stopping for 10 minutes. At the end of the time, students were instructed to count the number of words they had written and write it on a chart in order to graph and compare the number of words written with previous entries. Students were then given time to do sustained silent reading of their graded readers. Although the overall time was much shorter for the timed-writing group, the amount of sentence production was roughly the same as in the other two groups, with all students writing around 150 words each week (Table 2).

Table 2. Detailed Description of Groups

| Group | Hours | Year | *n* | Class *n* | Procedure for 90-min. class |
|---|---|---|---|---|---|
| SC | 100 | 1st | 42 | 30 | 1. Sentence combining (30 min.) |
| | | | | | 2. Textbook coverage (30 min.) |
| | | | | | 3. Paragraph writing (30 min.) |
| TR | 100 | 1st | 35 | 30 | 1. Translation writing (30 min.) |
| | | | | | 2. Textbook coverage (30 min.) |
| | | | | | 3. Paragraph writing (30 min.) |
| TW | 100 | 1st | 28 | 30 | 1. Timed writing (10 min.) |
| | | | | | 2. Silent reading (20 min.) |
| | | | | | 3. Textbook coverage (30 min.) |
| | | | | | 4. Paragraph writing (30 min.) |

*Note.* *n* = number of students in each group; class *n* = the number of students in each class; hours = hours of English class time; SC = sentence-combining group; TR = translation group; TW = timed-writing group.

### *Measures and Analyses*

A repeated measures analysis of variance (ANOVA) was used to analyze five dependent variables of syntactic complexity (see Table 3). Writing samples were analyzed for mean length of sentence (MLS), mean length of T-unit (MLT), mean length of clause (MLC), and clauses per T-unit (C/T). In Norris and Ortega's (2009) examination of measures of syntactic complexity, they recommended that researchers use MLT and MLC as measures of global or general complexity and C/T as a measure of subordination. Writing samples were analyzed using the L2 Syntactical Complexity Analyzer (L2SCA) developed by Lu (2010). The L2SCA is a computational system for automatic analysis of syntactic complexity in second language writing. The software analyzes text for up to 14 different syntactic structures and syntactic complexity indices.

Table 3. Syntactic Complexity Measures

| Measure | Description | Calculation |
|---|---|---|
| Mean length of sentence (MLS) | Global complexity measure of sentential unit | Total number of words divided by number of sentences |
| Mean length of T-unit (MLTU) | Global complexity measure of sentential unit | Total number of words divided by number of T-units |
| Mean length of clause (MLC) | Complexity measure of clausal and/or phrasal elaboration | Total number of words divided by number of clauses |
| Clauses per T-unit (C/T) | Complexity measure of subordination | Total number of clauses divided by number of T-units |

The results of the syntactic analysis using the L2SCA were tabulated and analyzed using the Statistical Package for Social Science (SPSS) version 21.0 software program. Comparisons were made between times and groups. It was predicted that there would be significant differences for time for all groups. Furthermore, it was predicted that the sentence-combining group would significantly outperform the other groups on all measures of syntactic complexity.

MLS was measured in the three groups at three times. The descriptive statistics for MLS are presented in Table 4. The results of the univariate test show no significant differences for within-subjects effects for time. Mauchly's test indicated that the assumption of sphericity had been violated $c^2(2) = 6.18$, $p = .049$. Sphericity is the condition of all of the levels of the independent variable being equal. As a result of the violation, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity .938. The results show that there was no significant within-subjects effect for time $F(1.87, 166.96) = 2.03$, $p = .137$. This suggests that sentence length did not change significantly overall for the combined groups over time. Tests of between-subjects effects revealed significant differences between groups, $p < .05$. Post hoc comparisons, which show differences using the Fisher LSD test, revealed significant differences between the translation group and the timed-writing group ($p < .05$) and between the sentence-combining group and the timed-writing group ($p < .05$). No significant differences were found between the translation and sentence-combining groups.

Table 4. Descriptive Statistics for Mean Length of Sentence

| Group | Time 1 | | Time 2 | | Time 3 | |
|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Translation | 12.13 | 2.74 | 12.24 | 2.37 | 12.41 | 2.54 |
| Sentence combining | 11.70 | 2.92 | 12.13 | 2.85 | 12.96 | 2.77 |
| Timed writing | 10.77 | 2.97 | 10.98 | 2.53 | 11.04 | 1.95 |

The results of the univariate test for MLT revealed no significant within-subjects effect for time. The descriptive statistics for MLT are presented in Table 5. Mauchly's test indicated that the assumption of sphericity had been met $c^2(2) = 2.15$, $p > .05$. There were no significant within-subjects effect for time $F(2, 178) = 2.74$, $p = .068$. This suggests that the mean number of T-units did not increase for groups over time. Tests of between-subjects effects revealed significant differences between groups ($p < .05$). Post hoc comparisons using the Fisher LSD test revealed significant differences between the translation group and the timed-writing group ($p < .05$) as well as significant differences between the sentence-combining group and the timed-writing group ($p < .05$). There were no significant differences between the translation and sentence-combining groups for MLS.

Table 5. Descriptive Statistics for Mean Length of T-Units

| Group | Time 1 | | Time 2 | | Time 3 | |
|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Translation | 10.17 | 1.59 | 10.32 | 1.10 | 10.49 | 1.63 |
| Sentence combining | 10.23 | 1.83 | 10.25 | 1.80 | 10.93 | 1.50 |
| Timed writing | 9.59 | 1.50 | 9.57 | 1.35 | 9.97 | 1.63 |

The results of the univariate test for MLC revealed significant within-subjects effects. The descriptive statistics for MLC are presented in Table 6. Mauchly's test indicated that the assumption of sphericity had been met $c^2(2) = 3.86$, $p > .05$. There were significant within-subjects effects for time $F(2, 178) = 30.55$, $p < .001$. This suggests that overall the

participants writing experienced significant growth at the clausal and/or phrasal levels. Tests of between-subjects effects revealed no significant differences between groups ($p$ = .077). The results suggest that although there were overall significant gains over time, there were no significant difference in gains between the three groups.

Table 6. Descriptive Statistics for Mean Length of Clause

| Group | Time 1 | | Time 2 | | Time 3 | |
|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Translation | 7.93 | 1.15 | 8.14 | 1.46 | 9.48 | 1.12 |
| Sentence combining | 7.71 | 1.13 | 8.21 | 1.83 | 9.05 | 1.21 |
| Timed-writing | 7.47 | 0.98 | 7.74 | 1.15 | 8.82 | 1.17 |

The results of the univariate tests for C/T revealed significant within-subjects effects. The descriptive statistics for C/T are presented in Table 7. Mauchly's test indicated that the assumption of sphericity had been violated $c^2$ (2) = 51.61, $p < .001$. Therefore, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity .693. There were significant within-subjects effects for time $F(1.38, 123.29) = 53.13$, $p < .001$. This suggests there were overall changes to C/T among all participants over time. Tests of between-subjects effects revealed no significant differences between groups ($p > .05$). The results suggest that although there were overall significant changes over time, there were no significant difference in changes between the three groups.

Table 7. Descriptive Statistics for Clauses per T-Unit.

| Group | Time 1 | | Time 2 | | Time 3 | |
|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Translation | 1.29 | 0.20 | 1.45 | 0.26 | 1.10 | 0.15 |
| Sentence combining | 1.33 | 0.22 | 1.46 | 0.24 | 1.21 | 0.22 |
| Timed-writing | 1.29 | 0.22 | 1.40 | 0.30 | 1.13 | 0.17 |

## Discussion

The objective of this study was to examine the effects of three different types of writing practice on the syntactic complexity of L2 English writing among groups of 1st-year Japanese university students. In examining MLS, the average sentence length of the translation and sentence-combining group demonstrated significant increases in comparison to the timed-writing group. A similar result was found with the MLT measurement, another global measure. There were no significant differences between translation and sentence-combining groups, but both groups made significant gains over the timed-writing group. Although MLS indicates development in syntactic complexity, it is a global measure that provides little indication of exactly where elaboration is occurring. Therefore, a measure of clausal or phrasal elaboration was also analyzed. All groups combined made significant gains in MLC over time, but there were no significant differences between any of the groups. Regarding the index of subordination, gains in C/T were made between Time 1 and Time 2, but then decreased between Time 2 and Time 3.

Therefore, the hypothesis that the sentence-combining group would outperform both the translation and timed-writing groups on all measures was not supported by the results. One interpretation of these results is that this type of translation is quite similar to sentence combining, which may explain why there were no significant differences between the two groups on any measures. However, both groups outperformed the timed-writing group on global measures of complexity, possibly indicating that form-focused practice is superior to fluency writing in terms of developing L2 complexity. Furthermore, the fact that MLC increased while C/T decreased significantly over time for all groups could indicate that there is a combination of clausal elaboration and coordination occurring simultaneously. Rather than subordination, writers at this level are beginning to elaborate clauses through adverbials and nominalizations and also starting to join more sentences through the use of conjunctions. These findings correspond with Crossley and McNamara's (2014) and Bulté and Housen's (2014) patterns of syntactic development in which at the end of each study learners produced more complex phrases, longer clauses, and less subordination.

## Implications, Limitations, and Conclusion

This study has some notable implications for L2 writing instruction and research. First, complexity in L2 writing is an aspect of writing that can be hastened through form-focused practice that specifically draws attention to syntactic patterns in sentence structures. This corroborates other research into L2 complexity in the field of SLA (Ortega, 2003). Second, translation and sentence combining as instructional techniques appear to

encourage the development of complexity in L2 writing development when compared to more free-form writing activities, such as timed writing, that encourage fluency and production. Third, translation in the form of Japanese to English and sentence combining might tap into similar cognitive processes that are conducive to developing complexity in L2 writing.

There were several limitations to this study, which might provide insights into the results as well as indicate directions for future research. The first limitation of this study is that the data were measured using specific quantitative measurements. Although these measures are useful in capturing developmental changes in syntactic complexity on several levels within the sentence across time, they cannot capture qualitative changes in the participants' writing. Even though the ability to use a variety of sentence structures can make writing less redundant, it is still arguable as to whether it makes it more interesting for readers, who enjoy writing for reasons beyond sentence complexity. Future research could combine these quantitative measurements with qualitative measures by human raters in order to investigate potential links between development in complexity and improvements in overall writing quality.

Second, the comic used for the writing prompts might have inadvertently limited the writing sample, provided some of the writing content, and focused more of the cognitive process on retrieving unknown vocabulary. The comics consisted of only four cells, which lead some participants to write one sentence for each cell for a total of four sentences. Moreover, the comics contained dialogue, which many participants used in their story writing. Finally, participants were forced to retrieve lexical items to describe the situations, objects, and events depicted in the comics. This may have focused the participants' cognitive processes on the lexical items rather than the syntactic forms of the sentence structures. Furthermore, the limited amount of time given for each comic (10 minutes) might have also lead to reduced writing production. Future researchers might want to investigate complexity using more open-ended personal narrative prompts in addition to a longer test time to encourage more thoughtful writing and more robust writing samples.

Finally, the results indicated that the sentence-combining and translation groups did not separate on the statistical measures of complexity. This could indicate that translation and sentence combining are tapping into a similar cognitive process in L2 writing. Future researchers might want to investigate the cognitive processes of each writing task using think-aloud protocols to understand any differences or similarities between the two tasks.

## Bio Data

**J. Paul Marlowe** is an associate professor at Kobe University's Graduate School of Maritime Sciences. He is currently a doctorate student at Temple University Japan. His main research interest focuses on developing complexity in L2 learner writing.

## References

Abdan, A. A. (1983). *The effect of sentence-combining practice on the written syntactic maturity, overall writing quality, and reading comprehension of EFL Saudi students* (Unpublished doctoral dissertation). University of Kansas, Lawrence, KS.

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

Casanave, C. P. (2004). *Controversies in second language writing: Dilemmas and decisions in research and instruction*. Ann Arbor: University of Michigan Press.

Cobb, T. (2015). Compleat web VP! [computer program]. Retrieved from <http://www.lextutor.ca/vp/comp/>

Cooper, C. R. (1973). An outline for writing sentence-combining problems. *English Journal, 62,* 96-108. http://dx.doi.org/10.2307/814087

Cooper, T. C. (1976). Measuring written syntactic patterns of second language learners of German. *The Journal of Educational Research*, *69*(5), 176-183. http://dx.doi.org/10.1080/00220671.1976.10884868

Cooper, T. C. (1981). Sentence combining: An experiment in teaching writing. *The Modern Language Journal, 65,* 158-165. http://dx.doi.org/10.1111/j.1540-4781.1981.tb00967.x

Cooper, T., & Morain, G. (1980). A study of sentence-combining techniques for developing written and oral fluency in French. *French Review, 53,* 411-423. Retrieved from <http://www.jstor.org/stable/390492>

Crossley, S. A., & McNamara, D. S. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing, 26,* 66-79. http://dx.doi.org/10.1016/j.jslw.2014.09.006

Gaies, S. J. (1976, May). *Sentence-combining: A technique for assessing proficiency in a second language*. Paper presented at the Conference on Perspectives on Language, Louisville, KY.

Gorsuch, G. J. (1998). *Yakudoku* EFL instruction in two Japanese high school classrooms: An exploratory study. *JALT Journal*, *20*, 6-32.

Hino, N. (1988). Yakudoku: Japan's dominant tradition in foreign language learning. *JALT Journal*, *10*, 45-55.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels*. NCTE Research Report No. 3.

Kameen, P. T. (1978). A mechanical, meaningful, and communicative framework for ESL sentence combining exercises. *TESOL Quarterly*, *12*, 395-401. http://dx.doi.org/10.2307/3586138

Kobayakawa, M. (2011). Analyzing writing tasks in Japanese high school English textbooks: English I, II, and writing. *JALT Journal*, *33*, 27-48.

Lawlor, J. (1983). Sentence combining: A sequence for instruction. *The Elementary School Journal*, *84*(1), 53-62. Retrieved from <http://www.jstor.org/stable/1001129>

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

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. http://dx.doi.org/10.1016/j.jslw.2015.06.003

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2009). Linguistic features of writing quality. *Written Communication, 27*(1), 57-86. http://dx.doi.org/10.1177/0741088309351547

Mellon, J. C. (1969). *Transformational sentence-combining: A method for enhancing the development of syntactic fluency in English composition*. (NCTE Research Report No. 10). Urbana, IL: National Council of Teachers of English.

Monroe, J. H. (1975). Measuring and enhancing syntactic fluency in French. *French Review, 48,* 1023-1031. Retrieved from <http://www.jstor.org/stable/389177>

Morenberg, M., Daiker, D., & Kerek, A. (1978). Sentence combining at the college level: An experimental study. *Research in the Teaching of English, 12,* 245-256. Retrieved from <http://www.jstor.org/stable/40170713>

Nation, I. S. P. (2008). *Teaching ESL/EFL reading and writing*. New York, NY: Routledge.

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31*(7), 9-13.

Nishino, T., & Watanabe, M. (2008). Communication-oriented policies versus classroom realities in Japan. *TESOL Quarterly, 42,* 133-138. http://dx.doi.org/0.1002/j.1545-7249.2008.tb00214.x

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30,* 555-578. http://dx.doi.org/10.1111/0023-8333.00136

O'Hare, F. (1973). *Sentence combining: Improving student writing without formal grammar instruction* (NCTE Research Report No. 15). Urbana, IL: National Council of Teachers of English.

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. http://dx.doi.org/10.1093/applin/24.4.492

Schachter, J. (1974). An error in error analysis. *Language Learning, 24,* 205-214. http://dx.doi.org/10.1111/j.1467-1770.1974.tb00502.x

Skehan, P. (1989). *Individual differences in second language learning*. London, UK: Edward Arnold.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.

Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity* (No. 17). Honolulu: University of Hawaii Press.