Measuring Syntactic Complexity in Spoken and Written Learner Language: Comparing the Incomparable?

Pekka Lintunen
University of Turku
pekka.lintunen@utu.fi

Mari Mäkilä
University of Turku
mkmaki@utu.fi

Abstract

Spoken and written language are two modes of language. When learners aim at higher skill levels, the expected outcome of successful second language learning is usually to become a fluent speaker and writer who can produce accurate and complex language in the target language. There is an axiomatic difference between speech and writing, but together they form the essential parts of learners' L2 skills. The two modes have their own characteristics, and there are differences between native and nonnative language use. For instance, hesitations and pauses are not visible in the end result of the writing process, but they are characteristic of nonnative spoken language use. The present study is based on the analysis of L2 English spoken and written productions of 18 L1 Finnish learners with focus on syntactic complexity. As earlier spoken language segmentation units mostly come from fluency studies, we conducted an experiment with a new unit, the U-unit, and examined how using this unit as the basis of spoken language segmentation affects the results. According to the analysis, written language was more complex than spoken language. However, the difference in the level of complexity was greatest when the traditional units, T-units and AS-units, were used in segmenting the data. Using the U-unit revealed that spoken language may, in fact, be closer to written language in its syntactic complexity than earlier studies had suggested. Therefore, further research is needed to discover whether the differences in spoken and written learner language are primarily due to the nature of these modes or, rather, to the units and measures used in the analysis.

Keywords: EFL, complexity, written language, spoken language

1. Introduction

The mastery of a second language (L2) usually refers to high proficiency in both spoken and written communication. Spoken and written language are seen as two modes of production. In second language acquisition (SLA) research, it is common to focus on one or the other. As the two modes are considered intrinsically different, common methods for comparing spoken and written production are not always easy to find. However,

spoken and written skills are, in addition to the receptive skills of reading and listening, two important aspects of a learner's L2 proficiency. Therefore, it is important to pay attention to the methods used in analysing these two skills. If we assume, for example, that a learner has achieved a higher proficiency in their spoken than written skills, we should also be able to measure the difference objectively.

Complexity, accuracy and fluency (or CAF) are nowadays seen as the qualitative dimensions of language use (Housen, Kuiken and Vedder 2012b). A proficient language user produces fluently accurate and complex language. The interplay between these dimensions interests many researchers, but they are also studied separately. This study focuses on complexity in spoken and written learner language. The concept of complexity is the most challenging component of the CAF framework to define and operationalise (e.g. Pallotti 2009: 592; Housen, Kuiken and Vedder 2012a: 10). With the help of this concept, researchers describe and study the proficiency of the L2 learner (Housen and Kuiken 2009: 461), and syntactic complexity measures have, for instance, been used to evaluate L2 writing development (Ortega 2003: 92). So far, in the majority of CAF studies, the focus has been on accuracy, fluency or the developmental aspect of complexity, whereas the present study focuses on syntactic complexity in two modes of production.

The purpose of this study is twofold: we want to examine the nature of syntactic complexity in spoken and written L2 production, and secondly, explore how the choice of the segmentation unit affects the complexity measure results. The data come from informal spoken monologues and short essays produced by the same subjects. Our theoretical underpinnings come from three partly related areas: the differences between spoken and written language, the concept of complexity as a quality dimension of learner language, and, finally, a critical review of the methods used in measuring L2 complexity in earlier studies.

2. Theoretical framework

2.1 The two modes: spoken and written language

The approaches to the relationship between spoken and written language have varied greatly with regard to the theoretical standpoint. Some linguists consider spoken and written language as closely related modes of production; others do not recognise the relationship at all (Cleland and Pickering 2006: 185). Many linguists have studied the fundamental differences between spoken and written language (e.g. Halliday 1979, 1989; Beaman 1984; Bourdin and Fayol 1994), and some of these studies have shown that writing is more demanding than speaking (e.g. Chafe 1982). However, in such studies, writing and speaking are often considered as cognitive tasks, and writing demands more cognitive resources. Moreover, higher expectations are set for accuracy in writing, whereas spoken language allows many inaccuracies in form. In the present study, the focus was set on the end product instead of the use of cognitive resources in real-time processing. As Cleland and Pickering (2004: 186) suggest, one of the differences in the end products is the complexity of the constructions used.

Baron (2000) has presented three approaches to the relationship between speech and writing. The *Opposition View* accentuates the difference between the modes of production, such as writing being highly structured, syntactically complex and formal, whereas speech being loosely structured, syntactically simple and informal. The *Cross-over View* brings the different types of writing and speech to the centre of focus, and the boundary between the modes blurs as features of one mode are adopted in the other, such as in "talking books". The *Continuum View* takes the surrounding context into consideration and places different types of writing and speech on a continuum. Written texts, such as law texts and academic writing, could be placed at the literary end of this continuum, whereas informal conversation could be at the oral end. These ends of the continuum differ greatly from each other, and therefore, the greatest differences in structural complexity can be found when comparing, for instance, casual conversations and academic writing where the differences in the way of producing, transmitting, and receiving the language are notable (see further e.g. Tanskanen 2006: 74–80). The contradictory results of the differences between spoken and written complexity in earlier studies are largely due to extraneous variables: in addition to the mode, the studies have also included the register or style of the productions as variables (see e.g. Beaman 1984: 51). That is, the productions studied have represented styles or registers that are too far from each other on the continuum.

Leech, Deuchar and Hoogenraad (1982: 139–140) present the typical features of speech and writing. In writing, these features include, for instance, explicitness, clear sentence boundaries, more complex structure and features reflecting formality, whereas speech is characterised by inexplicitness, lack of clear sentence boundaries, simple structure and interactional features. Such features vary between and within the mode of production (for a thorough description of the grammatical differences across registers of speech and writing, see Biber, Conrad and Leech 2002). Moreover, researchers have discovered that written texts contain complex syntactic structures, whereas in spoken language subordinate clauses are rare and occurrences of hesitations and unfinished utterances frequent, which makes spoken syntax less complex (Brown and Yule 1983: 1–10, Pietilä 1999: 6–7). Zhang (2013: 835) also notes that the sentences used in written language tend to be longer and more complex than the corresponding units used in spoken language.

However, for example Halliday (1979, 1989: 76, 79) points out that speech is no less structured (or less complex, in a sense) than written language, although its complexity is of different nature as the grammar itself is more complex in speech, while lexical complexity is more evident in writing. Writers usually aim to be precise and concise in their writing, whereas in speech ideas are elaborated "online", which can result in complex chains of thoughts (Beaman 1984: 50–51). Halliday (1989), while focusing on native language (L1), claims that syntactic structures can be even more complex in speech than they are in writing. In the studies that support this idea, the language studied is often L1 and the focus is on embedding, while the other aspects of complexity, such as clausal complexity, are neglected.

In the present study, these fundamental differences between speech and writing were taken into consideration with due regard paid to the difficulties in comparing the two different and much varied modes. However, it can also be noted that the difference between the two modes of production manifested in our study can be expected to be

somewhat smaller as the spoken samples were not dialogues, but planned monologues, and thus, contain a relatively limited number of characteristics typical of interactive situations, such as repetitions and unfinished utterances (Pietilä 1999: 7), and the written part consisted of rather informal essays.

It should also be noted in the context of L2 learning and teaching that writing has traditionally had a dominating role in educational systems around the world (see e.g. Zhang 2013: 836). This has an inevitable effect on the language that students learn. Additionally, teaching materials are mostly based on writing, and authentic spoken materials are less frequently used. In addition, the way of learning L2 (formal vs. informal) can have a great effect on the language skills in different domains: for example, learning through formal teaching may result in a better knowledge of grammatical rules and structures, whereas in the process of learning through informal L2 acquisition, understanding and speaking skills typically develop faster (Ortega 2009: 80). Our study focused on learners benefitting from rather formal and traditional education process, which, as it may be expected, would strengthen and emphasise their written skills.

2.2 L2 complexity

The triad of CAF and its theoretical applications stem from the fundamental question of the characteristics of a proficient language user (Housen, Kuiken and Vedder 2012a: 2). Although commonly used, CAF has its controversial aspects, one of the most disputable ones being the interplay of its components. Ellis and Barkhuizen (2005: 139–145) note that the component parts are interdependent, but as for the degree of interdependency, there is no clear consensus in relevant literature (see further, e.g. Skehan 1998; Robinson 2001; Gilabert 2006; Towell 2007; Ellis 2008; Tonkyn 2012). In general, L2 proficiency is just understood as multicomponential. The analysis of learner language using one of the components has been justified, and the notions of the triad can be used separately in the analysis of the L2 system and its development.

According to Bulté and Housen (2012: 22), the contradictory results in selected complexity studies can be partly explained by the vagueness of metalinguistic definitions, which relate to a varied number of different aspects that can be measured. For instance, Ellis and Barkhuizen (2005: 152–156) categorise the types of complexity measures according to the object of measurement into interactional, propositional, functional, grammatical and lexical. As pointed above, in the present study, the focus is solely on *grammatical complexity* (i.e. structural or syntactic complexity).

At a very basic level, according to Bulté and Housen (2012: 22), a satisfying definition is that complexity can refer to "(1) the number and the nature of the discrete components that the entity consists of, and (2) the number and the nature of the relationships between the constituent components". A frequently used definition for complexity is that the language user has the ability to produce linguistically, and thus cognitively, more demanding linguistic material (e.g. longer units with more complex embedding elements) (Pallotti 2009: 593). However, it is important to note that complexity cannot be totally paralleled with the difficulty of production. According to Pallotti, linguistic variation can be an important part of the notion of complexity and

some linguists add to this definition the fact that language learners learn to use such cognitively demanding material rather late in their learning process. However, Pallotti (2009: 593–594) also notes that the developmental aspect needs to be separated from the complexity definition due to the difference between the product and the process with complexity being a characteristic of linguistic production and development belonging in the process of learning.

In short, although relevant definitions are much varied, it is also a fact that most researchers believe that the notion of complexity can be used to describe the structural characteristics of learner language and to study linguistically demanding production. Thus, for the purposes of the present study the concept of complexity has been defined as follows: complexity is the use of linguistically demanding language, considering both its quantitative (length, ratios, frequency) and qualitative aspects (dependent clause type).

2.3 Measuring L2 complexity

According to Ellis and Barkhuizen (2005: 154), a common complexity measure is grammatical complexity, introduced as "grammatical variation and sophistication" (Wolfe-Quintero, Inagaki and Kim 1998: 69) or, if only a certain aspect such as embedding is considered, "the number, type and depth of embedding in a text" (Beaman 1984: 45). Thus, syntactic complexity means that varying structures with complex elements, such as embedded dependent clauses, are used. The analysis of syntactic complexity aims to describe such complex structures, i.e. to determine how smaller units and simple sentences are combined into more complex structures (Holger 2004: 3), and to study various phenomena, for example the effect of instruction or task complexity on performance in L2 writing, and especially, the developmental aspect of complexity in L2 (Ortega 2003: 492).

Various studies on syntactic complexity (cf. e.g. Wolfe-Quintero, Inagaki and Kim 1998, Norris and Ortega 2009) have also examined the ways of measuring syntactic complexity. As Hunt (1965) states, *more* is often considered to be a sign of complexity. Thus, length, amount of embedding, and frequency of certain sophisticated structures (e.g. non-finite clauses) can function as a basis for syntactic complexity. It follows that long production units (e.g. elongated sentences) can be considered more complex than short units. The number of subordinated structures is one of the most frequently used measures in syntactic complexity studies, especially in L2 studies (Ellis and Barkhuizen 2005: 154). Holger (2004: 3) notes that complex sentences originate from simple sentences that are gradually linked together, through coordination and subordination. This linking of production units makes the language more complex. Some linguistic elements are also considered more complex than others (e.g. infinitival phrases) (Bulté and Housen 2012: 31). Higher frequency of such elements can be seen as a sign of complex language use.

Bergman and Abrahamsson (2004: 611) have created a three-level scale for describing the syntactic structures in L2. For beginners, the sentence structures are simple and only the basic linking elements (such as *and, but, then*) are present. At the intermediate level, the use of complex sentence structures increases, especially

concerning dependent clauses, with variation in the use of linking elements. At this level, non-finite clauses begin to appear in learners' production. The advanced learners' level is characterised by a varied use of different sentence structures with multiple dependent and non-finite clauses. The subjects of the present study were all from a group with a target intermediate level, i.e. the mid-level of the above-mentioned scale, which led to the assumption that the syntactic structures used within the group should be quite complex.

Earlier studies often focused on the development of syntactic complexity (e.g. Bardovi-Harlig 1992; Mellow 2006) or the effect of task types on L2 written (e.g. Ishikawa 1995, 2007; Storch and Wigglesworth 2007; Kuiken and Vedder 2007; Robinson 2007) and spoken complexity (e.g. Skehan and Foster 2005; Tavakoli and Foster 2008). Syntactic complexity has also been studied by comparing spoken and written material. For instance, in the 1960s, studies comparing spoken and written language concentrated on word frequency counts (related to the length of sentences) and resulted in finding that writing was more complex than speech. It should be reiterated in this context that the existing inconsistency, or even contradictions, in research conclusions may have resulted from the inadequately phrased definitions (e.g. Tanskanen 2006: 74).

Silva, Abchi and Borzone (2010) studied the L1 syntactic complexity in oral and written retellings by Spanish children. In their analysis, the length and number of T-units and the number of subordinated clauses per T-unit were used as measures. The first two measures revealed a difference between the modes, whereas the last one did not. Beaman (1984) studied L1 syntactic complexity by comparing 20 spoken and written narratives with focus on coordination and subordination. Her results supported Halliday's (1979) earlier proposal that if subordination is the most important indicator of syntactic complexity, spoken production is as complex as written production.

Another example is provided by Larsen-Freeman (2006), who studied the development of L2 complexity, accuracy and fluency in the spoken and written production of five Chinese (higher-) intermediate learners of English. In her study, she repeated the same type of tasks with the same subjects four times during a six-month period. Syntactic complexity was measured with the average number of clauses per T-unit. The limited number of subjects and only one measure used in the assessment of complexity were the limitations of this study, but the main (quantitative) finding was that every CAF domain improved at the group level although individual differences were still significant (Larsen-Freeman 2006: 598–560).

To reiterate, different methods and definitions have been common in earlier studies. For example, there are varying definitions for a *clause* in relevant literature (cf. e.g. Iwashita 2006: 159; Ishikawa 2007: 142; Vyatkina 2013: 18), and quite often this unit is not at all defined (e.g. Polio 2001: 97; Bulté and Housen 2012: 39). Hunt (1965: 15) defines a clause as "a structure with a subject and a finite verb (a verb with a tense marker)", while Bulté and Housen (2012: 39) note that the disadvantage of this definition is that it excludes an essential use of complex structures, i.e. non-finite clauses, which are important and may further affect research results when certain verb constructions are analysed as two clauses. For the present study, the clause is defined after Foster, Tonkyn and Wigglesworth (2000), i.e. as a structure which does not need to include a finite verb. A clause is a structure that consists of a verbal element plus an

additional clause element, for example an object or an adverbial (Foster, Tonkyn and Wigglesworth 2000: 366).

In written language analysis, the category of the *sentence* has often been used in segmenting data into units. However, when analysing learner language, the use of sentence as a comparative unit could be challenging as punctuation is not always consistent. Hunt (1965) presents the *minimal terminable unit* (T-unit) as a valid comparative unit for measuring syntactic complexity in L1 writing development and Hunt (1965: 20, 49) defines such a unit as an entity that consists of one main clause and (optional) subordinate clauses (i.e. dependent clauses) and non-clausal units or sentence fragments attached to it. This means that a traditional sentence with two coordinated main clauses should be segmented into two T-units.

The concept of the T-unit has been used in several L2 studies (e.g. Bardovi-Harlig 1992: 390) and researchers have generally been content with the unit (Gaies 1980: 53– 54). As coordination is characteristic of lower proficiency levels, the sentence has been reintroduced as a comparative measure in studies with subjects of higher proficiency (Bardovi-Harlig 1992: 390). Gaies (1980: 59) notes, however, that the T-unit is usable also at higher proficiency levels if the researcher takes the limitations into account. Following this view, in the present study, the T-unit was used to enable comparison of the results with earlier studies. In addition, the sentence was also used as a comparative unit. This enabled us to examine the ratio of coordinated structures and the measure of sentence complexity ratio, and to study the effect of the segmentation unit on the complexity measures.

In selected studies, especially those focused on comparing spoken and written production, the T-unit was adopted as a unit for spoken data (e.g. Halleck 1995; Pietilä 1999; Larsen-Freeman 2006). However, the units that are originally based on (L1) written language can be problematic at the time of segmenting spoken L2 data; such units do not meet the requirements that an analysis of such complex language samples as L2 speech sets for the measurements (Foster, Tonkyn and Wigglesworth 2000: 354; Ellis and Barkhuizen 2005: 147). Thus, Foster, Tonkyn and Wigglesworth (2000: 365) introduce the *Analysis of Speech Unit* (AS-unit) for dividing spoken data into analysable units. In addition to its syntactic quality, the AS-unit has features related to intonation and semantics and is defined as "a single speaker's utterance consisting of *an independent clause*, *or sub-clausal unit*, together with any *subordinate clause(s)* associated with it". This definition was adopted for the present study and the exact methods used are presented in section 3.

3. Methodology

3.1 Research questions and subjects

The reported study had two main research questions: 1. What is the nature of syntactic complexity in spoken and written L2 production? 2. How does the choice of the segmentation unit affect the complexity measure results?

The subjects were 18 upper secondary school students (9 male, 9 female), 17–18 years old (average 17.6) in South-Western Finland. As a result, the data consisted of 18 written essays and 18 transcribed spoken productions. With regard research ethics, parental consent was obtained for underage subjects.

The subjects with expected intermediate educational level were chosen to match the methodology and the measures used in the study, which appeared to be most suitable for intermediate-level learners (cf. Norris and Ortega 2009). Formally, the subjects were expected to represent intermediate-level language learners (in the range of higher B1 and lower B2 on the CEFR scale according to the curricular target levels). The individual subjects were selected with the use of a number of criteria in order to make the research group as homogenous as possible.

All subjects had Finnish as their L1, none of them were bilingual, and none had spent more than a month abroad. The subjects studied English as their first foreign language in mainstream education and had received average grades. Thus, the subjects chosen for the study were approximately at the same proficiency level (based on the combination of the educational level and school grades). It is, naturally, probable that some subjects may have been more proficient writers than speakers or vice versa. Their school grades represented their overall skills and, among other factors, may also have been affected by a particular student's diligence.

3.2 Data collection

The data for the study were collected in connection with a larger research project on CAF and L2 English. The subjects were involved in two tasks. In the written part, the subjects were asked to write an informal essay of 150–250 words on one of three given topics. In the spoken part, the subjects were shown a cartoon strip with six frames, and their task was to tell a story based on the cartoon. The subjects had two minutes to familiarise themselves with the cartoon and to plan their story before performing the task. The subjects were able to see the cartoon while telling the story. The spoken samples were transcribed, and the length of pauses was measured with the help of the waveform functions of two types of software (Amadeus Lite for Mac and Transcriber for Windows). Intonation was assessed auditorily.

The total number of words in the data was 4240, the written part being somewhat larger than the spoken part (written 2353 words and spoken 1887 words, 16.9 minutes). The number of words in the spoken data contained only the words analysed (e.g. repetitions and hesitations excluded). The mean sample length was longer in the written samples (written sample on average 130.7 words and spoken 104.8 words).

3.3 Measurements

When we focus on learner language that includes many nonstandard forms, the importance of definitions is highlighted. For instance, the notion of the *word* needs to be clarified. In the present study, compounds were analysed according to their actual form, whether correct or not (e.g. *seventeenyear olds* was counted as two words). In the spoken

samples, only the words belonging to the units analysed were included in the word count (*e.g.* repetitions were excluded). As has been mentioned above, we followed Foster, Tonkyn & Wigglesworth's (2000) definitions of clauses and AS-units. T-units and sentences were also used as comparative units.

To examine the relationship between spoken and written complexity, we also introduced a new unit. In a sense, this unit can be seen as an equivalent to a written sentence. However, the aim was not to adapt the rules of writing to spoken production. When focusing on fluency and learner development, the tone unit or mean length of run (MLR) are often used as comparative units. These are not ideal when analysing complexity as the use of short pauses as boundary signals places the emphasis on fluency rather than on complexity. The fact that non-native speakers need more time for planning their production or for searching for an appropriate word or expression should not affect the assessment of the syntactic complexity of their production. Therefore, a unit with a more flexible definition for unit boundaries is needed when studying L2 complexity. In the unit introduced and piloted in the present study, the traditional criteria were applied but the pause limit was loosened and the use of a combination of several criteria was enabled to include the context of the stretch of speech in the segmentation. The new unit, called the *modified utterance* or the *U-unit*, is closly related to the concepts of T-unit, C-unit (a semantic unit used in spoken data, a stretch of speech containing a pragmatic meaning, Pica et al. 1989: 72) and utterance (e.g. Foster, Tonkyn and Wigglesworth 2000: 359), and therefore combines syntax, semantics and intonation as a basis for segmentation. The basic idea is similar to the idea unit as defined by Ellis and Barkhuizen (2005: 154), i.e. "a message segment consisting of a topic and comment that is separated from contiguous units syntactically and/or intonationally." The definition of the U-unit is more precise when it comes to the unit boundaries. We defined the U-unit as follows:

> one independent clause or several coordinated independent clauses, with all dependent clauses or fragmental structures attached to it, separated from the surrounding speech by a pause of 1.5 seconds or more, or, especially in occurrences of coordination, a clear change in intonation and a pause of 0.5 seconds or more (depending on the average length of boundary pauses in the sample), containing one semantic unity.

The syntactic starting point is that of independent clauses, coordination included, and dependent clauses. Coordination was allowed following earlier observations of learner language (e.g. Gaies 1980; Bardovi-Harlig 1992). As stated earlier, L2 speakers need more time to plan their speech and/or to search for the correct word. In the present study, the duration of 1.5 seconds was chosen, as the unit boundary is often clearly marked when stretches of speech are separated by such a long pause. If there was a clear change in intonation, which was rather infrequent in our learner language samples, the pause could be shorter. If the end of a stretch of speech was marked by a clear (usually falling) intonation pattern and followed by a pause of 0.5 seconds, this was considered a unit boundary, in fact similar to the AS-unit (Foster, Tonkyn and Wigglesworth 2000: 367). Individual differences were taken into account by examining the overall length of pauses in the sample. If the speaker had long intraclausal planning pauses, the criterion for the unit boundary pause was lengthened accordingly. The pauses had to be silent pauses between semantic units, i.e. pausing with hesitation (or filled pauses) within a semantic

unit was not considered a unit boundary. Examples 1 and 2 illustrate how semantics affects the segmentation into U-units and how it differs from AS-units. In the examples, the end of a unit is marked with square brackets. Pause lengths are in brackets. Micropauses (shorter than 0.4 seconds) are marked as (.).

(1) some (.) age (0.7) ago (0.5) there's a (0.7) big apple tree [AS] and they (0.9) sit (1.8) under (0.6) under the tree (.) together [AS U]

(2) then (0.6) they think that (0.9) the tree is (1.5) old enough [AS] (0.5) and (.) they build (.) a house (1.0) from the tree [AS U]

In the examples, two AS-units form a semantic unit (with varying lengths in pausing), and therefore constitute one U-unit. Coordination was not an automatic boundary signal for U-units. In Example 2, the speaker has a long planning pause before the subject predicate "old enough" but continues without a longer pause or other boundary signals into the coordinated clause, which is part of the same semantically coherent U-unit.

L2 syntactic complexity and the development of sentence structure complexity have been studied by using different methods of measuring complexity (see Polio 2001: 96–97; Tonkyn 2012: 222–223). For instance, Wolfe-Quintero, Inagaki and Kim (1998: 9–11) examine ways of measuring syntactic complexity in written production: their common methods include counting the length, measuring the frequency or ratio of linguistic elements. Some other studies (e.g. Scarborough 1990; Bardovi-Harlig 1992) have used index-based formulae. Wolfe-Quintero, Inagaki and Kim (1998: 119) conclude that the most reliable measures for syntactic complexity development are clauses per T-unit and the number of dependent clauses per clause, or the number of dependent clauses per T-unit, which are all based on ratios. However, it seems that by using more measures one can achieve a more reliable result, which is not affected by the overuse of a certain unit, for example. The measures used in this study are listed in Table 1 below.

Measure type		Measure for written language	Measure for spoken language
Overall measure	Length	1. Mean length of sentence (words per sentence, W/S) 2. Mean length of T-unit (words per T-unit, W/T)	1. Mean length of U-unit (words per U-unit, W/U) 2. Mean length of AS-unit (words per AS-unit, W/AS)
Ratio	Complexity ratio	3. Sentence complexity ratio (clauses per sentence, C/S)	3. U-unit complexity ratio (clauses per U-unit, C/U)
	Coordination	4. Coordinate clauses per sentence (Coord/S)	4. Coordinate clauses per U-unit (Coord/U)

Measure type		Measure for written language	Measure for spoken language
	Subordination	5. Dependent clause ratio (dependent clauses per clause, DC/CWr)	5. Dependent clause ratio (dependent clauses per clause, DC/CSp)
Intra-clausal	Length	6. Mean length of clause (words per clause, W/CWr)	6. Mean length of clause (words per clause, W/CSp)
	Frequency	7. Number of non-finite dependent clauses (NonFWr)	7. Number of non-finite dependent clauses (NonFSp)

Table 1: Measures

In the present study, the overall measures included the length of units, as length is widely considered to be a valid measure for overall complexity (Szmrecsányi 2004; Norris and Ortega 2009: 561). The mean length of the unit was calculated by dividing the total number of words by the number of comparative units. Ratio measures contained the complexity ratio, coordination and subordination (measured in dependent clauses in the present study). These three measures were used in the present study to check how the measures functioned when the modes of production were being compared. The number of non-finite dependent clauses was included in the analysis, as earlier studies indicated that a high frequency of non-finite dependent clauses indicated more complex language (Bergman and Abrahamsson 2004: 611). We compared the two modes by comparing the corresponding measures. The measures of T-unit and AS-unit complexity (C/T, C/AS) and coordination per T-unit and AS-unit (Coord/T and Coord/AS) were not part of the actual analysis, but were calculated for the discussion on the choice of comparison unit.

SPSS was used for statistical testing. The Shapiro-Wilk's test was used for testing the normal distribution. As all the measures were not normally distributed, for clarity's sake, only non-parametric tests are reported in this study. When comparing the modes, the Wilcoxon signed-ranks test was used for statistical significance (see Larson-Hall 2010: 251, 404).

4. Results

The size of the written sample was 2353 words and the spoken sample 1887 words. The written sample contained on average 9.4 T-units, 8.4 sentences and 21.3 clauses per subject, whereas the mean number of their spoken counterparts was 12.2 AS-units, 7.3 U-units and 17.8 clauses.

	Minimum	Maximum	Mean	Std. Deviation
W/S	9.18	27.00	16.53	5.01
W/T	9.18	20.20	14.57	3.47
W/CWr	4.21	9.21	6.25	1.07
C/S	1.67	4.57	2.67	.78
Coord/S	.00	.86	.29	.26
DC/C	.20	.68	.50	.14
NonFWr	.00	9.00	4.39	2.79

Table 2: Complexity in the written production

In written production, as Table 2 shows, the mean lengths of the comparative units were 16.53 words per sentence, 14.57 words per T-unit and 6.25 words per clause. The mean sentence complexity ratio (C/S) was 2.67, indicating that the linking of clauses into larger units was rather common. This, according to the definition of complexity, can be seen as a good indication of complex language. Coordination was used rather infrequently as a way of linking main clauses (0.29 coordinate clauses per sentence on average). This result reflects the relatively small difference between sentences and T-units, as the main difference in segmenting these units is based on coordination. The mean dependent clause ratio was 0.50, i.e. there was one dependent clause for every other clause, indicating rather frequent use of dependent clauses, which can be considered a sign of complexity. In addition, the complexity of the use of dependent clauses can be seen in the mean number of non-finite clauses.

	Minimum	Maximum	Mean	Std. Deviation
W/U	10.00	20.00	14.27	3.04
W/AS	5.38	15.20	9.08	2.29
W/CSp	4.64	7.23	5.89	.68
C/U	1.57	3.60	2.45	.56
Coord/U	.14	1.33	.57	.34
DC/CSp	.00	.61	.34	.16
NonFSp	.00	4.00	1.67	1.28

**Table 3**: Complexity in the spoken production

In spoken production, as can be seen in Table 3, the mean lengths of the comparative units were 14.27 words per U-unit, 9.08 words per AS-units and 5.89 words per clauses. The lengths of U-units and AS-units varied more than the length of clauses. The U-unit complexity showed rather complex linking between clauses (mean of 2.45 in C/U), and coordination was used relatively often in the spoken data. Dependent clauses and non-finite clauses were rather infrequent in the spoken data (means of 0.34 and 1.67, respectively).

The complexity of the two modes of production was compared at group level by examining the measure pairs (see Table 1). As mentioned above, the spoken measures of C/AS and Coord/AS, and their written counterparts of C/T and Coord/T were calculated to compare the traditional comparative units with the U-unit in the measures of unit complexity and coordination. Therefore, the total number of measure pairs examined was nine.

| Measures compared	Z	Sig.
W/S - W/U	-1.590	.112
W/T - W/AS	-3.375	.001**
W/CWr - W/CSp	-1.502	.133
C/S - C/U	-.849	.396
C/T – C/AS	-3.332	.001**
Coord/S - Coord/U	-2.680	.007**
Coord/T – Coord/AS	-1.677	.094
DC/CWr - DC/CSp	-2.504	.012**
NonFWr - NonFSp	-3.149	.002**

Table 4: Spoken and written language measures compared

Table 4 shows the statistical significance of the Wilcoxon signed-ranks tests used to compare the measures. What follows is a separate discussion of the length, ratio and intra-clausal measures. Figure 1 shows the comparison of the length measures.
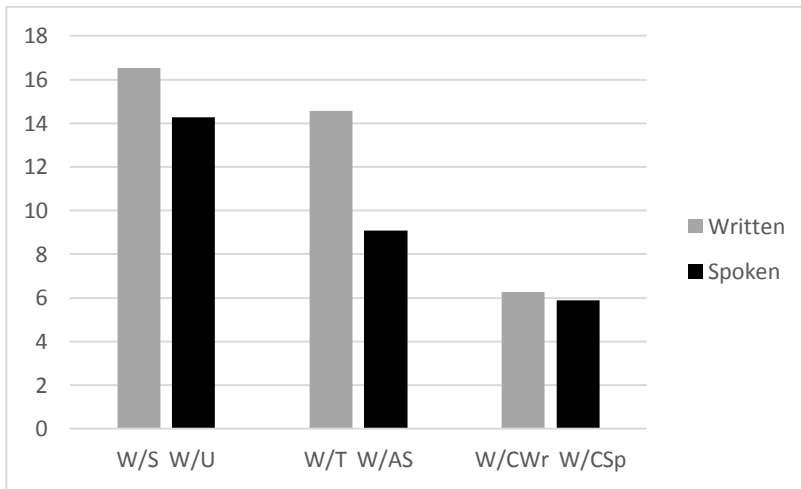
Figure 1: Means of length measures

As can be seen in Figure 1, all length measures implied greater complexity in the written mode. The difference was clearest between T-units and AS-units, the mean lengths being 14.57 and 9.08, respectively. The mean length of a single sentence was 16.53 words, whereas the mean length of the U-unit was 14.27. The difference between the clause lengths (mean length of 6.25 words per clause in the written and 5.89 words per clause in the spoken production) was notably smaller. The measure of mean length of T-unit vs. AS-unit was the only length measure that showed a statistically significant difference (p=0.001) between the modes of production (see Table 4).
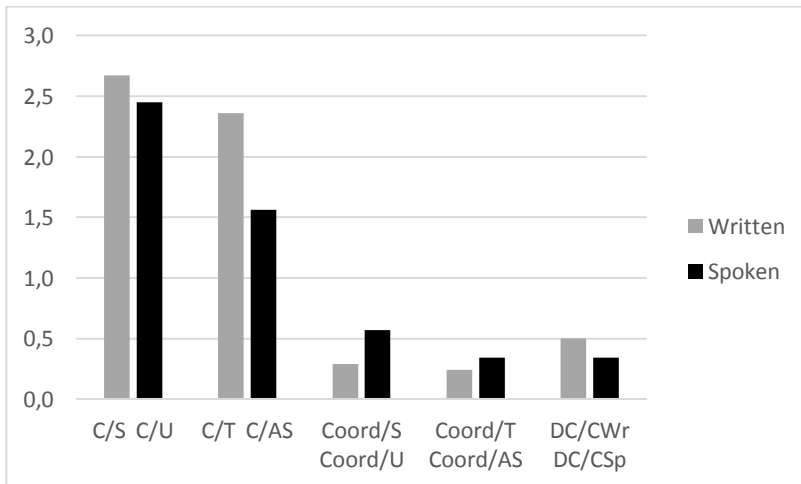
Figure 2: Means of ratio measures

Figure 2 presents the means of the ratio measures. Once more, the written samples were more complex than the spoken samples. The written complexity ratios were 2.67 (sentence complexity ratio, C/S) and 2.36 (T-unit complexity ratio, C/T), whereas the spoken complexity ratios were 2.45 (U-complexity ratio, C/U) and 1.56 (AS-complexity ratio, C/AS). As was the case with the length-based measures, using the U-unit in segmenting the data brought the spoken complexity closer to the written complexity.

The spoken production contained more coordination than the written production. In the measure pair of coordinate clauses per sentence (Coord/S, mean of 0.29) vs. coordinate clauses per U-unit (Coord/U, mean of 0.57), the difference was somewhat bigger than in the pair of coordinate clauses per T-unit (Coord/T, mean of 0.24) vs. coordinate clauses per AS-unit (Coord/AS, mean of 0.34). Once again, the difference between the AS-unit and T-unit was not as evident as when using the sentence and the U-unit categories as comparative units. The dependent clause ratios showed that the written productions (DC/CWr 0.50) contained a higher ratio of dependent clauses than the spoken productions (DC/CSp 0.34). In the complexity ratios, only the measure pair with comparative units of T-unit and AS-unit had a statistically significant (p=0.001) difference. Conversely, in the coordination ratios, only the coordination measure pair with sentence and U-unit as comparative units had a statistically significant (p=0.007) difference. In addition, the measure of dependent clause ratio had a statistically significant difference between the modes of production (p=0.012).

Lastly, complexity operationalized in terms of the intra-clausal measure was investigated. The measure of the number of non-finite dependent clauses indicated that the written mode of production was clearly more complex than the spoken mode. The mean number of non-finite dependent clauses in the written samples was 4.39, and 1.67 in the spoken samples. The difference in this measure was statistically significant (p=0.002).

The greatest differences in the complexity of the production modes were found on the measures of mean T-unit length vs. mean AS-unit length (p=0.001), T-unit complexity ratio vs. AS-unit complexity ratio (p=0.001), and the mean number of non-finite dependent clauses (p=0.002). Therefore, based on these measures, it may be concluded that written production was significantly more complex than spoken production. The greatest similarities were found between the mean clause lengths and the complexity ratios of the sentence and the U-unit. Thus, the clause length of written and spoken production did not differ much. The fact that the complexity ratios revealed a statistical difference between T-units and AS-units, but not between sentences and U-units, indicates that the choice of the segmentation unit affected the results greatly.

5. Discussion

In this section, we will compare the results of our study of the syntactic complexity of L2 speech and writing to earlier studies. First, we will focus on written language complexity. The mean lengths of the written units were 16.53 words per sentence, 14.57 words per T-unit and 6.25 words per clause. These lengths indicate a rather complex use of written language. In Storch and Wigglesworth (2007), the number of words per T-unit was 16.24, and the number of words per clause was 7.73. The numerical difference,

when compared to the results in the present study, is surprisingly small, as the subjects in Storch and Wigglesworth (2007) were university students who had to achieve a certain proficiency level in English. Larsen-Freeman (2006) reports a range of T-unit lengths between approximately 11 and 13 words in the written production by high intermediate learners. The assumed proficiency level of the subjects in the present study was somewhat lower, but the T-unit length indicates somewhat greater complexity. Wolfe-Quintero, Inagaki and Kim (1998: 23–25, 30–32) compared results from various L2 studies on sentence, T-unit and clause lengths. Despite the fact that in the studies the measures were used for measuring fluency, the variation found in the results offers some basis for comparison. In the studies compared, the mean number of words per sentence varied from 8.5 at the lower levels of learners to 23.59 at the advanced level, the mean number of words per T-unit from 6.0 to 23.0, and the mean number of words per clause from 5.20 to 10.83. The lengths of sentence and T-units in the present study can be placed approximately in the middle of these ranges, indicating that the level of complexity of our intermediate-level subjects was, in fact, intermediate. When it comes to the length of clauses, the mean length in the present study was somewhat shorter than the intermediate levels in the studies reported. For instance, according to Sharma (1980), low intermediate learners achieved on average 9.31 words per T-unit and 6.44 words per clause, and high intermediate learners 9.86 words per T-unit and 6.97 words per clause. The clause lengths seemed to be similar to the present study, but T-unit lengths showed a greater difference. Ishikawa (2007) also reported shorter T-unit lengths, ranging from 8.90 to 9.96. However, for instance Larsen-Freeman and Storm (1977) reported that average students had 12.92 words per T-unit, which is closer to the result of the present study. Although Ishikawa (2007) used a different definition of clause, the lengths of clauses, from 6.98 to 7.25, were similar to the mean lengths in the present study.

The ratios for the written production were 2.67 in sentence complexity ratio, 0.29 in coordination per sentence and 0.50 in dependent clause ratio. In Ishikawa (1995), the range of sentence complexity ratio was from 1.41 to 1.68. However, the subject group consisted of low-level learners. The T-unit ratio in the present study was 2.36, which indicated a somewhat less complex use of language than the sentence complexity ratio did. However, the T-unit ratio in the present study was slightly higher than the one in Storch and Wigglesworth (2007) (ratio of 2.10) and considerably higher than the one in Larsen-Freeman (2006) (a range from approximately 1.40 to 1.60 clauses per T-unit). The different definitions for the clause evidently affected the results, as could be seen especially in Ishikawa (2007), where the T-unit ratio varied from 1.27 to 1.37. The coordination ratio in the present study was 0.29. In Beaman (1984), coordinated sentences were the most common type of complex sentences both in written and spoken samples. The amount of written coordination was 0.38, which is somewhat more than in the present study. The frequency index for subordination found in Beaman (1984) was 54.2, which was a sign of frequent use of subordination. Ishikawa's (2007) subjects achieved dependent clause ratios of 0.20 and 0.26, which were considerably lower than the ratios in the present study. The same tendency can be found when comparing the results by Kuiken and Vedder (2007). Their dependent clause ratios were 0.36 for the first year students and 0.40 for the third year students. The main reasons for the variation in the results were the proficiency levels of the subjects and the varying, often inadequate definitions for the clause, among other factors. To reiterate, the inconsistency

in the definitions is an often-stated problem when discussing studies on syntactic complexity.

The mean number of non-finite dependent clauses was 4.39, indicating a rather frequent use of non-finite clauses. Robinson (2007: 209) noted that the increasing complexity in narratives led to a greater use of complex structures, such as infinitival structures. Therefore, it can be stated that the number of non-finite dependent clauses in the written data indicated complex language use. The average number of sentences, for instance, in the written essays was 8.4, and therefore, on average, there was a non-finite dependent clause for every other sentence.

With regard to the spoken language data, the mean lengths for spoken comparative units were 14.27 words per U-unit, 9.08 words per AS-unit and 5.89 words per clause. Tavakoli and Foster (2008) reported that the mean length of the AS-unit in their study varied from 7.72 to 10.86 words. The present study is thus in the mid-section of their range. Halleck (1995) used the T-unit as a length measure in spoken data, and reported a mean length of 8.02 words in the narrative part of the study at intermediate level, which is somewhat shorter than the mean length of AS-unit in the present study. Skehan and Foster (2005) reported a mean clause length of 5.39 to 5.50. Thus, the analysis of this measure in Skehan and Foster (2005) gave similar results with a similar subject group as in the present study.

The ratios in the spoken data were 2.45 in U-unit complexity ratio, 0.57 in coordination per U-unit and 0.34 in dependent clause ratio. The U-unit complexity ratio indicated that clausal linking was rather common in the spoken data. Skehan and Foster (2005) studied clausal linking in their data by examining the ratio of clauses per AS-unit. The AS-unit complexity ratio in their data was between 1.28 and 1.38, which is somewhat smaller than the ratio of 1.56 in AS-unit complexity ratio in the present study. The common linking at clausal level was in accordance with earlier studies accentuating that spoken language can be highly structured, especially when it comes to complexity in clausal linking (see Halliday 1979, 1989). However, the rather infrequent use of dependent clauses in relation to clauses in general indicated that embedding was not as common in the present data as some earlier studies reported (e.g. Beaman 1984: 78). Also Pietilä (1999), who used the analysis of subordination per T-unit in spoken data, found that the ratio of complex T-units was relatively small. Coordination is in general used relatively often in spoken language. Especially in coordination per U-unit in the data, coordination could be seen as a rather frequent way of combining main clauses into larger units. This could be interpreted as the use of less complex language, or simply as evidence for a characteristic of spoken language. Both arguments seem to have been supported by earlier research.

The mean number of non-finite dependent clauses per spoken sample was 1.67, which revealed an infrequent use of non-finite dependent clauses. Beaman (1984: 78) suggested that embedding is frequent both in written and spoken language and in writing there are more non-finite dependent clauses. The present study corroborates these arguments. However, Beaman further suggested that the overall number of dependent clauses was higher in spoken than written production, which claim is contradictory to our findings.

In conclusion, compared to earlier studies, the mean length of sentences and T-units in our study was somewhat longer than earlier studies suggest for intermediate level,

whereas the written clauses were somewhat shorter. The ratio-based measures revealed somewhat more complex language use than that reported in earlier studies. However, this finding must not be accepted without criticism as the varying definitions may have affected the overall results. The length of spoken units was significantly close to the results of earlier studies. The ratio-based measures suggested frequent clausal linking in the spoken data.

6. Conclusion

The purpose of this study was to examine the nature of syntactic complexity in spoken and written L2 production in our sample group and to explore how the choice of the segmentation unit might affect the results. The comparison of the modes showed that, in general, the written mode is more complex than the spoken mode of production. This finding is in accordance with numerous earlier studies on the differences between written and spoken complexity (cf. Tanskanen 2006: 74–80). The modes of production differed significantly on the measures of mean length of T-unit and AS-unit (p=0.001), T-unit and AS-unit complexity ratio (p=0.001), coordination per sentence and U-unit (p=0.007), dependent clause ratio (p=0.012) and the number of non-finite dependent clauses (p=0.002). The lack of statistical significance in the measures of mean length of the sentence and U-unit, mean length of clauses, sentence and U-unit complexity ratio and coordination per T-unit and AS-unit, indicated that the choice of segmentation unit strongly affected the results, and that spoken language complexity may not be as different from written language complexity as it had been claimed in several earlier studies.

    The results support partly the mainstream thinking that the differences of syntactic complexity in written and spoken language are evident, but also the Hallidayan idea that spoken language can be as complex as its written counterpart, although with different aspects of complexity in focus. It seems that even the complexity measured with the use of the same measures in written and spoken production can be closer to one another than earlier studies claimed, which would be against Halliday's notion that different modes are characterized by complexity of different nature. The reason for this seems to be, first of all, the choice of units used in segmenting the data. But as the marginal difference that the measure of mean length of clauses (the means of 6.25 words in written and 5.89 words in spoken production) revealed between the modes indicates, there are rather significant similarities in the complexities of written and spoken production. This is especially intriguing as embedding was often highlighted by researchers who agreed with Halliday (1979, 1989) that spoken production was more complex, however, the analysis of the mean length of clauses in the present study showed similarities in clausal complexity across the two modes.

    The individual examination of the samples revealed fairly great inter-individual and intra-individual variation in the present study. Despite the general tendency that writing would be more complex than speech, with certain individual subjects, the spoken production was more complex than their written production. Thus, there should be further studies that would focus on individual variation more attentively. On the basis of the present results, it can be stated that the fundamental differences in the modes of

production result in certain differences in syntactic complexity, but, on the other hand, the placement on the orality-literacy continuum has some effect on the syntactic structures used. In the case of L2 users, the proficiency level in writing and speech has an additional bearing on syntactic complexity. Although the registers of the data types in the present study were rather similar, the style of the task (narrative in the spoken and a more or less argumentative style in the written task) may have influenced the structures that the subjects used.

As the second aim of the present study, a new unit for segmenting spoken data was piloted. This unit was not solely bound to pausing and intonation that, in fact, would often indicate fluency rather than complexity, but it would provide an opportunity to include more syntactic structures that are meant to form a more complex unit. The U-unit seems to bring spoken language complexity closer to written language complexity, and appears to be able to do that more efficiently than the AS-unit does. Therefore, the extent to which spoken and written language complexity differ from each other seems to depend not so much on the measure used, as it does on the units used in segmenting the data. The U-unit might reveal the learner's intended idea better than the traditionally used spoken language units. The fact that learners have frequent and long pauses in their spoken production should not affect the analysis of the syntactic complexity of their spoken language. In addition, as the U-unit allows coordination more efficiently than some other units, it is closer to the concept of a sentence as used with reference to written language.

After this piloting phase, this experimental unit should be examined further. A larger set of data must be analysed before the inter-rater reliability could be efficiently tested. In further studies, a comprehensive qualitative analysis on inter- and intra-individual variability is also needed. However, as one of the aims of the present study was to experiment with the units and measures, this study hopefully functioned as a starting point for discussion and further research. The potential of measuring learner's development is another aspect related to our new unit. This is especially important as the proficiency levels of an individual learner might significantly differ in speaking and in writing. The present study also indicates the potential of future research into using different task types and their possible effects on the results, and, finally, into wider consistent experiments and studies in the two modes of production, speech and writing.

On the basis of this study, it can be concluded that the differences between written and spoken complexity seem to be partly a result of the nature of the mode and partly a result of the choice of theoretical units used in segmenting the samples. Many units used in segmenting spoken data seem to focus too much on the aspect of fluency and cannot be fully adapted to the study of complexity. Therefore, research is needed, both on the units and on their application in the measures, to study more reliably the differences between the modes of production.

References

Bardovi-Harlig K. 1992. A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly* 26 (2): 390–395. DOI: 10.2307/3587016
Baron N. 2000. *Alphabet to Email*. London: Routledge.

Beaman K. 1984. Coordination and subordination revisited: syntactic complexity in spoken and written narrative discourse. In D. Tannen (ed.) *Spoken and Written Language: exploring orality and literacy*. Norwood, N.J.: Ablex: 45–80.

Bergman P. and T. Abrahamsson. 2004. Bedömning av språkfärdigheten hos andraspråkselever. In K. Hyltenstam and I. Lindberg (eds) *Svenska som andraspråk – i forskning, undervisning och samhälle*. Lund: Studentlitteratur: 597–626.

Biber D., S. Conrad and G. Leech. 2002. *Longman Student Grammar of Spoken and Written English*. Harlow: Pearson Education.

Bourdin B. and M. Fayol. 1994. Is written language production more difficult than oral language production? A working memory approach. *International Journal of Psychology* 29 (5): 591–620. DOI: 10.1080/00207599408248175

Brown G. and G. Yule. 1983. *Teaching the Spoken Language*. Cambridge: Cambridge University Press.

Bulté B. and A. Housen. 2012. Defining and operationalising L2 complexity. In A. Housen, F. Kuiken and I. Vedder (eds) *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*. Amsterdam: Benjamins: 21–46.

Chafe W. 1982. Integration and involvement in speaking, writing and oral literature. In D. Tannen (ed.) *Spoken and Written Language: exploring orality and literacy*. Norwood, N.J.: Ablex: 35–53.

Cleland A. A. and M. J. Pickering. 2006. Do writing and speaking employ the same syntactic representations? *Journal of Memory and Language* 54: 185–198. DOI:10.1016/j.jml.2005.10.003

Ellis R. 2008. *The Study of Second Language Acquisition*. 2nd edition. Oxford: Oxford University Press.

Ellis R. and G. Barkhuizen. 2005. *Analysing Learner Language*. Oxford: Oxford University Press.

Foster P., A. Tonkyn and G. Wigglesworth. 2000. Measuring spoken language: a unit for all reasons. *Applied Linguistics* 21 (3): 354–375. DOI: 10.1093/applin/21.3.354

Gaies S. 1980. T-Unit analysis in second language research: Applications, problems and limitations. *TESOL Quarterly* 14 (1): 53–60. DOI: 10.2307/3586808

Gilabert R. 2007. The simultaneous manipulation of task complexity along planning time and (+/- here-and-now): Effects on oral production. In M. P. Garcia Mayo (ed.) *Investigating Tasks in Formal Language Learning*. Clevedon: Multilingual Matters: 44–68.

Halleck G. 1995. Assessing oral proficiency: A comparison of holistic and objective measures. *The Modern Language Journal* 79 (2): 223–234. DOI: 10.1111/j.1540-4781.1995.tb05434.x

Halliday M. A. K. 1979. Differences between spoken and written language: Some implications for literacy teaching. In G. Page, J. Elkins and B. O'Connor (eds) *Communication through reading: Proceedings of the Fourth Australian Reading Conference Vol. 2*. Adelaide: Australian Reading Association: 37–52.

Halliday M. A. K. 1989. *Spoken and Written Language*. Oxford: Oxford University Press.

Holger D. 2004. *The Acquisition of Complex Sentences*. New York: Cambridge University Press.

Housen A. and F. Kuiken. 2009. Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics* 30 (4): 461–473. DOI: 10.1093/applin/amp048

Housen A., F. Kuiken and I. Vedder. 2012a. Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken and I. Vedder (eds) *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*. Amsterdam: Benjamins: 1-20.

Housen A., F. Kuiken and I. Vedder (eds). 2012b. *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*. Amsterdam: Benjamins.

Hunt K. 1965. *Grammatical Structures Written at Three Grade Levels*. Champaign, IL: National Council of Teachers of English.

Ishikawa S. 1995. Objective measurement of low-proficiency EFL narrative writing. *Journal of Second Language Writing* 4 (1): 51–69. DOI: 10.1016/1060-3743(95)90023-3

Ishikawa T. 2007. The effect of manipulating task complexity along the [+/- Here-and-Now] dimension on L2 written narrative discourse. In M. P. García Mayo (ed.) *Investigating Tasks in Formal Language Learning*. Clevedon: Multilingual Matters: 136–156.

Iwashita N. 2006. Syntactic complexity measures and their relation to oral proficiency in Japanese as a foreign language. *Language Assessment Quarterly* 2 (4): 151–170. DOI: 10.1207/s15434311laq0302_4

Kuiken F. and I. Vedder. 2007. Cognitive task complexity and linguistic performance in French L2 writing. In M. P. García Mayo (ed.) *Investigating Tasks in Formal Language Learning*. Clevedon: Multilingual Matters: 117–135.

Larsen-Freeman D. 2006. The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics* 27 (4): 590–619. DOI: 10.1093/applin/aml029

Larsen-Freeman D. and V. Storm. 1977. The construction of a second language acquisition index of development. *Language Learning* 27 (1): 123–134. DOI: 10.1111/j.1467-1770.1977.tb00296.x

Larson-Hall J. 2010. *A Guide to Doing Statistics in Second Language Research Using SPSS*. New York: Routledge.

Leech G., M. Deuchar and R. Hoogenraad. 1982. *English Grammar for Today*. Basingstoke: Macmillan.

Mellow D. J. 2006. The emergence of second language syntax: A case study of the acquisition of relative clauses. *Applied Linguistics* 27: 645–670. DOI: 10.1093/applin/aml031

Norris J. M. and L. Ortega. 2009. Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics* 30 (4): 555–578. DOI: 10.1093/applin/amp044

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. DOI: 10.1093/applin/24.4.492

Ortega L. 2009. *Understanding Second Language Acquisition*. London: Hodder Education.

Pallotti G. 2009. CAF: Defining, refining and differentiating constructs. *Applied Linguistics* 30 (4): 590–601. DOI: 10.1093/applin/amp045

Pica T., L. Halliday, N. Lewis and L. Morgenthaler. 1989. Comprehensible outputs as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition* 11 (1): 63–90. DOI: 10.1017/S027226310000783X

Pietilä P. 1999. *L2 Speech: Oral Proficiency of Students of English at University Level.* Anglicana Turkuensia No 19. Turku: University of Turku.

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*. Mahwah, N.J.: Lawrence Erlbaum: 91–115.

Robinson P. 2001. Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics* 22 (1): 27–57. DOI: 10.1093/applin/22.1.27

Robinson P. 2007. Task complexity, theory of mind, and intentional reasoning. Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics* 45 (3): 193–213. DOI: 10.1515/iral.2007.009

Scarborough H. S. 1990. Index of productive syntax. *Applied Psycholinguistics* 11: 1–22. DOI: 10.1017/S0142716400008262

Sharma A. 1980. Syntactic maturity: Assessing writing proficiency in a second language. *Language Learning 37*: 469–481.

Silva M. L., V. Sánchez Abchi and A. Borzone. 2010. Subordinated clauses usage and assessment of syntactic maturity: A comparison of oral and written retellings in beginning writers. *Journal of Writing Research* 2 (1): 47–64. DOI: 10.17239/jowr-2010.02.01.2

Skehan P. 1998. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Skehan P. and P. Foster. 2005. Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis (ed.) *Planning and Task Performance in a Second Language*. Amsterdam: John Benjamins: 193–216.

Storch N. and G. Wigglesworth. 2007. Writing tasks: The effects of collaboration. In M. P. García Mayo (ed.) *Investigating Tasks on Formal Language Learning*. Clevedon: Multilingual Matters: 157–177.

Szmrecsányi B. M. 2004. On operationalizing syntactic complexity. *JADT: 7e Journées internationales d'Analyse statistique des Données Textuelles*. Retrieved October 2014 from www.benszm.net/omnibuslit/Szmrecsanyi2004.pdf.

Tanskanen S.-K. 2006. *Collaborating Towards Coherence: Lexical Cohesion in English Discourse*. Philadelphia: John Benjamins.

Tavakoli P. and P. Foster. 2008. Task design and second language performance: The effect of narrative type on learner output. *Language Learning* 5 (2) 439–473. DOI: 10.1111/j.1467-9922.2008.00446.x

Towell R. 2007. Complexity, accuracy and fluency in second language acquisition research. In S. Van Daele, A. Housen, F. Kuiken, M. Pierrard and I. Vedder (eds) *Complexity, Accuracy and Fluency in Second Language Use, Learning and Teaching*. Brussels: Contactforum: 260–272.

Tonkyn A. 2012. Measuring and perceiving changes in oral complexity, accuracy and fluency. In A. Housen, F. Kuiken and I. Vedder (eds) *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*. Amsterdam: Benjamins: 221–245.

Vyatkina N. 2013. Specific syntactic complexity: Developmental profiling of individuals based on an annotated learner corpus. *The Modern Language Journal* 97 (S1): 11–30. DOI: 10.1111/j.1540-4781.2012.01421.x

Wolfe-Quintero K., S. Inagaki and H.-Y. Kim. 1998. *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. Honolulu: University of Hawaii Press.

Zhang B. 2013. An analysis of spoken language and written language and how they affect English language learning and teaching. *Journal of Language Teaching and Research* 4 (4): 834–838. DOI:10.4304/jltr.4.4.834-838