

A corpus-based comparison of syntactic complexity in NNS and NS university students' writing

Haiyang Ai and Xiaofei Lu

This paper reports on a corpus-based comparison of syntactic complexity in NNS and NS university students' writing. We analysed 600 essays from the *Written English Corpus of Chinese Learners* and the *Louvain Corpus of Native English Essays* using 10 syntactic complexity measures to investigate whether and the extent to which NNS and NS university students' writing differs with respect to length of production unit, amount of subordination, amount of coordination, and degree of phrasal sophistication. Results showed significant differences in all four dimensions of syntactic complexity between the writing of NNS students at both low and high proficiency levels and that of NS students. This gap calls for the design of pedagogical interventions to enhance NNS university students' syntactic development.

1. Introduction

Syntactic complexity, i.e. the range and degree of sophistication of syntactic structures that surface in language production, has been recognized as a very important construct in second language writing teaching and research (e.g. Ortega 2003). A large variety of syntactic complexity measures have been proposed in the second language writing development literature, and numerous second language writing development studies have been conducted to determine which of them constitute valid and reliable developmental indices that can be used by second language teachers and researchers to objectively gauge second language learners' developmental level or global proficiency in the target language (e.g. Larsen-Freeman 1978, 2009; Wolfe-Quintero et al. 1998; Ortega 2003; Lu 2011).

Whereas researchers have approached second language writing proficiency with different conceptualizations, a native-speaker (NS) baseline appears to be a rather neglected dimension in the examination and assessment of the performance and developmental level of non-native speakers (NNS) in the target language

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY

(e.g. Foster & Tavakoli 2009). Comparing NNS and NS performances not only allows us to determine whether and the extent to which NNS performance deviates from or approximates NS performance, but also provides valuable information that could be used by second language teachers and course material developers to devise appropriate pedagogical interventions that target specific problem areas (e.g. Hinkel 2003). Although some studies have used an NS baseline to examine NNS performance in second language writing (e.g. Reid 1992; Ferris 1994), studies that systematically compare syntactic complexity in NNS and NS students' writing are rather scarce.

One of the factors that may have contributed to this situation is the lack of computational tools for automating syntactic complexity analysis in second language writing and the labour-intensiveness of manual analysis. As a result, the scale of the few studies that examined syntactic complexity in NNS and NS students' writing (e.g. Foster & Tavakoli 2009) tended to be on the small side, with a limited number of syntactic complexity measures applied to relatively small amounts of data. This problem is not unique to comparative studies, however, but common among studies that examined syntactic complexity in second language writing in one way or another. For example, over 20 second language writing studies were reviewed in a research synthesis by Ortega (2003), most of which applied one to three different syntactic complexity measures to an average of fewer than 100 writing samples, with a few notable exceptions.¹ This situation remains largely unchanged in more recent research. For example, Stockwell & Harrington (2003) examined one syntactic complexity measure, clauses per T-unit, on approximately 300 email messages; Ellis & Yuan (2004) applied the same measure to 52 narratives; and Beers & Nagy (2009) focused on two measures, mean length of clause and clauses per T-unit, in an analysis of 41 essays. Briefly, a *T-unit* consists of a main clause and any dependent clause or nonclausal structure attached or embedded in it (Hunt 1970). While there is no denying that previous studies have offered very useful insight into the relationship of syntactic complexity to second language writing development, the quantitative results derived from studies that used small datasets need to be interpreted with caution. Moreover, as Wolfe-Quintero et al. (1998) and Ortega (2003) noted, it is difficult to pool results from the literature to examine the relative performance of different measures, as there is considerable variability and inconsistency in the choice and definition of syntactic complexity measures as well as the various aspects of research design among previous studies.

1. The units of analysis in studies that examine syntactic complexity in second language writing are usually complete learner essays, as syntactic complexity measures are generally computed as ratios of one structure to another in a text. The number of complete essays or texts analysed is therefore a more informative indicator of the sample size than the number of words analysed in this case.

Following Lu (2011), the current study constitutes another effort to remedy this situation. Taking advantage of the availability of the newly developed *L2 Syntactic Complexity Analyser* (Lu 2010), a computational system designed to automate the analysis of syntactic complexity of writing samples produced by college-level L2 English learners using a wide range of measures, we apply a comprehensive set of 10 syntactic complexity measures to large-scale college-level NNS and NS writing data from the *Written English Corpus of Chinese Learners Version 2.0* (WECCL 2.0) (Wen, Liang & Yan 2008) and the *Louvain Corpus of Native English Essays* (LOCNESS) (see Granger 1996). This is done with the aim to systematically investigate whether and the extent to which NNS and NS university students' writing differs in syntactic complexity, conceptualized here as a multifaceted construct encompassing length of production unit, amount of subordination, amount of coordination, and degree of phrasal sophistication. It is our hope that findings from this large-scale comprehensive comparison will contribute useful and reliable insight into college-level EFL learners' syntactic development.

The rest of the paper is organized as follows. We first review previous second language studies that examined the relationship between syntactic complexity and second language writing and the role of syntactic complexity in second language writing instruction and assessment. We then discuss how syntactic complexity has been quantified in previous research and how it is quantified in the present study. Next, the method section describes the two corpora of college-level writing data analysed in this study, the specific research questions addressed, and the analytical procedure. This is followed by a presentation of the results and a discussion of the research findings. Finally, we conclude the paper with a discussion of the pedagogical implications of our findings and directions for future research.

2. Syntactic complexity in second language writing

Along with measures of accuracy, fluency, and lexical complexity, measures of syntactic complexity have received extensive attention from second language writing development researchers in the search for valid and reliable developmental measures that can be used to objectively gauge second language learners' developmental level or global proficiency in the target language (e.g. Larsen-Freeman 1978, 2009; Wolfe-Quintero et al. 1998; Ortega 2003; Norris & Ortega 2009; Lu 2011). A major strand of research in this area has focused on the relationship between proficiency and syntactic complexity in second language writing. Some longitudinal studies in this strand have investigated changes in syntactic complexity in second language writing over a certain period of time (e.g. Hunt 1970; Casanave 1994; Ortega 2000; Stockwell & Harrington 2003; Stockwell 2005; Norrby 2007). For

example, Stockwell & Harrington (2003) collected email exchanges between college-level learners of Japanese and native Japanese college students over a five-week period to examine the effect of email interactions on incidental syntactic development, and reported a reliable increase in syntactic development quantified using several metrics of structural mastery. Norrby (2007) analysed sentence length, subordination, and nominal vs. verbal style in spoken and written data from adult learners of Swedish over a one-year period to understand how linguistic complexity interacts with learners' morpho-syntactic development. Other studies in this strand adopted a cross-sectional design to assess the extent to which different measures of syntactic complexity correlate with or have an effect on proficiency. For example, in an early study, Larsen-Freeman (1978) analysed 212 compositions written by college students at five ESL proficiency levels using several measures based on the T-unit in order to identify the best discriminators of these levels. Ferris (1994) examined 160 ESL compositions produced by students at low and advanced levels in an ESL program and presented several syntactic features that significantly differentiated these two levels. Using the L2 Syntactic Complexity Analyser, Lu (2011) evaluated a range of measures of syntactic complexity as indices of college-level ESL writers' language development using large-scale written data of Chinese EFL learners, and recommended several measures that discriminated learners in different school years. These cross-sectional and longitudinal studies varied significantly in terms of the choice and definition of the specific measures examined, the operationalization of proficiency (e.g. using program level, holistic ratings, or standardized test scores), the number and type of learners and writing samples analysed, as well as the results reported on the same measures, making it challenging to synthesize the cumulative knowledge presented (Wolfe-Quintero et al. 1998; Ortega 2003; Lu 2011).

Another major strand in the research on the relationship between syntactic complexity and second language writing development has systematically evaluated the role various learner-, task-, and context-related variables play in this relationship. Previous studies have shown that such variables as writing tasks and prompts (e.g. Way et al. 2000), genre (e.g. Lu 2011), mode of computer-mediated communication (e.g. Sotillo 2000), pretask planning (e.g. Ellis & Yuan 2004), timing condition (e.g. Lu 2011), and instructional setting (e.g. Ortega 2003) have various effects on this relationship.

Studies that systematically compared syntactic complexity in NS and NNS performances are scarce, with a few notable exceptions. For example, Hinkel (2003), in her quantitative analysis of 1,083 NS and NNS English academic texts, found that advanced NNS students in U.S. universities tended to overuse simple syntactic constructions. Recognizing the importance of an NS-baseline as an important dimension in studying second language performance, Foster & Tavakoli (2009)

examined the effect of task features on syntactic complexity, operationalized as syntactic subordination and mean length of utterance, in NS oral narratives and compared the results with those from a parallel study on NNS oral narratives (Tavakoli & Foster 2008). The results revealed that the effect of task design on syntactic complexity differs between the NNS and NS groups. However, for most of the task types, there is a clear gap in syntactic complexity between the two groups.

In addition to the relationship between syntactic complexity and second language writing development, the role of syntactic complexity in second language writing instruction and assessment has also received considerable attention (e.g. Buckingham 1979; Perkins, 1983; Silva 1993; Hinkel 2003). For example, Hinkel (2003) considered possible reasons driving the prevalent tendency for second language writers to rely heavily on simple syntactic features in academic essays, and put forward several instructional methods for addressing this problem. Perkins (1983) discussed the strengths and limitations for employing each of several syntactic complexity measures directly or indirectly in assessing ESL learners' writing proficiency.

3. Measuring L2 syntactic complexity

A large variety of measures have been proposed for characterizing syntactic complexity in the second language writing development literature. Most of these measures gauge syntactic complexity by quantifying one of the following: length of production unit, amount of subordination or embedding, amount of coordination, range of syntactic structures, and degree of phrasal sophistication. A comprehensive review of these measures can be found in a book-length research synthesis by Wolfe-Quintero et al. (1998), who examined more than 100 measures of accuracy, fluency and complexity in 39 second language writing development studies. As Lu (2010, 2011) noted, the set of measures reviewed in this research synthesis represent a fairly complete picture of the range of measures that have been adopted in second language writing research to date. To alleviate the labour-intensiveness of manual analysis, Lu (2010) designed the *L2 Syntactic Complexity Analyser*, a computational system for automatic measurement of syntactic complexity of English writing samples produced by college-level English learners with 14 measures selected from this set. All of the 14 measures were either shown by at least one previous study to have a significant effect for proficiency or were recommended by Wolfe-Quintero et al. (1998) for further research.

The present study will take advantage of the automatic processing capability of the *L2 Syntactic Complexity Analyser*, and will focus on 10 of the 14 measures that are included in the *L2 Syntactic Complexity Analyser*. The four measures – complex

Table 1. Syntactic complexity measures investigated

| Measure | Code | Formula |
|---|------|--------------------------------------|
| <i>Length of production unit</i> | | |
| Mean length of clause | MLC | # of words/# of clauses |
| Mean length of sentence | MLS | # of words/# of sentences |
| Mean length of T-unit | MLT | # of words/# of T-units |
| <i>Amount of subordination</i> | | |
| Dependent clauses per clause | DC/C | # of dependent clauses/# of clauses |
| Dependent clauses per T-unit | DC/T | # of dependent clauses/# of T-units |
| <i>Amount of coordination</i> | | |
| Coordinate phrases per clause | CP/C | # of coordinate phrases/# of clauses |
| Coordinate phrases per T-unit | CP/T | # of coordinate phrases/# of T-units |
| T-units per sentence | T/S | # of T-units/# of sentences |
| <i>Degree of phrasal sophistication</i> | | |
| Complex nominals per clause | CN/C | # of complex nominals/# of clauses |
| Complex nominals per T-unit | CN/T | # of complex nominals/# of T-units |

T-units per clause (CT/C), complex T-units per T-unit (CT/T), clauses per sentence (C/S), and verb phrases per T-unit (VP/T) – were not included in this study because they were shown to be poor candidates for developmental indices (see Lu 2011). The remaining 10 measures and their formulas are summarized in Table 1. Lu (2010, 2011) also emphasized the importance of providing and using explicit and consistent definitions of the production units and syntactic structures that are involved in calculating one or more of the syntactic complexity measures. The definitions of the six production units and syntactic structures involved in the measures examined in the present study are recaptured below.

1. Sentence: A sentence is defined as a group of words (including sentence fragments) punctuated with a sentence-final punctuation mark, including a period, exclamation mark, question mark, and occasionally elliptical marks or closing quotation marks.
2. Clause: A clause is a structure with a subject and a finite verb, including independent, adjective, adverbial, and nominal clauses, but not non-finite verb phrases, which are included in the definition of verb phrases instead (Hunt 1965; Polio 1997).
3. Dependent clause: A dependent clause is defined as a finite adverbial, adjective, or nominal clause (Hunt 1965; Cooper 1976; Kameen 1979).
4. T-unit: A T-unit consists of a main clause and any dependent clause or non-clausal structure attached or embedded in it (Hunt 1970).

5. Coordinate phrase: Coordinate adjective, adverb, noun, and verb phrases are counted as coordinate phrases.
6. Complex nominal: Complex nominals include (1) noun phrases with one or more of the following pre- or post-modifiers: adjective, possessive, prepositional phrase, adjective clause, participle, or appositive; (2) nominal clauses, and 3) gerunds and infinitives in subject position (Cooper 1976).

4. Method

4.1 Data

The present study draws English writing samples by NNS and NS university students from the *Written English Corpus of Chinese Learners Version 2.0* (WECCL 2.0) (Wen et al., 2008) and the *Louvain Corpus of Native English Essays* (LOCNESS) (Granger 1996), released by the Centre for English Corpus Linguistics (CECL), *Université Catholique de Louvain*.

WECCL 2.0 consists of 4,950 essays written by EFL learners from more than 20 different colleges in China. Both English major students and non-English major students from all four school years in college are represented. A total of 27 different topics were used across the corpus, including 26 topics for argumentative essays, and one for expository essays. The prompts were generally brief, and those for argumentative essays presented either one view or two opposing views on an issue and asked the students to state their own views, e.g. "Some people think that education is a life-long process, while others don't agree. Write an essay to state your own opinion". Each essay in the corpus is annotated with a header that encodes information about the genre (argumentative or expository), topic and timing condition (timed or untimed) of the essay; the school level (first, second, third, or fourth year in college) and year of admission (2003 through 2007) of the student; and whether the student majors in English.

LOCNESS comprises 436 essays written by native English speakers on a range of topics, including 232 by American university students, 90 by British university students, and 114 by British A-Level (General Certificate of Education Advanced Level) students. This corpus is originally collected as a control corpus for comparing college-level NNS writing data in the International Corpus of Learner English (Granger et al. 2009) with a native English corpus. WECCL 2.0 is used in the present study instead of ICLE because it affords a better diversity of proficiency levels. Given that the age group of students represented and the types of essays included in LOCNESS are comparable to those in WECCL 2.0, LOCNESS constitutes an appropriate control corpus for comparing data in WECCL 2.0 with a native English corpus, too.

4.2 Research questions

Drawing upon data from the two corpora described above, the present research aims to compare whether and the extent to which NNS and NS university students' writing differ in syntactic complexity, conceptualized as a multifaceted construct encompassing length of production unit, amount of subordination, amount of coordination, and degree of phrasal sophistication. Specifically, we seek to answer the following two research questions:

1. Are there significant differences between NNS and NS university students' writing in syntactic complexity, and if yes, in which aspects and to what degree?
2. Is the writing of NNS university students at higher proficiency levels closer to that of NS university students in syntactic complexity than the writing of NNS university students at lower proficiency levels?

4.3 Analysis

To ensure homogeneity of the NNS and NS groups, only essays written by English majors were sampled from WECCL 2.0, and only essays written by American university students were sampled from LOCNESS. Our final dataset consisted of a total of 600 essays, including 200 essays randomly sampled from each of the following three student groups: (1) first- and second-year English major students (100 each) in WECCL 2.0 (NNS-low); (2) third- and fourth-year English major students (100 each) in WECCL 2.0 (NNS-high); and (3) American university students in LOCNESS (NS). In so doing, we considered Chinese EFL learners in the first two years of college to be at a relatively lower proficiency level than those in the last two years of college. This loose conceptualization of proficiency aligns with the curriculum expectations for English majors in the first and last two years of college. For example, all English majors in four-year colleges in China are required to sit for and pass the *Test for English Majors Band 4* (TEM-4), an obligatory high-stakes English proficiency test, towards the end of the second year and then sit for *Test for English Majors Band 8* (TEM-8) towards the end of the fourth year. Table 2 summarizes the details of the essays sampled from the three groups. Note that the NS essays are considerably longer than the NNS essays. This difference, however, should not affect the type of comparison being pursued here, as the syntactic complexity measures considered are all computed as ratios of one syntactic structure to another in complete texts. The essays in the final dataset were analysed using the *L2 Syntactic Complexity Analyser* (Lu 2010). Given an essay in plain text format as input, the analyzer returns 14 indices of syntactic complexity of the essay based on 14 different measures, including all of the 10 measures considered in the present study.

Table 2. Summary of data

| Group | NNS-Low | NNS-High | NS |
|------------------------------|---------|----------|---------|
| Number of essays | 200 | 200 | 200 |
| Average length of essay | 251.90 | 289.59 | 812.95 |
| Standard deviation of length | 77.12 | 85.40 | 464.60 |
| Total number of words | 50,380 | 57,918 | 162,590 |

5. Results and discussion

5.1 Research question 1

Table 3 summarizes the mean values of syntactic complexity measures of the following four groups: NNS-low, NNS-high, NNS (i.e. the combination of NNS-low and NNS-high), and NS. In the first research question, we are focusing on determining whether there is any significant difference in syntactic complexity between the NNS group and the NS group, and if yes, in which aspects. As the results in Table 3 show, for 9 of the 10 syntactic complexity measures (i.e. all but the T/S measure), the mean value of the NNS group is lower than that of the NS group. An independent-samples *t* test is run to determine whether the mean complexity values for the NNS and NS groups differ significantly. Given that we are investigating 10 measures and therefore performing 10 tests on the same dataset simultaneously, we employ the Bonferroni correction to avoid spurious positives. This sets the alpha value for each comparison to $.05/10$, or $.005$, where $.05$ is the significance level for the complete set of tests, and 10 is the number of individual tests being performed. The *t* test results are summarized in the “NNS vs. NS” column of Table 4. These results reveal statistically significant differences ($p < .005$) in the mean values of 8 out of the 10 measures (i.e. all but two coordination measures, CP/C and T/S) between the NNS group and the NS group.

Our results suggest significant differences in all four aspects of syntactic complexity between NNS and NS university students' writing, including length of production unit, amount of subordination, amount of coordination, and degree of phrasal sophistication. In terms of length of production unit, the mean lengths of clause, sentence, and T-unit of the NNS group are all significantly shorter than those of the NS group. These results appear to be consistent with the results reported by Foster & Tavakoli (2009), who found that NNS students produced significantly shorter utterances than NS students in oral narratives. For amount of subordination (measured using DC/C and DC/T), our results indicate that the NNS group uses a significantly smaller proportion of dependent clauses and

complex T-units than the NS group. These results are again consistent with the results reported by Foster & Tavakoli (2009), who found that NNS group engaged a significantly smaller amount of syntactic subordination than the NS group in oral narratives. With respect to amount of coordination, our results show that the NNS group differs from the NS group only on the amount of phrasal coordination (measured using CP/T), but not on the amount of sentential coordination (measured using T/S). Finally, the results also show that the NNS group uses a significantly smaller proportion of complex nominals than the NS group.

5.2 Research question 2

Having established that there are statistically significant differences in syntactic complexity between NNS and NS university students' writing, we further examine whether the writing of NNS university students at higher proficiency levels better approximates that of NS university students in syntactic complexity than the writing of NNS university students at lower proficiency levels.

As can be seen in Table 3, except for the CP/C and T/S measures, the mean values of all of the other 10 syntactic complexity measures increase linearly across the following three groups: NNS-low, NNS-high, and NS. A one-way ANOVA is run to determine whether the mean complexity values for the NNS-low, NNS-high, and NS groups differ significantly. As is the case with the independent-samples

Table 3. Mean complexity values for the NNS and NS groups

| Measure | Code | NNS-L | NNS-H† | NNS | NS |
|---|------|--------|--------|--------|--------|
| <i>Length of production unit</i> | | | | | |
| Mean length of clause | MLC | 9.041 | 9.423 | 9.232 | 9.942 |
| Mean length of sentence | MLS | 15.132 | 16.632 | 15.882 | 19.153 |
| Mean length of T-unit | MLT | 13.531 | 14.815 | 14.173 | 17.072 |
| <i>Amount of subordination</i> | | | | | |
| Dependent clauses per clause | DC/C | 0.319 | 0.346 | 0.333 | 0.404 |
| Dependent clauses per T-unit | DC/T | 0.508 | 0.568 | 0.538 | 0.726 |
| <i>Amount of coordination</i> | | | | | |
| Coordinate phrases per clause | CP/C | 0.235 | 0.231 | 0.233 | 0.254 |
| Coordinate phrases per T-unit | CP/T | 0.350 | 0.365 | 0.358 | 0.430 |
| T-units per sentence | T/S | 1.118 | 1.126 | 1.122 | 1.121 |
| <i>Degree of phrasal sophistication</i> | | | | | |
| Complex nominals per clause | CN/C | 0.966 | 1.064 | 1.015 | 1.222 |
| Complex nominals per T-unit | CN/T | 1.446 | 1.669 | 1.558 | 2.089 |

† NNS-L and NNS-H denote the NNS-low and NNS-high group, respectively.

Table 4. Differences in mean complexity values among the NS, NNS-low, NNS-high, and NNS groups

| Measure | Code | NNS vs. NS | NNS-L vs. NS | NNS-H† vs. NS | NNS-L vs. NNS-H |
|---|------|---------------|-----------------|------------------|--------------------|
| <i>Length of production unit</i> | | | | | |
| Mean length of clause | MLC | * | * | * | – |
| Mean length of sentence | MLS | * | * | * | * |
| Mean length of T-unit | MLT | * | * | * | * |
| <i>Amount of subordination</i> | | | | | |
| Dependent clauses per clause | DC/C | * | * | * | – |
| Dependent clauses per T-unit | DC/T | * | * | * | – |
| <i>Amount of coordination</i> | | | | | |
| Coordinate phrases per clause | CP/C | – | – | – | – |
| Coordinate phrases per T-unit | CP/T | * | * | * | – |
| T-units per sentence | T/S | – | – | – | – |
| <i>Degree of phrasal sophistication</i> | | | | | |
| Complex nominals per clause | CN/C | * | * | * | – |
| Complex nominals per T-unit | CN/T | * | * | * | * |

† NNS-L and NNS-H denote the NNS-low and NNS-high group, respectively. * indicates a statistically significant difference ($p < .005$); – indicates a non-significant difference ($p \geq .005$)

t test discussed above, the same Bonferroni correction is applied here, which again adjusts the alpha value for each of the 10 comparisons to .05/10, or .005. The one-way ANOVA shows statistically significant differences ($p < .005$) in the mean values for 8 of the 10 measures (i.e. all but the CP/C and T/S measures) among the NNS-low, NNS-high, and NS groups. For the 8 measures that show statistically significant between-group differences, the Bonferroni test, a post hoc multiple comparison test, is run to determine whether significant differences exist between any two of the three groups. Results of the post hoc Bonferroni tests are summarized in the last three columns in Table 4.

5.2.1 Length of production unit

The results show significant differences in the mean values of all three measures of length of production unit, i.e. MLC, MLS, and MLT, between the NNS-low group and the NS group as well as between the NNS-high group and the NS group. The mean values of MLS and MLT both increase significantly from the NNS-low group to the NNS-high group. However, although the mean value of the MLC measure also increases from the NNS-low group to the NNS-high group, this increase is not found to be statistically significant. These results suggest that,

in terms of the average length of sentences and T-units (but not clauses) produced in writing, NNS university students at higher proficiency levels approximate NS university students significantly better than those at lower proficiency levels. This is largely consistent with previous observations that the average length of production units generally increases with proficiency levels (e.g. Wolfe-Quintero et al. 1998; Lu 2011).

5.2.2 *Amount of subordination*

The results reveal significant differences in the mean values of the two measures that reflect the amount of subordination, i.e. DC/C and DC/T, between the NNS-low group and the NS group as well as between the NNS-high group and the NS group. However, although there is an increase in the amount of subordination used from the NNS-low group to the NNS-high group, this increase is not found to be statistically significant for either of the two subordination measures. These results suggest that NNS university students at higher proficiency levels do not use significantly more subordination in writing than those at lower proficiency levels. Lu (2011) found that NNS university students' use of subordination in timed argumentative essays increases significantly from the first year to the second year, but then decreases from the second year to the fourth year. The results in the present study differ slightly from those reported in Lu (2011) in that a non-significant increase, instead of decrease, is found from the NNS-low group to the NNS-high group. Nevertheless, both results show that NNS university students seem not to engage syntactic complexity more at the clausal level as they advance to higher levels of proficiency.

5.2.3 *Amount of coordination*

With respect to the amount of coordination, significant differences are found in the mean values of the CP/T measure between the NNS-low and the NS group as well as between the NNS-high and the NS group, but not between the NNS-low group and the NNS-high group. No significant differences are found in the mean values of the CP/C and T/S measures among the three groups. These results suggest that NNS university students use significantly fewer coordinate phrases per T-unit than NS university students in writing, but not coordinate phrases per clause or T-units per sentence. Furthermore, NNS university students at lower and higher proficiency levels do not differ significantly from each other in terms of the amount of coordination they use in writing. These results are consistent with Lu's (2011) finding that the T/S measure does not differentiate between school levels, but differ somewhat from Lu's (2011) finding that both the CP/C and CP/T measures discriminate certain nonadjacent school years.

5.2.4 *Degree of phrasal sophistication*

Finally, significant differences are found in the mean values of the two measures that gauge the degree of phrasal sophistication between the NNS-low group and the NS group as well as between the NNS-high group and the NS group. A significant increase from the NNS-low group to the NNS-high group is observed for the CN/T measure, indicating that NNS university students at higher proficiency levels use more complex nominals per T-unit than those at lower proficiency levels and therefore better approximate NS university students in this regard. The mean value of the CN/C measures also increases from the NNS-low group to the NNS-high group, but the increase is not found to be statistically significant. These results are again largely consistent with the findings reported in Lu (2011), although the CN/C measure was also found to discriminate certain adjacent and non-adjacent school years in that study.

6. Conclusions and implications

Adopting a corpus-based cross-sectional research design, this study has provided a comprehensive comparison of syntactic complexity in NNS and NS university students' writing in the following four areas: length of production unit, amount of subordination, amount of coordination, and degree of phrasal sophistication. This comparison has allowed us to answer important questions on whether and the extent to which NNS and NS university students' writing differ in each of these four areas of syntactic complexity, and whether NNS university students' writing better approximates the writing of NS university students in syntactic complexity as they advance to higher levels of proficiency in the target language.

Our results showed significant differences in all four areas of syntactic complexity between NNS and NS university students' writing. On average, NNS university students produce shorter clauses, sentences, and T-units, a smaller amount of subordination, and a smaller proportion of complex nominals than NS university students. They also use a smaller number of coordinate phrases per T-unit. No significant difference is found in sentential coordination between the NNS group and the NS group, suggesting that college-level English learners have largely mastered this structure. This is probably not too surprising, as sentence coordination is usually introduced in the early stages of English instruction. The same patterns of difference between NNS and NS students apply to both NNS university students in the lower and higher proficiency groups. However, there is some evidence that the higher-proficiency NNS group approximates the NS group significantly better than the lower-proficiency NNS group in the areas of length of production unit and degree of phrasal sophistication.

Our results on NNS-NS differences in syntactic complexity are consistent with previous findings reported by Foster and Tavakoli (2009), but more comprehensive. The results on differences between the lower and higher proficiency NNS groups are also largely consistent with the results Lu (2011) reported on the differences among NNS students in different school years, with some exceptions. In particular, some measures that were found to discriminate school levels by Lu (2011) were not found to discriminate the lower and higher proficiency NNS groups in this study, such as MLC and CN/C. These differences could be due to the division of the NNS university students into different number of groups in the two studies: two in the present study and four in Lu (2011) (i.e. one for each school year). In addition, Lu (2011) used all timed argumentative essays from one specific institution, whereas the present study randomly sampled 400 essays from the entire corpus.

Findings from this study point to the importance for second language writing teachers to be aware of the significant gap in all four aspects of syntactic complexity between NNS university students at both low and high proficiency levels and NS university students. This gap calls for the design of relevant pedagogical interventions to enhance NNS university students' syntactic development.

Given the scope of this research and the information available in the WECCL 2.0 and LOCNESS corpora, several important issues were not taken up in this study. First, the WECCL 2.0 learner corpus contains samples produced by L1 Chinese learners only. Future studies will benefit from examining whether the same patterns of difference are found between NS students and NNS students of other L1 backgrounds. This investigation will allow us to ascertain the effect of L1 on NNS students' syntactic development. Second, it will be useful to systematically examine the effects of different learner-, task-, and context-related factors (e.g. timing condition and institution, among others) on NNS-NS differences in syntactic complexity and to control for factors with significant effects in the analysis. Finally, it would also be very useful to assess whether and the conditions under which highly advanced NNS writers eventually obtain the same level of syntactic complexity as NS writers.

References

- Beers, S.F. & Nagy, W.E. 2009. Syntactic complexity as a predictor of adolescent writing quality: Which measures? which genre? *Reading and Writing* 22(2): 185–200.
- Buckingham, T. 1979. The goals of advanced composition instruction. *TESOL Quarterly* 13(2): 241–254.
- Casanave, C.P. 1994. Language development in students' journals. *Journal of Second Language Writing* 3(3): 179–201.

- Cooper, T.C. 1976. Measuring written syntactic patterns of second language learners of German. *The Journal of Educational Research* 69(5): 176–183.
- Ellis, R. & Yuan, F. 2004. The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition* 26: 59–84.
- Ferris, D.R. 1994. Rhetorical strategies in student persuasive writing: Differences between native and non-native English speakers. *Research in the Teaching of English* 28(1): 45–65.
- Foster, P. & Tavakoli, P. 2009. Native speakers and task performance: Comparing effects on complexity, fluency and lexical diversity. *Language Learning* 59(4): 866–896.
- Granger, S. 1996. From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In *Languages in Contrast*, K. Aijmer, B. Altenberg & M. Johansson (eds), 37–51. Lund: Lund University Press.
- Granger, S., Dagneaux, E., Meunier, F. & Paquot, M. 2009. *International Corpus of Learner English* (Version 2.0). Louvain-la-Neuve: Presses Universitaires de Louvain.
- Hinkel, E. 2003. Simplicity without elegance: Features of sentences in L1 and L2 academic texts. *TESOL Quarterly* 37(2): 275–301.
- Hunt, K.W. 1965. *Grammatical Structures Written at Three Grade Levels*. Champaign IL: National Council of Teachers of English.
- Hunt, K.W. 1970. Do sentences in the second language grow like those in the first? *TESOL Quarterly* 4(3): 195–202.
- Kameen, P.T. 1979. Syntactic skill and ESL writing quality. In *On TESOL '79: The Learner in Focus*, C. Yorio, K. Perkins & J. Schachter (eds), 343–364. Washington DC: TESOL.
- Larsen-Freeman, D. 1978. An ESL index of development. *TESOL Quarterly* 12: 439–448.
- Larsen-Freeman, D. 2009. Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics* 30(4): 579–589.
- Lu, X. 2010. Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics* 15(4): 474–496.
- 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.
- Norrby, C. 2007. The interaction of complexity and grammatical processability: The case of Swedish as a foreign language. *International Review of Applied Linguistics in Language Teaching* 45(1): 45–68.
- Norris, J.M. & Ortega, L. 2009. Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics* 30(4): 555–578.
- Ortega, L. 2000. Understanding syntactic complexity: The measurement of change in the syntax of instructed L2 Spanish learners. PhD dissertation, University of Hawaii.
- Ortega, L. 2003. Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics* 24(4): 492–518.
- Perkins, K. 1983. On the use of composition scoring techniques, objective measures, and objective tests to evaluate ESL writing ability. *TESOL Quarterly* 17(4): 651–671.
- Polio, C.G. 1997. Measures of linguistic accuracy in second language writing research. *Language Learning* 47(1): 101–143.
- Reid, J. 1992. A computer text analysis of four cohesion devices in English discourse by native and nonnative writers. *Journal of Second Language Writing* 1(2): 79–107.
- Silva, T. 1993. Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL Quarterly* 27(4): 657–677.
- Sotillo, S.M. 2000. Discourse functions and syntactic complexity in synchronous and asynchronous communication. *Language Learning and Technology* 4(1): 82–119.

- Stockwell, G. 2005. Syntactic and lexical development in NNS-NNS asynchronous CMC. *The JALT CALL Journal* 1(3): 33–49.
- Stockwell, G. & Harrington, M. 2003. The incidental development of L2 proficiency in NS-NNS email interactions. *CALICO Journal* 20(2): 337–359.
- Tavakoli, P. & Foster, P. 2008. Task design and second language performance: The effect of narrative type on learner output. *Language Learning* 58(2): 439–473.
- Way, D.P., Joiner, E.G. & Seaman, M.A. 2000. Writing in the secondary foreign language classroom: The effects of prompts and tasks on novice learners of French. *The Modern Language Journal* 84(2): 171–184.

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY