

Wijers, Martje. 2018. The Role of Variation in L2 Syntactic Complexity: A Case Study on Subordinate Clauses in Swedish as a Foreign Language. *Nordic Journal of Linguistics* 41(1), 75–116.

The Role of Variation in L2 Syntactic Complexity: A Case Study on Subordinate Clauses in Swedish as a Foreign Language

Martje Wijers

This paper critically examines the exclusive use of the relative frequency of subordinate clauses as a measure of syntactic complexity in Second Language Acquisition (SLA) L2 acquisition and as an indicator of Second Language (L2) development. Following Lambert & Kormos (2014), it is argued that it is important to also take into account: (i) different subordinate clause types, (ii) item-based frequencies, and (iii) text genre. Longitudinal written data was collected among 21 Dutch-speaking foreign language learners of Swedish. Based on these data, the study shows that the subordination ratio (a common measure for syntactic complexity) alone was found to be insufficient as an indicator of syntactic complexity and L2 development, as hardly any significant differences were observed in the subordination ratios of the learners and native speakers. The study shows that other aspects are also relevant in determining the learners' level of syntactic complexity as well as their L2 development, such as the internal structure and context of subordinate clauses, subordinate clause types, and especially the type-token ratio of subordinators used. The results showed a significant negative correlation between the subordination ratio of texts and the type-token ratio of subordinators. The aim of the study is to acknowledge the variation in the learners' use of subordinate clause, in addition to the general subordination ratio, in order to arrive at a more nuanced view of syntactic complexity in second language acquisition.

Keywords: subordination, Swedish, foreign language acquisition, development, syntactic complexity

University of Ghent, Department of Linguistics, Institute for Scandinavian studies, Ghent, 9000, Belgium. Martje.Wijers@Ugent.be

1. INTRODUCTION

This paper examines the role of variation in syntactic complexity in Second Language Acquisition (SLA). Special attention will be paid to the role of variation in the development of Second Language (L2) subordinate clause usage and its implications for theory on L2 complexity. The study is longitudinal and is concerned with written production of 21 beginning L2 learners of Swedish as a foreign language.

In various fields of linguistics, clausal embedding is linked to syntactic complexity. In SLA specifically, the SUBORDINATION RATIO (i.e. number of

subordinate clauses divided by the total number of clauses or T-units) is one of the most frequently used measurements for syntactic complexity in SLA (Norris & Ortega 2009). In addition, researchers using the subordination ratio have assumed that language learners master relations between clauses and structures that go across clause boundaries only at later stages of development, and that, consequently, the use of subordinate clauses must be complex (Skehan & Foster 1999; Ellis 2009). Also, for native speakers, when more subordinate clauses are being used, language is generally considered to be more complex or harder to process. (Lord 2002).

However, a number of researchers have recently questioned the validity of the subordination ratio as a measure of syntactic complexity and/or an indicator of development (Baten & Håkansson 2015; Bulté & Housen 2012; Norris & Ortega 2009; Pallotti 2009, 2014; Lambert & Kormos 2014). Their criticism is based on the definition, operationalization, and the level of analysis of the construct. The present study discusses and empirically investigates these criticisms. It will be argued that the factor variation has often been overlooked in studies on syntactic complexity, possibly because it is harder to operationalize and qualify compared to frequency measures. Besides the frequency of subordinate clauses or a deeper level of syntactic embedding, a text may be complex in that it has a high level of syntactic diversity. Syntactic variation is, therefore, important to investigate in order to fully understand syntactic complexity in L2 language production.

Lambert & Kormos (2014) question the validity of complexity as it has been used in research on L2 development and suggest other measures that might be more representative for SLA. Their three main points of critique are the following:

1. The current measures do not differ between different types of subordination. Indeed, native speakers use different subordinate clause types according to age, at least for English (see Nippold et al. 2005). Analogously, L2 learners of English also use different subordinate clause types according to proficiency level (see Norris & Ortega 2009).
2. There is no control for frequent item-based use of superordinate verbs. If some superordinate verbs are used very frequently or exclusively, this is not at all reflected in the overall subordination ratio.
3. The current measures do not take into account differences in task types or text genre and mode of production (spoken or written production).

The first two points can be connected to the level of analysis or the lack of attention to syntactic variety, or to the interaction between syntax and semantics in research on syntactic complexity. If systemic grammatical complexity is defined as “elaboration, size, range, variation, ‘breadth’ of L2 grammar” (Bulté & Housen 2012, 27), which will be used as the definition of complexity in this paper, then the measures (such as mean length of clause and frequency of subordinate clauses) that are currently used do not measure all of the above-mentioned types of complexity.

In research on lexical complexity, the factor variation is already integrated (Bulté & Housen 2012), but for syntactic complexity, this is not the case. Furthermore, in all of the mentioned studies, researchers make a strict distinction between syntactic and lexical/semantic complexity, although it is highly disputable whether it is even possible to disentangle these two concepts.

In addition to the subordination ratio, the present study investigates the structural syntactic complexity of texts written by L2 learners by means of the three points mentioned by Lambert & Kormos (2014): (i) the type of subordinate clause used, (ii) item-based (lexical) frequencies, and (iii) text genre. These factors will be put to the test on the basis of longitudinal written L2 production data to investigate whether they are indeed representative of complexity in second language development. The focus will be on different types of subordinate clauses and item-based use of subordinators. Different text types will also be discussed, but only in passing, as it is hard to distinguish between time of writing (point in development) and text type in the current study design. The learner data will also be compared to written native speaker baseline data, consisting of multiple text genres, because baseline data is noticeably lacking in most L2 complexity studies.

The paper is organized as follows. Section 2 starts with background information on linguistic complexity in general and introduces and critically examines the notion of subordination ratio in SLA. The structure of subordinate clauses in Swedish and previous studies on subordination in L2 Swedish are briefly presented in Section 2.2. The methodology is presented in Section 3. The results are presented in Section 4, with more thorough discussion of these in Section 5. Section 6 concludes the article.

2. SUBORDINATION AND SYNTACTIC COMPLEXITY

Complexity is a much-debated concept in many fields of linguistics, such as theoretical linguistics and typology, language contact, historical linguistics, and language acquisition. Several book volumes on linguistic or syntactic complexity have been published in recent years (see, e.g., Miestamo, Sinnemäki & Karlsson 2008; Sampson, Gil & Trudgill 2009; Newmeyer & Preston 2014; Trotske & Bayer 2015). In the following section, the main issues in the study of subordination and complexity and the link between these concepts will be discussed.

2.1. *Main issues in the study of subordination and complexity*

Two main issues can be discerned in the large body of research on the topic of linguistic complexity. The first issue concerns what is often called ‘the trade-off problem’, which refers to the long-standing and only recently challenged

hypothesis that all languages are equally complex and that complexity in one domain is compensated by simplicity in others (Sampson et al. 2009:1). The second issue that caused a renewed interest in linguistic complexity was the major hypothesis of the minimalist program that recursion is the only “uniquely human component of the faculty of language” (Hauser et al. 2002:1569). Many theoretical/typological studies focus on clausal embedding or subordination and have implicitly or explicitly linked clausal embedding to complexity. The common challenge in these studies is to find ways to objectively measure and define syntactic complexity.

As is often the case with basic concepts within linguistics, there is no universally acknowledged definition of complexity. Miestamo (2008) argues that two notions of linguistic complexity should be distinguished: (1) absolute complexity—referring to an objective property of the system, and (2) relative complexity—complexity as cost/difficulty to language users. Culicover (2014:148) also makes a distinction between formal and processing complexity but links higher processing complexity to competing constructions. He hypothesizes that lower processing cost is usually preferred over similar constructions that have a higher processing complexity by speakers. The less complex construction should, therefore, be more frequent than constructions that require more processing cost. Kusters (2003) assigns an important role to SLA in measuring grammatical complexity. He advocates that relative complexity should be applied to measure differences in linguistic complexity between languages. According to Kusters, it is more difficult for L2 learners to acquire complex linguistic structures.

However, the definition of complexity as it pertains to SLA is often unclear. Pallotti (2014) argues that different uses of “complexity” need to be disentangled. According to Pallotti (2014), complexity is used in at least three different ways:

1. inherent complexity
2. difficulty or cognitive complexity
3. processing difficulty (developmental order)

The first definition refers to formal properties of the linguistic system irrespective of the language learner and has also been described as purely structural objective difficulty (Pallotti 2009). The second definition has to do with issues of processing cost for language users and depends on how demanding linguistic items are to process. This definition would be what Bulté & Housen (2012) refer to as cognitive complexity. The third definition is directly connected to language acquisition—it builds on the idea that complex grammatical structures are acquired late, and, as a consequence, that more advanced learners use more demanding or difficult linguistic features. This is, for instance, the definition of complexity used by Skehan & Foster (1999) and Ellis (2009). They link complexity to the use of more advanced language, which in

turn may lead to less control, more risk-taking, and ultimately to development in the interlanguage.

Taking into account the different uses of complexity, Pallotti (2014) argues that inherent linguistic complexity in particular should be kept separate from processing difficulty or developmental order, because language does not necessarily become more complex over time. Along similar lines, Bulté & Housen (2012) remark with regard to the second and third definition that “*difficulty* is a different construct from *structural* complexity, and the correspondence between the two constructs still has to be demonstrated rather than *a priori* assumed.” Others, like Menn et al. (2014), on the other hand, argue that (structural) complexity measures should successfully predict neurological activity. It is, therefore, crucial to critically investigate the currently used measures of syntactic complexity in SLA and other fields in linguistics to get closer to answers to the more general linguistic debates on linguistic complexity. In this study, complexity refers to structural complexity (definition 1) in case it is not indicated otherwise. Lambert & Kormos (2014) suggest that complexity can be linked to learner development, but not with the measures that are currently used most frequently.

Naturally, the argument to keep different notions of complexity separate from each other may have consequences in terms of operationalization. Different notions of complexity require different measures. Subordination is also a prominent feature in Complexity, Accuracy, and Fluency (CAF) research in SLA (Housen & Kuiken 2009; Bulté & Housen 2012). In a meta-study on complexity, Norris & Ortega (2009) found that the subordination ratio was used in all the examined complexity studies as a measure of syntactic complexity. In the studies that they investigated, the subordination ratio was used as the dependent/outcome variable, meaning that the possible effects of factors like task type, instruction, or task planning on the measures of syntactic complexity were investigated. The subordination ratio has also been used as an independent/predictor variable in second language development, even though the relation between the subordination ratio and L2 development is not clear (Housen & Kuiken 2009). If there is no correspondence between the two variables, the results of these studies would be invalid. Remarkably, in none of the mentioned studies or the studies investigated by Norris & Ortega (2009) were subordination ratios in language use of learners compared to native speakers’ use of subordination. There is, in other words, no reference point or baseline data.

Norris & Ortega (2009:566) express the view that “exclusive reliance on subordination is worrisome” (see also Bulté & Housen 2012). As the frequency of subordinate clauses is only one aspect of linguistic complexity, complexity should be investigated using a broader range of factors, such as variety, sophistication, and acquisitional timing of syntactic forms (Norris & Ortega 2009:561–562). Baten & Håkansson (2015) argue that the subordination ratio is not the right measure to indicate structural syntactic complexity and developmental timing in SLA. Instead,

it is the internal structure of the subordinate clause, rather than its frequency ratio, that is relevant when analyzing language development. Their claim builds on a study of word order patterns in subordinate clauses, more precisely the V-End rule in German and the ADV + Verb rule in Swedish in second language learners' production.

Researchers in the CAF field have yet to investigate the role of 'variety' and 'sophistication' in complexity more systematically. Only a few researchers (e.g. Ellis 2005; Sangarun 2005) have tried to include these factors in their studies on complexity, counting selected forms which are considered to be more sophisticated (Norris & Ortega 2009:562). This is, of course, a harder criterion to examine objectively, which is probably why it has not been done very often.

2.2. Subordinate clauses in Swedish

This study follows the classification of subordinate clauses by Teleman et al. (1999). In Swedish, a prototypical subordinate clause is traditionally defined as any subordinate clause introduced by a subordinating conjunction and containing a finite verb (Teleman et al. 1999:4, 462).¹

Three main types of subordinate clauses are usually distinguished, based on their grammatical function in the sentence (Teleman et al. 1999:4, 462–472): noun clauses, adjectival clauses, and adverbial clauses.

NOUN CLAUSES function as arguments to the verb and can thus be subjects, direct objects, predicates, appositives, or objects of the preposition. Noun clauses most prominently include *att*-clauses (comparable to *that*-clauses in English) and interrogative clauses. An example of an *att*-clause is given in (1).

- (1) Det var snällt av dig [att du kom.]
 it be.PST kind of you that you come.PST
 'It was kind of you that you came.' (noun cl.)
 (Teleman et al. 1999:469)

ADJECTIVAL CLAUSES are mainly relative clauses, such as the subordinate clause in example (2), but the more general term 'adjectival clause' is used here to emphasize that the emphasis in the analysis is on the function of subordinate clauses, rather than their form. For instance, the group of adjectival clauses also includes appositive noun clauses, *att*-clauses that are used in apposition to a noun or a pronoun as in the standard analysis taken in Teleman et al. (1999:471), illustrated in example (3).

- (2) Vi tog en bok [som du inte har last.]
 We take.PST a book that you not have.PRS read.SUP
 'We took a book that you haven't read.'
 (Teleman et al. 1999:471) (adj. cl.)

- (3) Det faktum [att Olle inte kan komma] är knappast en
 the fact that Olle not can.PRS come.INF be.PRS hardly a
 överraskning.
 surprise
 ‘The fact that Olle isn’t coming is hardly a surprise.’
 (Teleman et al. 1999:471)

ADVERBIAL CLAUSES can, for instance, express time, condition, purpose, reason, place, or manner. Whereas noun clauses have a direct relation to the main verb in the main clause, adverbial clauses mainly anchor the proposition in the main clause in time or space. An example of an adverbial clause is given in (4) below.

- (4) Han glömde allt [medan han läste boken.]
 He forget.PST everything while he read.PST book.DEF
 ‘He forgot everything while he was reading the book.’ (adv. cl.)

An important difference between subordinate clauses and main clauses in Swedish is word order. Swedish has V2 in main clauses, whereas the finite verb follows both the subject and sentential adverbs in subordinate clauses, as is illustrated in the examples in (5). In example (5a), the adverb *aldrig* ‘never’ follows the finite verb *talar*, whereas it typically precedes the finite verb in the subordinate clause in example (5b).²

- (5) a. Det har jag **aldrig** sett.
 That have.PRS I never see.SUP
 ‘I have never seen that.’
 b. Jag ska visa dig något som du **aldrig** har
 I will show.INF you something that you never have.PRS
 sett.
 see.SUP
 ‘I will show you something you’ve never seen.’

However, it should be pointed out that there is an important structural difference between subordinate clause types in Swedish. Some subordinate clauses, namely ‘asserted’ *att*-clauses, are different from prototypical subordinate clauses. These clauses frequently have embedded V2 and have been said to be syntactically subordinate clauses, but semantically main clauses (for an extensive literature overview on embedded V2 in Swedish, see Chapter 4 in Petersson 2014).

2.3. Subordination in L2 Swedish

Quite a number of studies on L2 Swedish have investigated subordinate clause word order in learners of L2 Swedish (Baten & Håkansson 2015; Håkansson & Nettelbladt 1993; Håkansson & Norrby 2010; Hammarberg & Viberg 1977; Hyltenstam 1977),

which is typically acquired late (Glahn et al. 2001). The studies do not reveal, however, whether non-native speakers have difficulties acquiring the subordinate clause word order rule or whether they do not understand the difference between main and subordinate clauses in the target language (in other words, whether it is a matter of annotating or exchanging grammatical information).

Baten & Håkansson (2015:532) argue that the acquisition of subordination is not difficult per se. They state that in a language like English, for example, there is no difference between most main and subordinate clauses with regard to internal structure. Thus, they conclude that “[i]n L2 English, learners merely need to acquire (the meaning of) subordinators, to which they can add a new clause” (Baten & Håkansson 2015:32).

This view seems somewhat simplistic: even though structural differences between English main and subordinate clauses are not always visible, they do exist (Hooper & Thompson 1973). Moreover, even if it is easier for the learner of English to make use of the internal grammatical structure of subordinate clauses, we cannot be certain whether the learner analyzes these clauses as subordinate. Even in the absence of structural word order differences, the learner still has to acquire how different clause types can be used. It may be true that the structural distinction between clause types is not as relevant or more semantically based for learners of English compared to learners of Swedish. L2 learners of English, like learners of Swedish, might not realize when they use subordinate clauses. In English, this does not directly result in ungrammatical clauses, whereas this can be the case in Swedish, if a sentence adverbial is produced and placed after the finite verb in a non-asserted subordinate clause. Learning to understand in what cases the subordinate clause word order should be applied might in fact be more challenging than the subordinate word order pattern itself.

As word order is different in main and subordinate clauses in Swedish, checking word order is a way of distinguishing between main and subordinate clauses for native speakers of Swedish. L2 learners of Swedish, however, cannot rely on word order when they produce their own sentences if they have not acquired the word order yet. To acquire the subordinate clause word order, L2 learners have to signal the occurrence of subordinate clause structures. This does not mean that they have to be able to (explicitly) label them, but that they have to (implicitly) categorize clauses into main and subordinate clause constructions. This signaling can be helped by paying attention to subordinators or certain complement-taking predicates, the number of finite verbs in a sentence, and the semantic relation between the verbs. These aspects are, therefore, closely linked to the position of sentence adverbials in L2 language production. By investigating the use of subordinators and finiteness, I attempt to show that the word order rule per se is not the only problem but rather the distinction between main and subordinate clauses in general.

Text	Time of writing (after time of onset)	Text type	Topic
1	4 weeks	description	present yourself
2	8 weeks	narrative	my travels/holidays
3	12 weeks	description	description of a person
4	6 months	exposition	why behavior is also part of language
5	12 months	exposition	advantages and disadvantages of a topic of choice
6	13.5 months	narrative	a childhood memory
7	15 months	response	response to a popular science article

Table 1. Topics of writing assignments.

3. METHODOLOGY

3.1. *Participants*

In this longitudinal study, 21 Dutch-speaking learners of Swedish as a foreign language participated, starting from the absolute beginner level. The group of learners consisted of students at Ghent University in Belgium. All learners had Dutch as their L1 and English as their strongest L2. According to self-judgement in survey questions, English was the L2 the participants used most frequently besides their L1. All participants recorded that they were proficient in English. Many participants also had knowledge of other languages (most importantly French). The learners' age of onset for Swedish was between 17 and 22 years. Students with any prior knowledge of Swedish were excluded from the study. During the first semester, the participants received 6 hours of classroom instruction on Swedish per week for 13 weeks. During the second semester, they only had 1.5 hours of language instruction in a language proficiency course per week, but they received 4.5 hours of Swedish input in two other courses that used Swedish as language of instruction. Also during the third term, students received 6 hours of Swedish instruction per week for 13 weeks, of which 3 hours per week was in language proficiency.

3.2. *Data elicitation*

The data consists of free writing assignments that were obtained at seven different points in time, over a time span of 1 year and 3 months. Only the topic of writing was given, without further instruction. [Table 1](#) provides an overview of these topics together with the timing of the writing assignment, counted from the start of the learning process.

The topics resulted in four distinctive text types. Texts 1 and 3 were descriptive texts in which the learners described either themselves or another person. Texts 2 and 6 were both narrative texts, namely a recount of an event that happened during a holiday and a childhood memory. Texts 4 and 5 were expositions, and text 7 was a response to an article. Each text contained approximately 300 words, and the learners had a dictionary at their disposal while writing. The texts were written on a computer at home. They were not written specifically for this study but were part of course assignments, so the learners were not aware of the focus of the current study. This also means that the text types were not chosen for the purpose of the present study but rather reflect a representative order of writing assignments in a foreign language classroom, with descriptive and narrative texts preceding the more demanding text types (expositions and responses). The last assignment may have been less representative, because it was a response to another text. This means that the learners often used verbs of indirect speech, which automatically involves clausal subordination. This type of text can thus be considered an elicitation technique for subordinate clauses. Text 7 was also the only assignment for which the learners received specific information on the text genre required, namely response. They did not receive any example texts before writing.

The native speakers' texts used in this study were collected to compare L2 learners to Swedish native speakers. Texts from five different genres were studied:

1. 10 narrative texts written by 17-year-old high school students (see Johansson 2009)
2. 10 expositions written by 17-year-old high school students (see Johansson 2009)
3. 10 responses written by 17-year-old high school students (see Holmberg & Wirdeñäs 2010)
4. 10 randomly picked blog texts (recounts) written by young adults (aged between 18 and 30 years)
5. 10 randomly picked newspaper texts (recounts) written by professional journalists from various newspapers

The first two groups of texts were collected by Victoria Johansson. The selected students had a monolingual Swedish background and were likely to proceed in higher education. For a detailed description of the data collection, see Johansson (2009).³ The responses were collected by Per Holmberg as part of the material collected in the project Text- och kunskapsutveckling i skolan (TOKIS) 2007–2010, financed by Vetenskapsrådet (see Holmberg & Wirdeñäs 2010). The participants all had a monolingual Swedish background.⁴ The various text types were chosen to be as comparable to the genres of the L2 learners' texts as possible. The average age of the native speakers was only slightly younger than the L2 learners when they started learning Swedish (average age: 18). The blog texts were selected from the top 100 most frequently read blogs in Sweden. The newspaper texts were selected from the

front pages of the following major Swedish newspapers: *Svenska Dagbladet*, *Dagens Nyheter*, *Expressen*, and *Aftonbladet*. All texts were recounts (reporting objectively on current news events) with roughly the same number of words as the learner texts.

3.3. Data analysis

Although a prototypical subordinate clause in Swedish is introduced by a subordinating conjunction and contains a finite verb, these are hardly criteria that can be used to identify subordinate clauses in L2 Swedish. As we have seen earlier, the subordinator can be left out in some cases. Furthermore, L2 learners frequently use non-finite verb forms where a finite verb would be the target form, and the other way around: finite verb forms where a non-finite form would be the target form. Both criteria are thus unreliable. The same can be said of word order, which is different from the main clause in some subordinate clauses, but not in all. If the word order in the subordinate clause produced by an L2 learner is not target-like, this does not mean that the clause is not embedded. In other words, it is not easy to define the criteria for subordinate clauses in L2 writing. Therefore, word order is not used as a criterion for defining subordinate clauses in L2 writing. In the L2 writing in this study, a clause is broadly defined as a linguistic unit including at least a subject and a verb, regardless of whether the verb has a finite or a non-finite form. The clause is defined as subordinate if:

- (i) the linked clauses are not coordinated by means of a coordinating conjunction or punctuation;
- (ii) the clause functions as a constituent in another clause, or in case of adjectival clauses, modifies a noun phrase.

The data was analyzed in four steps. First, the number and type of main and subordinate clauses were analyzed for each single sentence in the learner texts. Second, the subordination ratio was computed by dividing the number of subordinate clauses by the total number of clauses. The subordinate clauses in all texts were analyzed and classified according to the three aforementioned types: noun clauses, adjectival clauses, and adverbial clauses. All frequencies were normalized for word count so that frequencies indicate the number of occurrences per 100 words. Furthermore, the subordinators used to introduce all subordinate clauses were counted and their type-token ratios were computed for each text. Some texts contained quotations from other texts. These were excluded from the analysis.⁵

4. RESULTS

In the following section, the results of the current study are reported. In Section 4.1, the results of the overall frequencies of subordinate clauses are given. Section 4.2

Text	Mean	Std. deviation	Number
text 1 (LT1)	0.028	0.056	21
text 2 (LT2)	0.120	0.101	21
text 3 (LT3)	0.181	0.081	19
text 4 (LT4)	0.391	0.088	15
text 5 (LT5)	0.345	0.095	20
text 6 (LT6)	0.304	0.102	20
text 7 (LT7)	0.469	0.083	19

Table 2. Descriptive statistics for the subordination ratio.

reports the results of the subordinate clause types. Finally, in Section 4.3, the results of the use of subordinators are presented.

4.1. The subordination ratio

In general, most learners started to use subordinate clauses more frequently in text 2. In this text, the majority of the learners produced subordinate clauses (16 of 21). Only a few learners produced subordinate clauses already in text 1. [Figure 1](#) shows the group mean of the subordination ratios for each text.

LT stands for Learner Text, and the numbers refer to assignments 1 to 7. The full lines in the boxes represent the median of the subordination ratio, whereas the dotted lines show the means. The whiskers indicate the variation between learners. Dots above or underneath the whiskers are outliers. The subordination ratio is a relative frequency, meaning that, for example in text 1, 2.8% of the clauses that the average learner wrote were subordinate clauses.

The subordination ratio displays a clear increase in the first four texts, ranging from a mean of almost zero (0.028) in the first text, 0.120 in text 2, 0.181 in text 3, and reaching 0.391 in text 4. However, the mean subordination ratio drops again in text 5 (0.345) and is even lower in text 6 (0.304), but it ultimately reaches a peak of 0.469 in text 7. The descriptive statistics for the successive groups are given in [Table 2](#). The complete comparison for all groups can be found in the Appendix. A repeated measures ANOVA was conducted to compare the effect of the time of writing (and/or possibly text type) on the subordination ratio. The effect was found to be significant using the Greenhouse-Geisser correction ($F_{3,43, 41.12} = 65.11, p < 0.001$). This means that there was a statistical difference between groups.

For the sake of clarity, the relevant statistical values are given in [Table 3](#).

Post hoc comparisons using Tukey's contrasts found a statistical difference between the subordination ratio in the texts indicated with asterisks in [Table 3](#). The mean subordination ratios in text 1 differed significantly from the mean subordination ratios of all other texts. The mean subordination ratio in texts 2 and 3 showed no

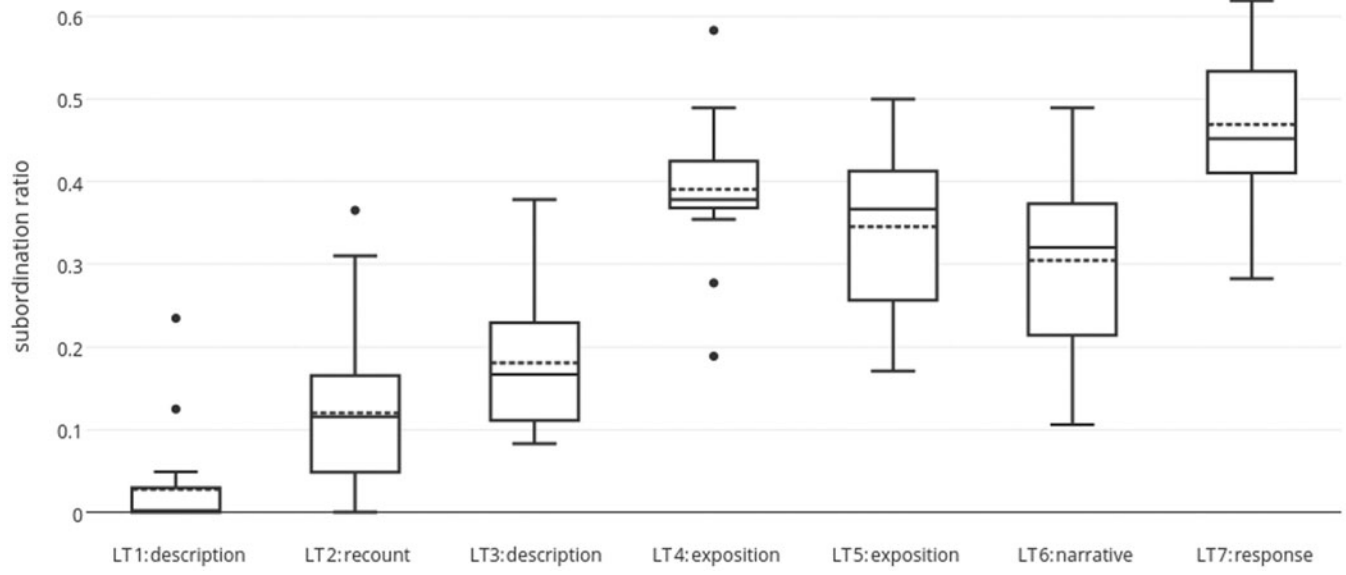


Figure 1. The subordination ratio in L2 learners of Swedish in investigated texts.

Texts	Estimate	Std. Error	<i>t</i> value	Pr(< <i>t</i>)
LT2 - LT1 == 0	0.092	0.025	3.755	0.003**
LT3 - LT2 == 0	0.069	0.025	2.773	not significant
LT4 - LT3 == 0	0.210	0.030	6.908	<0.001***
LT5 - LT4 == 0	-0.043	0.028	-1.519	not significant
LT6 - LT5 == 0	-0.041	0.025	-1.622	not significant
LT7 - LT6 == 0	0.163	0.026	6.383	<0.001***

Asterisks indicate a statistically significant difference between groups.

Table 3. Group comparisons of subordination ratio using Tukey's contrasts.

significant difference compared to each other, but they did differ significantly from the means of all other texts. However, the mean differences are smaller in the texts that were written after text 3. There were no significant mean differences between texts 4, 5, and 6. Although the comparison between texts 5 and 6, on the one hand, and text 7, on the other hand, showed significant results, there were no significant differences between the means of subordination ratios in text 4 and text 7. Effect sizes for all the comparisons showed rather strong effects for each significant comparison (with Cohen's *d* ranging between 1.13 and 2.85).

These statistical results show that, although the mean subordination indeed increases significantly in the very beginning of the language learning process, after 6 months it reaches a point where there are no significant differences between assignments. Only toward the last text does a significant increase still occur, but this increase does not necessarily indicate learners' progress in the complexity of their language use. In all likelihood, the subordination ratio can also be influenced by other factors, such as task type or task complexity.

To give an indication of what the subordination ratio is in Swedish native speakers' writing and what the possible influence of text type might be, the learner texts were compared to the texts written by native Swedish speakers. Although the sample is rather small, it can give a fair estimation of the subordination ratio in native written Swedish. In [Figure 2](#), the variation and mean of the subordination ratios between learners are visualized for each text, as well as the mean and variation within the native speakers' texts in five different text types.

[Figure 2](#) clearly reveals a different picture than the results in [Figure 1](#) with language learners only. Except for the first three texts, the mean subordination ratio for the native speakers' text types is lower than or similar to all non-native speakers' texts. The means of the subordination ratios for native speakers' and non-native speakers' texts were compared using a one-way ANOVA test. Significant differences were found between the three first learner texts and all native speakers' texts, whereas no significant results were found for differences between either of the native speakers' texts and the learners' texts 4 and 5. There was a significant difference between learner text 6 (narrative) and two of the native speakers' texts: expositions ($F = 3.59$,

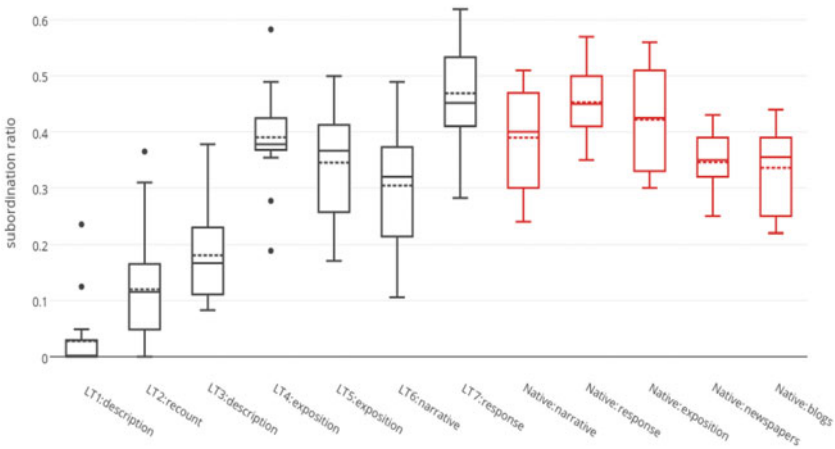


Figure 2. (Colour online) The subordination ratio in L2 learners and L1 speakers of Swedish in investigated texts.

$p < 0.02$), and responses ($F = 4.42, p < 0.01$). No significant differences were obtained between learner text 6 and the other three native speakers' texts (blogs, newspapers, and narratives). This means that learners increase the frequency of subordinate clauses from text 1 to text 3, but from text 4 onwards, after only 6 months of study, they have reached a native-like level. Comparisons using Tukey's contrasts found a statistical difference between LT 7 and two of the native speakers' registers (native speakers blog texts –LT 7, $F = -3.92, p < 0.01$; Native speakers Newspapers –LT 7, $F = -3.63, p < 0.02$). So, the learners even used a significantly higher degree of subordinate clauses in LT 7 than the native speakers in two text types.

In summary, the tests show that the learners in this study do not score significantly different from native speakers' writing on the subordination ratio in Swedish after 6 months of learner development. This indicates that the use of the subordination ratio is not very informative for language development without additional information. It is important to look at the subordinate clauses used by L2 learners in more detail in order to see if there are additional measures that may show differences between texts. In the remainder of the current paper, variations in subordinate clause types and subordinator use are investigated and related to subordination ratio.

4.2. Subordinate clause types

As Lambert & Kormos (2014) point out, one of the problems with the use of the subordination ratio as a measurement for syntactic complexity in L2 development is that all types of clausal subordination are treated in the same way. The measure does not take into account what kind of subordinate clauses the learners use. A

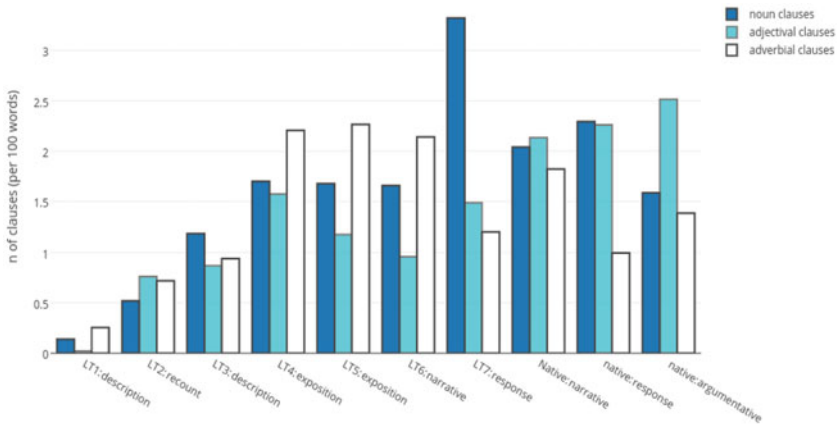


Figure 3. (Colour online) Differences between mean number of subordinate clause types.

high frequency of subordinate clauses does not necessarily mean that the learner has acquired a broad range of subordinate clause types or subordinating conjunctions.

In this study, the subordinate clauses were categorized according to the three subordinate clause types: noun clauses, adjectival clauses, and adverbial clauses. Clause types were not always equally distributed when looking at individual results. In general, most learners started to use subordinate clauses more frequently in text 2 or 3, but many of them used only one or two different clause types, or they clearly used one type more frequently than others even after that. Later on, the learners start to vary more between clause types. Figure 3 shows the average number of all subordinate clause types per text.

In the first two texts, subordinate clauses are used sporadically. In text 1, adjectival clauses are hardly used at all. Some learners use temporal adverbial clauses, and there are some individual examples of noun clauses. It is important to mention that only three learners produced subordinate clauses in text 1. In text 2, more learners started to use subordinate clauses, and 16 learners used subordinate clauses from text 2 onwards. However, there is no clear difference in the frequency of different subordinate clause types. Five learners are late developers, producing their first subordinate clauses in text 3 or 4. In text 3, the learners produced, on average, slightly more noun clauses than adjectival or adverbial clauses. In texts 4, 5, and 6, there is a clear pattern: adverbial clauses are used most frequently, occurring over two times per 100 words. The adverbial clauses are followed by noun clauses, and adjectival clauses occur least frequently. In text 7, there is a big shift, because in this text, noun clauses occur almost three times as often as adverbial clauses, and more than twice as often as adjectival clauses.

If we compare these frequencies to the frequency of clause types in the native speaker texts, we find that adverbial clauses have, in contrast with the learner texts,

Learner	Noun clause	Adjectival clause	Adverbial clause
CDU	3	2	4
CVZ	2	3	4
DBL	3	3	4
DSE	3	2	2
EDS	3	2	2
GMO	4	2	2
HVL	2	2	1
IGE	1	2	1
JGO	5	3	5**
LDS	2	2	3
LGA	2	2	2
LVB	3	2	2
MDM	3	2	4
MLA	3	4	2
MSC	4	4	4*
RTA	5	5	5*/**
SDN	4	1	3
SDS	3	2	2
TVC	3	2	3
VVP	5	3	3
VVV	2	4	2

*Missing data for text 3; **missing data for text 4.

Table 4. Texts with first use of subordinate clause types.

the lowest frequency in all text types. The adjectival clauses are instead the most frequently produced clause types for native speakers, except for the responses, where noun clauses are slightly more frequent.

So, looking at subordinate clause types, we find that there is indeed a difference between learners and native speakers and that there also is a clear difference between text types. Table 4 indicates for each learner in which text they start to use the various subordinate clauses.

Overall, only three learners produced subordinate clauses in text 1. From text 2 onwards, 16 learners used subordinate clauses. Five learners are late developers, producing their first subordinate clauses in text 3 or 4. Two learners did not write text 3, and for another two, text 4 was missing. These learners are marked with asterisks.

When it comes to the first appearance of subordinate clause types, the first adjectival clauses are produced already in text 2 by the majority of learners (16 of 21).⁶ Eight learners used their first adverbial clauses in the same text. For four learners, the first attestation of adverbial clauses was in text 3. For another five learners, this first attestation occurred in text 4. Overall, the first noun clauses of most

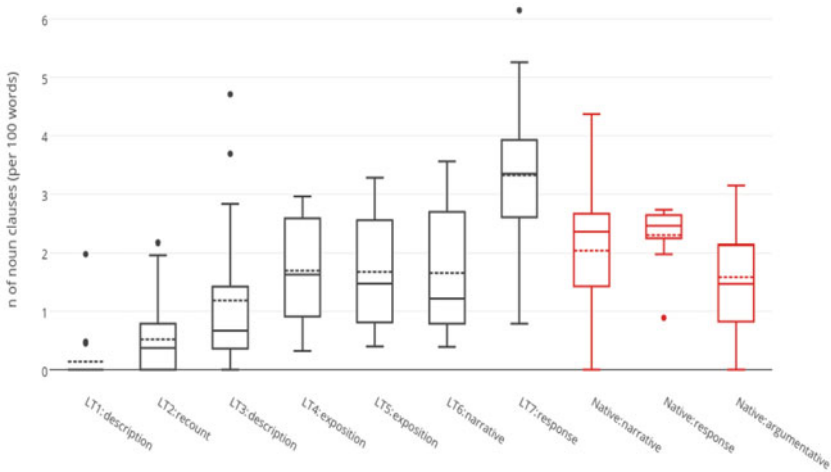


Figure 4. (Colour online) Normalized frequency of noun clauses produced by learners and native speakers of Swedish.

learners are attested later compared to the other clause types. Nine learners start using noun clauses in text 3. Three of them use noun clauses for the first time in text 4, and another three in text 5. Once the learners start using noun clauses, however, they use them, on average, more frequently than adjectival clauses. By text 5, all learners have used all subordinate clause types at least once.

We now look at the production of different clause types in more detail.

4.2.1. Noun clauses

The normalized frequency of noun clauses in the data is shown in [Figure 4](#)

The median is indicated by the full line, whereas the dotted lines show the mean frequencies. In learner text 1, there are barely any instances of noun clauses. There is a gradual increase of noun clauses from learner text 2 to learner text 4. Recall that all learners produced subordinate clauses in text 4, which increases frequencies in general. After text 4, however, there is a slight decrease in the frequency of noun clauses for the majority of learners in texts 5 and 6. The mean is almost identical in texts 4, 5, and 6. The fact that the boxes become bigger and the whiskers stand further apart over time means that there is more variation between learners. The mean and median of noun clauses in text 7 are exceptionally high. The whiskers are very far apart, meaning that there is much individual variation.

If we look at the native speakers' texts, we find a clear difference between text types. The means in all text types lie around two noun clauses per 100 words, which is a somewhat higher frequency than the number of noun clauses in all learner texts, except for text 7. However, the variation between texts is very limited in the responses,

Text	Mean	St. deviation	Data: <i>n</i>
LT1	0.139	0.444	21
LT2	0.520	0.680	21
LT3	1.188	1.348	19
LT4	1.699	0.910	15
LT5	1.678	0.921	20
LT6	1.658	1.069	20
LT7	3.323	1.303	19
Native: response	2.295	0.545	10
Native: exposition	1.587	1.010	10
Native: narrative	2.040	1.318	10

Table 5. Descriptive statistics noun clauses.

especially compared to the narrative texts, where there is much more variation. The median and mean of the native speakers is lowest in the expositions.

The descriptive statistics for the normalized number of noun clauses is outlined in [Table 5](#).

The mean frequency of noun clauses for the three native speakers' text types is higher than or similar to all non-native speakers' texts, except for learner text 7. Because a normal ANOVA does not take into account that the same subjects are tested multiple times, the data should be analyzed with a repeated measures ANOVA or a mixed-effects model. In this analysis, the longitudinal learner data were analyzed using mixed-effects model in *R* with subjects as a random factor, as advised by Larson-Hall (2010). The native speakers' texts were compared to the learner text using one-way ANOVA test.

The means of the normalized number of noun clauses in the learner texts were compared in order to analyze the effect of the time of writing on the normalized number of noun clauses. There was a significant main effect of developmental timing/text ($F(4, 10) = 30.87, p < 0.0001$) with a large effect size (partial eta squared = 0.93). This means that there is a significant difference between the normalized frequency of noun clauses in the learner texts.

Comparisons using Tukey's contrasts found a statistical difference between learner text 1 and all other texts except for text 2. The normalized mean of noun clauses in text 2 was found to be statistically different compared to texts 4, 5, 6, and 7, but there was no significant difference between texts 2 and 3. Another statistical difference was found between text 7 and all other learner texts. There were no statistical differences between texts 3, 4, 5, and 6. For the detailed statistics, see [Table 10](#) in the Appendix.

The fact that learners used more noun clauses in text 7 does not necessarily mean that they also produced more complex texts or more native-like clauses. The native

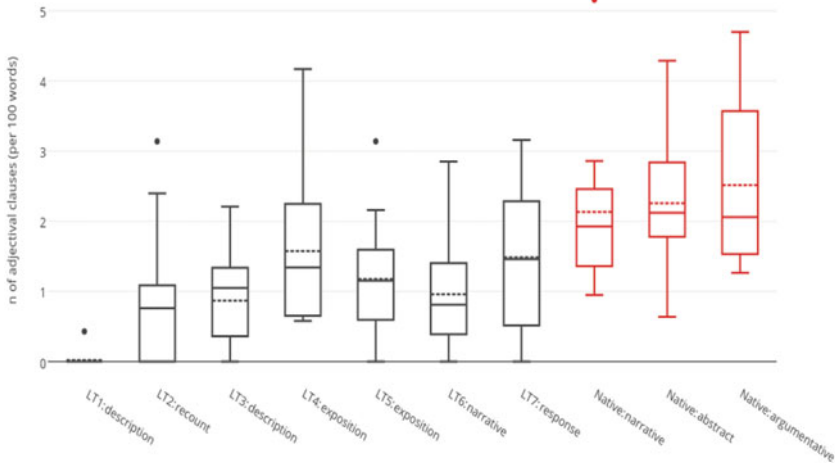


Figure 5. (Colour online) Normalized frequency of adjectival clauses produced by learners and native speakers of Swedish.

speakers' and non-native speakers' texts were compared using a one-way ANOVA test. Again, there was a significant main effect of developmental timing/text ($F(9, 155) = 15.2, p < 0.001$) with a large effect size (partial eta squared = 0.47). Post hoc comparisons using Tukey's contrasts found a statistical difference between learner text 1 and all native speakers' texts. Native speakers produced significantly more noun clauses than learners in text 1. Native speakers also produced statistically more noun clauses in the responses and the narrative texts compared to the learners in text 2. No significant results were found for differences between either of the native speakers' texts and the learners' texts 3, 4, 5, and 6. This means that learners increase the frequency of noun clauses from text 1 to text 3, but from text 4 onwards, after only 6 months of study, they do not produce statistically more or fewer noun clauses compared to native speakers of Swedish. A statistical difference was found between learner text 7 and two of the native speakers' texts: expositions ($F = -4.47, p < 0.001$) and narrative texts ($F = -3.30, p < 0.05$). This reveals that learners in text 7 used more noun clauses than the native speakers did in these text types. The relevant statistics are given in Table 11 in the Appendix.

4.2.2. Adjectival clauses

Let us now turn to the results for adjectival clauses. The results for the normalized frequencies of adjectival clauses are visualized in Figure 5, and the descriptive statistics are given in Table 6.

The longitudinal learner data were again analyzed using a mixed-effects model in *R* with subjects as a random factor. The native speakers' texts were compared to

Text	Mean	St. deviation	Data: <i>n</i>
LT1	0.021	0.094	21
LT2	0.869	0.678	19
LT3	0.869	0.678	19
LT4	1.575	1.022	15
LT5	1.178	0.761	20
LT6	0.960	0.818	20
LT7	1.487	0.968	19
Native: response	2.259	1.028	10
Native: exposition	2.517	1.249	10
Native: narrative	2.135	1.221	10

Table 6. Descriptive statistics adjectival clauses.

Null hypothesis	Estimate	Std. Error	<i>t</i> value	Sign.
LT2 - LT1	0.849	0.211	4.034	0.001**
LT3 - LT1	0.849	0.211	4.034	0.001**
LT4 - LT1	1.562	0.285	6.913	<0.001***
LT5 - LT1	1.179	0.207	5.683	<0.001***
LT6 - LT1	0.961	0.207	4.632	<0.001***
LT7 - LT1	1.467	0.211	6.989	<0.001***
LT4 - LT2	0.713	0.231	3.087	0.033*
LT4 - LT3	0.713	0.231	3.087	0.033*

* $P < 0.05$; ** $P = 0.001$; *** $P < 0.001$.

Table 7. Significant group differences between learners: adjectival clauses.

the learner text using a one-way ANOVA test with fixed effects only. In both tests there was a significant main effect of text/time, $F(4, 10) = 54.3$ and $F(1, 6) = 11.2$, respectively, both with a large effect size (partial eta squared = 0.96/0.65). Post hoc comparisons using Tukey's contrasts found statistical differences between the texts that are given in Table 7.

Table 7 clearly shows that there is a statistical difference between learner text 1 and all other texts. This means that the normalized frequency of adjectival clauses was lower in text 1 compared to the frequency of adjectival clauses in all other texts. Post hoc comparisons also showed that learners produced statistically more adjectival clauses in text 4 than in learner text 2 and learner text 3. No significant differences were found between other learner texts.

Post hoc comparisons for the one-way ANOVA test showed that native speakers produced statistically more adjectival clauses than the learners in texts 1, 2, 3, and 6 regardless of native speakers' text type. Significant differences were also found between learner text 5 (exposition) on the one hand and two of the native

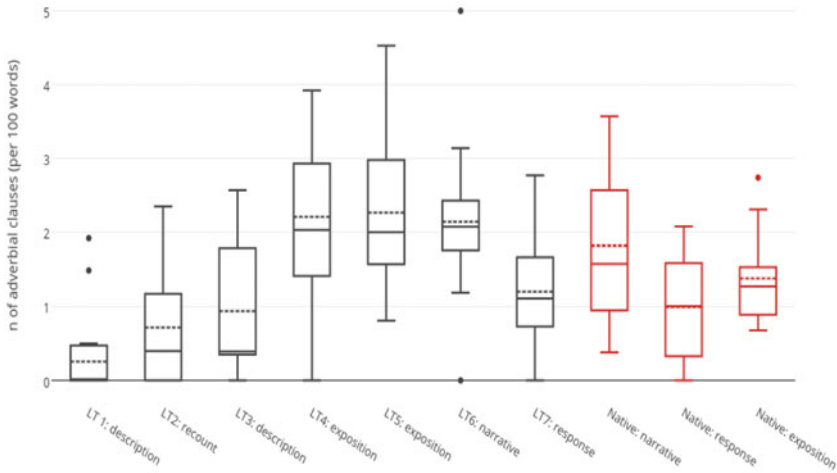


Figure 6. (Colour online) Normalized frequency of adverbial clauses produced by learners and native speakers of Swedish.

Text	Mean	St. deviation	Data: <i>n</i>
LT1	0.021	0.094	21
LT2	0.869	0.678	19
LT3	0.869	0.678	19
LT4	1.575	1.022	15
LT5	1.178	0.761	20
LT6	0.960	0.818	20
LT7	1.487	0.968	19
Native: response	2.259	1.028	10
Native: exposition	2.517	1.249	10
Native: narrative	2.135	1.221	10

Table 8. Descriptive statistics adverbial clauses.

speaker texts: responses ($F = 3.31$, $p < 0.05$) and expositions ($F = 4.10$, $p < 0.05$). However, no statistical difference was found between learner text 5 and the native speakers' narrative texts. There were no significant differences between either of the native speakers' texts and learner texts 4 or 7. No statistical differences were found between the different native speakers' text types. See Table 12 in the Appendix for full statistics.

4.2.3. Adverbial clauses

Finally, we look at the frequency of adverbial clauses in the data. The results are visualized in the boxplots in Figure 6. The descriptive statistics are listed in Table 8.

The learner texts were compared using mixed-model ANOVA with participants as a random factor, as was done for the other two clause types. Also for adverbial clauses, text/time was reported as a significant factor, $F(1, 6) = 18.7, p < 0.0001$. Post hoc tests using Tukey's contrasts showed that there were significant differences between the means of learner text 1, on the one hand, and texts 4, 5, 6, and 7, on the other hand. More adverbial clauses were used in the latter. The normalized frequency of adverbial clauses was also found to be statistically lower in text 2 than in learner texts 4, 5, and 6. The mean number of adverbial clauses in these three texts was also significantly higher than the means of learner text 3. Perhaps more surprisingly, the normalized means in learner text 7 were also reported to be statistically lower than the means of the three aforementioned texts (4, 5, and 6).

If we look at the differences between the learner texts and the native speaker texts, we find significant differences for the same texts as we did for the adjectival clauses. Post hoc comparisons with Tukey's contrast found statistical differences between all native speakers' texts, on the one hand, and learner texts 1, 2, 3, and 6, on the other hand. Statistical differences were also found between learner text 5 (exposition) and two of the native speakers' texts: expositions ($F = 3.31, p < 0.05$) and responses ($F = 4.098, p < 0.01$). But instead of more clauses, the native speakers produced significantly fewer adverbial clauses than the learners did in learner texts 5 and 6. The relevant data are summarized in [Table 14](#) in the Appendix.

4.2.4. Interim summary

Overall, these figures confirm the picture in [Figures 1](#) and [2](#) for all subordinate clauses. From texts 1 to 4, the use of all subordinate clause types increases. After text 4, the number of subordinate clauses drops for all clause types. After that, the number of nominal clauses stabilizes and reaches a peak at text 7. The number of adjectival clauses also shows a slight increase in text 7, but it does not top out in the last text. The most remarkable result is perhaps the picture drawn for adverbial clauses. The number of adverbial clauses peaks at text 4, after which it drops in all of the following texts. This indicates that in the development there are differences between the texts even after text 3, which becomes clear only once we look beyond the subordination ratio.

4.3. Choice of subordinators

In order to test the second claim by Lambert & Kormos (2014) on the role of item-based frequencies, we will take a closer look at the use of subordinators. Subordinators play an important role in clause linkage. They are important syntactic linking devices and indicators for the recognition of subordinate clauses for the language learners, while at the same time adding a semantic aspect to the acquisition of subordination. Although Lambert & Kormos (2014) mainly refer to frequently occurring verbs that

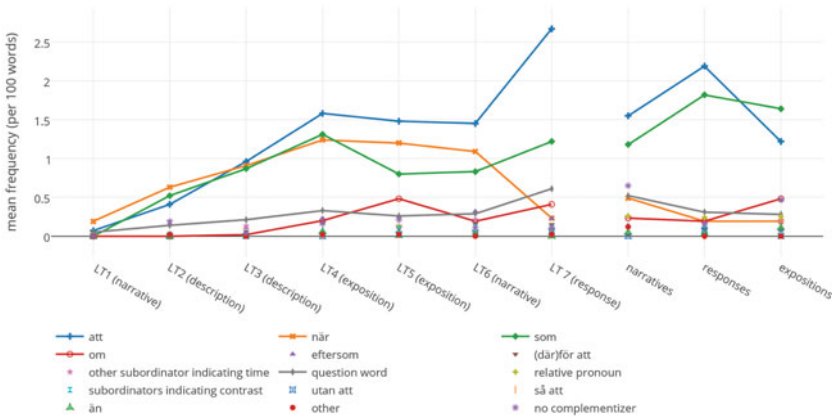


Figure 7. (Colour online) Normalized frequencies of subordinators used in the texts by learners and native speakers.

usually are accompanied by subordinate clauses, subordinators might give a better overall picture, because they are present in all subordinate clause types. This is not the case for complement-taking predicates, which only give insight into noun clauses. The use of complement taking predicates is studied elaborately in Wijers (forthcoming). In the current paper, the results for the usage of subordinators by learners are analyzed and compared to the subordinators that native speakers used in the data.

Overall, most learners tend to use a limited variety of subordinators within a text. Three subordinators were used most frequently in the majority of learner texts:

- (i) the subordinating conjunction *att* ‘that’, typically used to introduce a noun clause
- (ii) the relative pronoun *som* ‘which/who’, typically used to introduce an adjectival clause
- (iii) the adverb *när* ‘when’ introducing an adverbial clause⁷

Figures 7 and 8 show the frequencies of subordinators used for each text (i.e. actual use, not necessarily target-like use). In Figure 7, the normalized actual frequencies are given, whereas Figure 8 presents the relative frequencies of subordinators used for each learner text, as well as for the native speakers.

As we have seen in the previous results, there were hardly any subordinate clauses in the learner text 1. But for the few learners that already used subordinate clauses, *när* ‘when’ is the subordinator that is used most often. It occurs 0.19 times per 100 words, on average, and accounts for 48% of the total number of subordinators used. *Att* ‘that’ is ranked second, but it occurs only 0.07 times per 100 words (18% of all subordinators). In learner text 2, when more learners started using subordinate clauses, *när* is still the most frequent subordinator (together with *som*). It is used 0.63

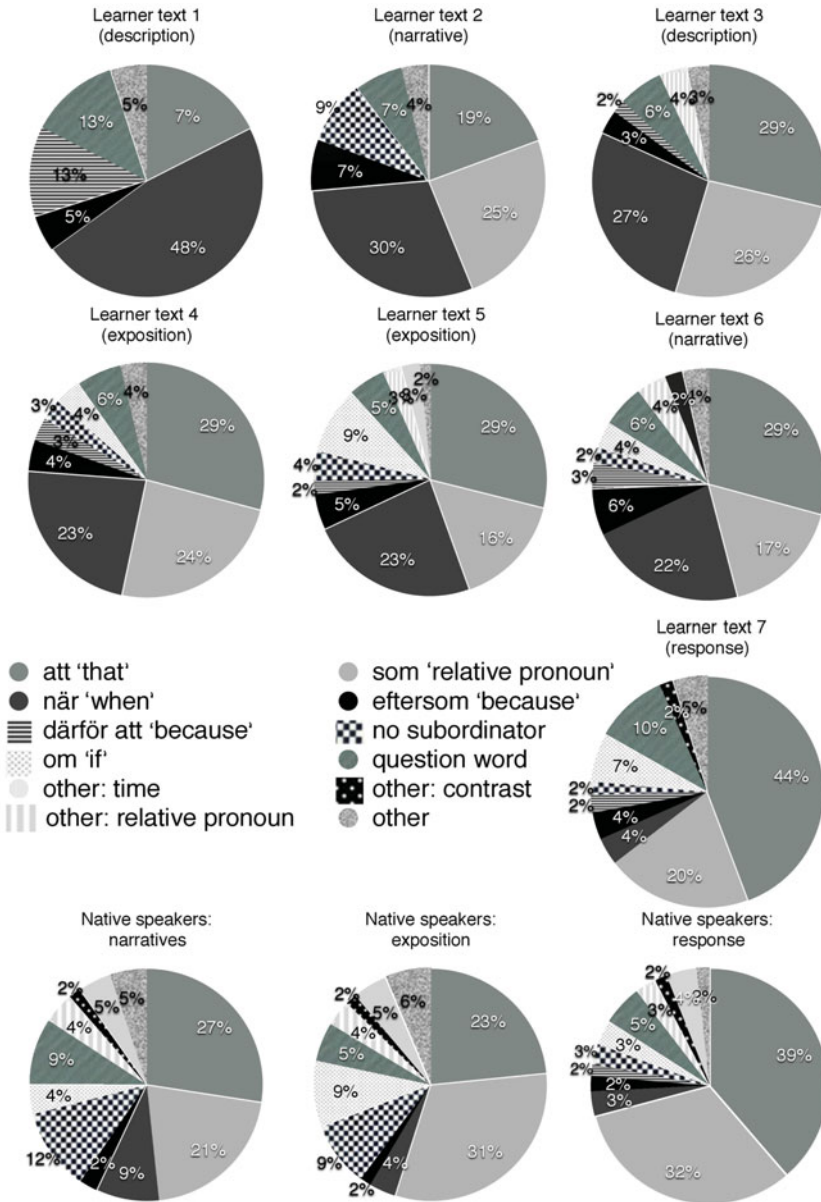


Figure 8. (Colour online) Percentages of subordinators used in the texts by learners and native speakers.

times per 100 words, on average, and it accounts for 30% of the subordinators. The fact that the percentage decreases actually means that more subordinators are being used and the variety of subordinators increases. The relative pronoun *som* occurs a little less frequently than *när*. At that point, the third most frequent subordinator is *att* ‘that,’ occurring 0.41 times per text, making up 19% of all subordinators. From text 3 onwards, *att* becomes the most frequent subordinator, on average, and will remain so in all learner texts. *När* and *som* follow in second and third position, respectively. For each text, the range of different subordinators increases, although most subordinators (other than *att*, *när*, and *som*) are used more sporadically. Gradually, more subordinators start to occur in the written texts, such as *därför att* and *eftersom* ‘because’ from text 3 onwards, as well as the subordinating conjunction *om* ‘if,’ which occurs for the first time in text 4 (although it only appears as one of the most frequent conjunctions in text 5). *Att* covers a much higher proportion of the total number of subordinators used in text 7 than in all other texts: it occurs 2.67 times per 100 words, on average, which equals 44% of all subordinate clauses.

Let us now look at the use of subordinators among native speakers. Their use differs in two respects from the learners. First, *som* was used relatively more often by native speakers than by the learners in this study. In the expositions, *som* was even the most frequently used subordinator, occurring on average 1.64 times per text and in 31% of the subordinate clauses. *Att* comes in second position with 23%, occurring 1.22 times per 100 words, on average. In the responses, *som* occurred 1.82 times per 100 words in 32% of the subordinate clauses produced by native speakers, which is significantly more frequent than in learner text 1 ($F = -6.13, p < 0.01$), learner text 2 ($F = -4.38, p < 0.01$), learner text 5 ($F = -3.43, p < 0.05$), and learner text 6 ($F = -3.31, p < 0.05$). *Att* occurred somewhat more frequently (2.19 times per 100 words; 39% of all subordinate clauses) in the responses, but the frequencies were comparable to the occurrence of *att* in the learner texts. The narrative texts contained the least occurrences of *som*, namely 1.18 times per 100 words (21%), which is only significantly different from the frequency of *som* in learner text 1 ($F = -3.96, p < 0.01$).

The second important difference between native speakers and learners is that *när* is much more frequent than in the texts written by native speakers. In learner texts 1 and 2, *när* was the most frequently used subordinator, and in the other texts, the number of occurrences of *när* increased and remained among the three most produced subordinators. In the native speakers’ texts, this is not the case. *Att* and *som* were clearly used more often than all other subordinators. The frequency of *när* does not stand out like it does in the texts written by learners.

The third difference between the learners and native speakers in this study is that the native speakers express the subordinator much less often than the learners do. Especially in the narrative texts, native speakers did not use any subordinator in 12% of all subordinate clauses (0.65 times per 100 words). In the expositions,

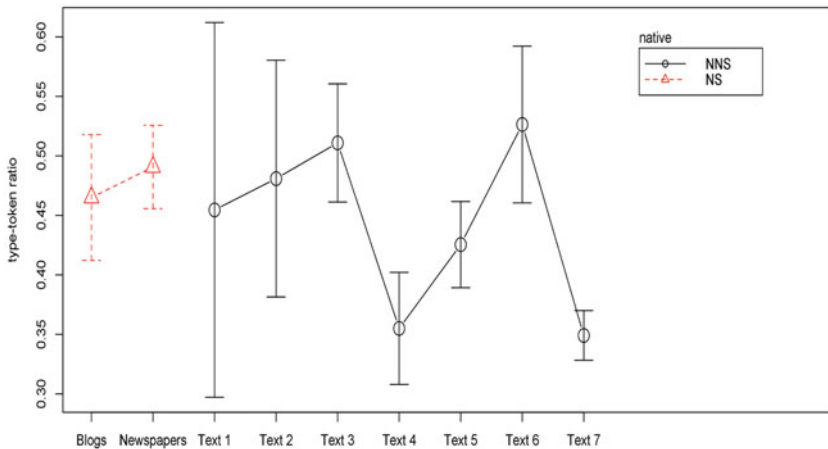


Figure 9. (Colour online) Type-token ratios of subordinators in learners' and native speakers' texts.

no subordinators were produced in 9% of the subordinate clauses (0.47 times per 100 words), and in the responses, the subordinator was not expressed in 3% of the subordinate clauses (0.16 times per 100 words). Learners, on the other hand, left out the subordinator 0.21 times per 100 words at the most (in text 5). In learner text 2, the subordinator is left out in 9% of the subordinate clauses (0.19 times per 100 words).

The variety of subordinators used can also be measured by looking at type-token ratios. The average type-token ratios for subordinators used are given in Figure 9. A high type-token ratio indicates a high level of variation.

There is a slight increase from text 1 to text 3, followed by a strong decrease in text 4. In texts 5 and 6, the type-token ratio increases again followed by another drop in text 7. Remarkable in this graph are the low means for two of the learner texts: texts 4 and 7. These texts were exactly the ones that had a very high subordination ratio.

A Pearson product-moment correlation coefficient was computed to assess the relationship between the subordination ratio and the type-token ratio of subordinators in the learner texts in which almost all learners used subordinate clauses (text 3 — text 7). There was a significant negative correlation between the two variables with a large effect size ($r = -0.4554$, $n = 53$, $p = 0.0006$). This means that the higher the subordination ratio of a text, the lower the type-token ratio of subordinators was. No significant correlation was found in the texts written by native speakers. A scatter plot summarizes the results (see Figure 10).

If we look at both factors, frequency of subordination and variety of subordinators/subordinate clause types used, we arrive at a different picture. In Figure 11, the subordination ratios for each text were multiplied by the type-token

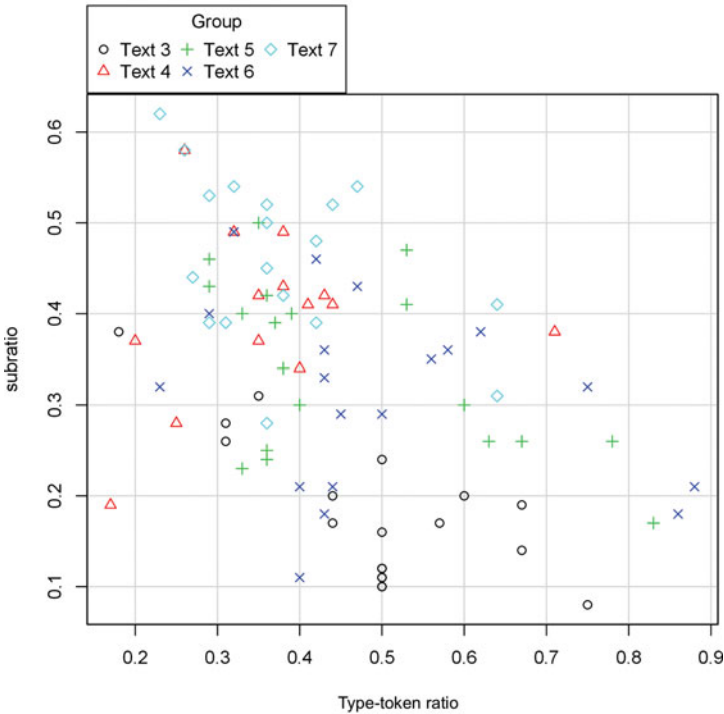


Figure 10. (Colour online) Correlation between subordination ratio and type-token frequency of subordinators in intermediate L2 learner texts.

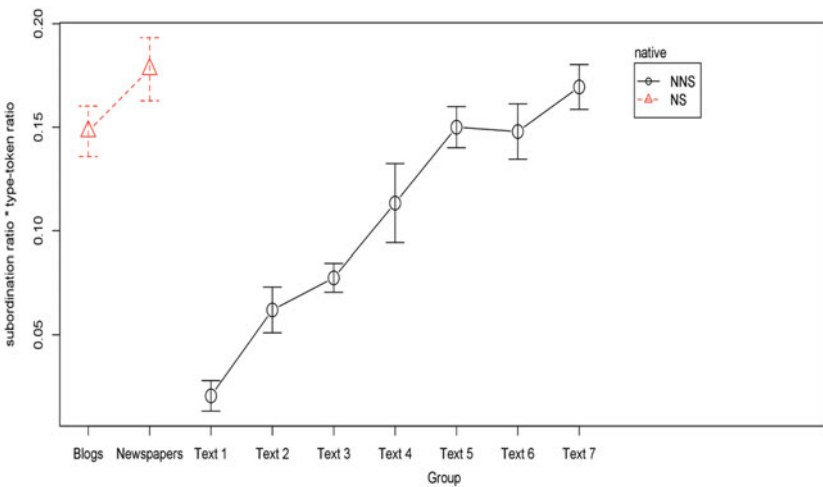


Figure 11. (Colour online) Subordination ratio multiplied by type-token ratios of subordinators in learners' and native speakers' texts.

ratios. The result is a more linear line and an increase for each text (except for text 6, which is at the same level as text 5). This shows that there are changes visible over time and that learners actually do develop, albeit in different areas. Combining these factors could be a possible solution for getting around the problem of text types.

5. DISCUSSION

The results of this study show that the subordination ratio increases in the earlier stage of development and quickly reaches a plateau. This pattern is reminiscent of earlier studies (Monroe 1975; Cooper 1976; Flahive & Snow 1980; Perkins 1980; Bardovi-Harlig & Bofman 1989). These results reflect Pallotti's arguments that complexity does not have to increase over time in learner development and that complexity is not the same as developmental or cognitive complexity (Pallotti 2014:2–4). In line with these arguments, it is worthwhile to look at native speaker data. If we assume that learners develop toward more native-like language use, then mean subordination ratios for native speakers should be investigated as well. As Pallotti (2009:598) argues:

[E]specially for fluency and complexity, native speakers' baseline data are crucial, not because learners' aim is necessarily to behave like native speakers, but because looking at what native speakers do may overcome the researchers' bias toward seeing learners as defective language users, who always need to 'do more'.

After six months of learner development, the data in this study showed no significant differences between the subordination ratio of learners and native speakers' writing. In other words, this suggests that the learners do not have to 'do more' when it comes to the frequency of subordinate clauses. The question is whether there are other additional measures that do reflect learner development even beyond the point where no clear changes are found in the frequency of subordinate clauses.

If we look at subordinate clause types, there are differences between learners and native speakers in this study. The learners in this study started producing noun clauses later, but from the moment they do use them, noun clauses quickly become more frequent than adjectival clauses. The difference between adverbial and adjectival clauses, on the one hand, and noun clauses, on the other hand, is that noun clauses function as arguments to the verb in the main clause, whereas adjectival and adverbial clauses do not. Because certain verbs require noun clauses as their complements, it is not remarkable that noun clauses occur later in the data than the other two clause types for most learners. If we look at the first verbs that were acquired in the L2, we find that these verbs typically do not require clausal complements. Adverbial clauses,

however, are used to anchor situations in time and space, and many writers intend to do this from early on.

The data also show that it is hard to prove whether differences between texts are due to development or text genre. Differences between native speakers in different text types were often bigger than the differences between learners and native speakers. Therefore, it is interesting to look at some examples of individual variation between learners. The individual analyses often reveal that the use of many subordinate clauses does not necessarily entail syntactic complexity. In other words, a text with a high subordination ratio is not necessarily more complex. Examples from two of the learners, SDN and CDU, will illustrate this. Learner SDN almost exclusively uses one particular type of clause in text 3. This is also the first text in which she systematically uses subordinate clauses. Some examples from text 3 are given in examples (6a–c).

- (6) a. När han skrattar, han har gropar i kinderna.
 when he laugh.PRS he have.PRS dimples in cheeks.DEF
 ‘When he laughs, he has dimples in his cheeks.’
- b. När vi går på restaurang med vår släkten, han ha
 when we go.PRS on restaurant with our family.DEF he have.INF
 en svart kostym på sig.
 a black costume on REFL
 ‘When we go to a restaurant with our family, he wears a black costume.’
- c. När är vi på resa, då tror man som vi är
 when be.PRS we on trip then believe.PRS people SUBR we be.PRS
 tvillingar.
 twins
 ‘When we are on a trip, people think that we are twins.’

In examples (6a–c), the sentence starts with an adverbial clause introduced by the subordinating adverb *när* ‘when’, and is followed by the main clause. The first two sentences have V3 word order, as opposed to the obligatory V2 word order in the target language Swedish. Thus, SDN does not apply inversion in the main clause. Interestingly, Ganuza (2011) shows that V3 word order after clause-initial subordinate clauses is common in the Swedish of adolescents in multilingual urban settings, too. However, in example (6c), the learner uses the correct word order after the adverbial clause, after she inserted an adverb *då* ‘then’. Hence, the learner seems to be aware of the V2 rule, but it is possible that the learner considers the adverbial clause and the main clause to be two separate main clauses instead of a main and subordinate clause combined in a complex sentence. Another explanation could be that *när* triggers main clause word order, because *när* is also a frequently used question word in main clauses. The learners have probably seen *när* combined with main clause structures much more frequently than subordinate clauses at this point. Although the subordinate clauses in examples (6a) and (6b) are not ungrammatical (the word order

cannot be checked because clauses are subject-initial and lack adverbials), the V3 word order seems to reveal that the difference between subordinate clauses and main clauses is not acquired yet. It is unclear whether these clauses should be analyzed as subordinate clauses or not. The fact that learner SDN uses commas between the main and the subordinate clause, which is not common in Swedish (or in the L1 Dutch), seems to point out that the learner analyzes the two clauses as two separate main clauses. Another possibility is that the use of commas is transfer from English punctuation.

In the absence of any sentence adverbials, the word order of subordinate clauses is also seemingly target-like in the sentence (7a–c) produced by learner CDU. The learner who produced the examples in (7a–c) started using subordinate clauses in text 2. Seven out of 8 subordinate clauses are restrictive relative clauses in this text.

- (7) a. Den är en stad var gamla romarna bodde förr!
 that be.PRS a city where old romans live.PST before
 ‘That is a city where old Romans lived before!’
- b. Det är slottet var kung Henry VIII bodde förr.
 that be.PRS castle.DEF where king Henry VIII live.PST before
 ‘That is the castle where King Henry VIII lived before.’
- c. London var den sista stad som vi besökte under vår resa.
 London be.PST the last city SUBR we visit.PST during our trip
 ‘London was the last city which we visited during our trip.’

The sentences in examples (7a–c) are structurally very similar; they all start with the subject followed by a copula verb and the predicate which is specified by a relative clause. Overall, many learners turned out to be rather conservative in their choice of subordinate clause types and in their overall sentence structures. A high relative frequency of subordinate clauses does, in other words, not reveal how varied the learner’s use of subordinate clauses is. A text with a high subordination ratio is, therefore, not necessarily more complex, because a text containing only sentences with the same syntactic structure can hardly be called syntactically complex.

Also in the use of subordinators, we find differences between learners at different stages and between learners and native speakers. This is especially apparent for the subordinator *när* ‘when’ in this study. This subordinator seems to fill a gap for the learners. Both noun clauses and adjectival clauses have one particular subordinator that is clearly used more frequently than other possible subordinators (*att* and *som*, respectively). However, adverbial clauses can be introduced by a number of subordinators. For most learners, *när* fulfills this function for adverbial subordinate clauses. Transfer might also play a role in some cases. In Dutch, the subordinator *als* can be used both as a temporal subordinator ‘when’ and as a conditional subordinator ‘if,’ whereas Swedish has two separate subordinators for this. This could also explain why the learners in this study use *när* more often.

With regard to *som*, there seems to be some confusion about the subordinator due to transfer from either Dutch or English where *that* can be used for both relative and *that*-clauses. In example (6c), for instance, the learner wrote *som*, the subordinating pronoun that is commonly used to introduce relative clauses instead of the grammatical *att* ‘that,’ the subordinator that is used in Swedish to introduce noun clauses. Another example of clear transfer is the use of *var* in several learner texts, such as in examples (7a) and (7b). In the learners’ native language Dutch, as well as in the L2 English, *waar/where* is used as a locative adverbial to introduce spatial relative clauses. However, in the target language Swedish *där* ‘there’ is the subordinating equivalent of the adverb *var*.

The fact that some subordinators, like *därför att* and *eftersom* ‘because’ and *om* ‘if’ are hardly produced before learner texts 4 and 5, can partly be explained by the text type, which are descriptive and narrative in the first texts, whereas subordinators of causality are logically used more frequently in expositions. This is also illustrated by the fact that *att* covers a much higher proportion of the total number of subordinators used in text 7 than in all other texts: it is used to introduce 44.3% of all subordinate clauses. This is not surprising given the reported high number of noun clauses in this text as a result of the many reporting verbs that occur in this text type. These reporting verbs often require noun clauses as an argument. However, the high frequency is not only due to text type. Especially, *eftersom* is used relatively frequently in all texts in which learners produced subordinate clauses and occurs in the top five of most used subordinators in all of them. To a lesser extent, this is also true for *om* ‘if’.

When it comes to subordinators, the results showed another important difference between native speakers and learners. The native speakers expressed the subordinator much less often than the learners did. This reinforces the idea that learners need subordinators to build subordinate clauses, because subordinators help them to recognize subordinate clauses. It seems as if learners first need to learn which subordinators to use to introduce different types of subordinate clauses before they can start leaving them out.

There are different constructions where the subordinator can be covert in Swedish. Some examples from the native speakers’ texts are shown in (8a) and (8b). The two examples here show the most common occurrences of subordinator omission.

- (8) a. De tyckte (att) det var pinsamt att förlora,
 They think.PST (that) it be.PST embarrassing to lose.INF
 tror jag.
 believe.PRS I
 ‘They thought (that) it was embarrassing to lose, I guess.’
- b. Men ärligt talat, är det inte viktigare att prata
 but honestly speak.PRF is it not more.important to talk

om de utmaningar (som) Sverige står inför?
 about DEF challenges (that) Sweden stand before
 'But honestly, isn't it more important to talk about the challenges (that)
 Sweden is facing?'

In example (8a), the complementizer *att* is left out of the noun clause, which is possible after certain predicates (see Teleman 1999:46–82). In example (8b), the relative pronoun *som* is not expressed in the adjectival clause, which is grammatical in restrictive relative clauses if the subject is overt. Interestingly, both examples would be grammatical in English, but not in the learners' L1 Dutch. This means that the overt expression of subordinators could also be caused by L1 transfer.

Interestingly, in almost all cases where the subordinator was left out in the learner texts, the omission was target-like. Only half of the learners omitted a subordinator in one or more texts, and only eight learners left out subordinators in more than one text. In almost half of the cases of subordinator omission (44%), these learners left out *som* in restrictive relative clauses, such as in examples (9) and (10).

- (9) Det var första gången [vi besökte Egypten].
 It be.PST first time.DEF we visit.PST Egypt
 'It was the first time we visited Egypt.'
- (10) En sak [jag älskade där] var vackra blommorna.
 One thing I love.PST there be.PST beautiful flowers.DEF
 'One thing I loved there were the beautiful flowers.'

In 19% of the cases where learners omitted the subordinator, *att* was left out. This happened only after the verbs *tycka* 'think, be of the opinion' or *tro* 'believe, think.' Two examples are given in (11) and (12).

- (11) Svenskarna trodde det var jätterolig.
 Swedes.DEF believe.PST it be.PST very.funny
 'The Swedes thought it was very funny.'
- (12) Jag tycker det är konstigt.
 I find.PRS it be.PRS weird
 'I think it is weird.'

The other 38% of the subordinator omissions occurred in coordinated subordinate clauses. An example of this type is presented in (13).

- (13) Det skulle vara bättre [när man har bara en säck] och
 it will.PST be.INF better when one have.PRS just one bag and
 [använder den alltid].
 use.PRS it always
 'It would be better when you have only one bag and use it all the time.'

The second coordinated subordinate clause is not introduced by a subordinator, so this clause is interpreted as an adverbial clause, like the first subordinate clause. Out of all the subordinators that the native speakers left out, 19% were coordinated clauses and 11% was an omission of *att* after the verbs *tycka* ‘think, be of the opinion’ or *tro* ‘believe, think.’ In the majority of omissions, 64% occurred before restrictive relative (*som*-)clauses.

The differences that were found between subordinate clause types and the use of subordinators give interesting additional information beyond general subordinate clauses and frequencies. The results of this study suggest that type-token ratios of subordinate clause types and subordinators used are factors that could be used in addition to the subordination ratio to measure structural complexity. Taking into these aspects can clearly compensate for factors that remain obscured in general frequency measures, most prominently variation in the use of subordinate clauses. In fact, the text with the highest subordination ratio scored significantly lower on the type-token ratio.

However, we have to remain cautious as to whether this also reflects development, because text type in particular has a great influence on the results as well. More research is needed to investigate whether, for instance, different target languages and different L1s in the same learner context yield similar results. We also have to bear in mind that type-token ratio is sensitive to quantity. Hence, a higher number of subordinate clauses by default entails a higher number of subordinators, which in turn entails that their type-token ratio will decrease.⁸ I suggest that the association of these factors be investigated in future studies. Moreover, it would be interesting to look at the interaction between text type and development over time. Overall, syntactic complexity should be investigated as a functional concept if we want to apply it to L2 studies.

6. CONCLUSION

Following the criticism of a number of scholars (Baten & Håkansson 2015; Bulté & Housen 2012; Norris & Ortega 2009; Pallotti 2009, 2014; Lambert & Kormos 2014), the present study examined whether there are other factors besides the frequently used subordination ratio that are valid indicators of written L2 development. In line with Lambert & Kormos (2014), the study focused on variation in the use of subordinate clauses. In particular, the following three points, which were mentioned by Lambert & Kormos (2014) as directions that CAF research should take in order to better reflect L2 development and performance in measuring complexity, were put to the test: (i) the type of subordinate clause used, (ii) item-based (lexical) frequencies, and (iii) text genre. The data used in the current study consisted of longitudinal written production from 21 L2 learners of Swedish as a foreign language and 50 native speakers’ texts in various genres, all comparable to the genres of the L2 texts.

First, the results show that many of the subordinate clauses were simple and similar to each other. Learners already reached a stable subordination ratio value after about 6 months from the start of learner development. Furthermore, no significant differences were found after this point, when L2 learners were compared to native speakers. The results showed that a higher subordination ratio indeed does not equal language development, as has already been pointed out by Pallotti (2009, 2014) and Baten & Håkansson (2015).

Second, the study showed that it is important to assess, besides the general frequency measures, the breadth of variation present in the learner data, which has been suggested by Lambert & Kormos (2014). A closer examination of the different types of subordinate clauses used showed that there were differences between texts that were written during different moments of language development, and between learners and native speakers. The biggest difference between learners and native speakers was found in the use of adjectival clauses. The differences are not necessarily caused by development but could also be linked to text type. This was confirmed by the differences between various native speaker text types. Differences between native speakers in different text types were often bigger than the differences between learners and native speakers.

Item-based frequencies were investigated by means of subordinators rather than verbs to include all types of subordinate clauses. Most remarkably, the variety of subordinators used increased over time. Native speakers used a broader range of subordinators than learners did, on average. In fact, subordination ratio turned out to be negatively correlated with type-token ratio for the use of subordinators. Combined, the subordination ratio and type-token ratio of different subordinators showed a linear development over time. Differences were also found regarding the use of certain specific subordinators. On average, native speakers left out the subordinator much more often than the L2 learners, who seem to benefit from the overt expression of subordinators.

In sum, the present study confirmed that the three aspects mentioned by Lambert & Kormos (2014) are important factors to take into account in syntactic complexity research in SLA. Furthermore, the importance of baseline data of native speakers was stressed to show that the differences between beginning or intermediate learners and native speakers are not always as big as predicted. To arrive at a fully representative view of subordination and its linguistic complexity, it is interesting and necessary to take into consideration factors such as syntactic and semantic variation in subordinate clause types use and language-specific factors.

ACKNOWLEDGEMENTS

I am grateful to Anders Agebjörn, Kristof Baten, Johan Brandtler, Gisela Håkansson and four anonymous *Nordic Journal of Linguistics* referees for valuable comments

on an earlier version of this paper. I would also like to thank Victoria Johansson and Per Holmberg for letting me use their data, and Svenska Institutet for funding a research stay at Gothenburg University. Finally, I thank Eric Lander for proofreading this paper.

APPENDIX

Statistics: group comparisons

Null hypothesis	Estimate	Std. Error	<i>t</i> value	Sign.
LT2 - LT1 == 0	0.09227	0.02710	3.404	0.0150*
LT3 - LT1 == 0	0.15314	0.02781	5.507	<0.001***
LT4 - LT1 == 0	0.36270	0.02969	12.216	<0.001***
LT5 - LT1 == 0	0.31748	0.02744	11.570	<0.001***
LT6 - LT1 == 0	0.27651	0.02744	10.077	<0.001***
LT7 - LT1 == 0	0.44110	0.02781	15.862	<0.001***
LT3 - LT2 == 0	0.06087	0.02781	2.189	0.3087
LT4 - LT2 == 0	0.27043	0.02969	9.108	<0.001***
LT5 - LT2 == 0	0.22521	0.02744	8.207	<0.001***
LT6 - LT2 == 0	0.18424	0.02744	6.714	<0.001***
LT7 - LT2 == 0	0.34883	0.02781	12.544	<0.001***
LT4 - LT3 == 0	0.20956	0.03034	6.908	<0.001***
LT5 - LT3 == 0	0.16434	0.02814	5.841	<0.001***
LT6 - LT3 == 0	0.12337	0.02814	4.385	<0.001***
LT7 - LT3 == 0	0.28796	0.02849	10.106	<0.001***
LT5 - LT4 == 0	-0.04522	0.03000	-1.507	0.7396
LT6 - LT4 == 0	-0.08619	0.03000	-2.873	0.0694
LT7 - LT4 == 0	0.07840	0.03034	2.584	0.1394
LT6 - LT5 == 0	-0.04097	0.02777	-1.475	0.7585
LT7 - LT5 == 0	0.12362	0.02814	4.393	<0.001***
LT7 - LT6 == 0	0.16459	0.02814	5.850	<0.001***

Table 9. Group comparisons of subordination ratio using Tukey's contrasts.

Null hypothesis	Estimate	Std. Error	<i>t</i> value	Sign.
LT3 - LT1	1.027	0.292	3.511	0.008**
LT4 - LT1	1.546	0.313	4.933	<0.001***
LT5 - LT1	1.559	0.288	5.408	<0.001***
LT6 - LT1	1.539	0.288	5.337	<0.001***
LT7 - LT1	3.195	0.292	10.926	<0.001***
LT4 - LT2	1.133	0.320	3.536	0.007**
LT5 - LT2	1.146	0.296	3.869	0.002**
LT6 - LT2	1.126	0.296	3.800	0.003**
LT7 - LT2	2.781	0.300	9.282	<0.001***
LT7 - LT3	2.168	0.300	7.235	<0.001***
LT7 - LT4	1.649	0.321	5.136	<0.001***
LT7 - LT5	1.636	0.296	5.534	<0.001***
LT7 - LT6	1.656	0.296	5.603	<0.001***

Table 10. Significant group differences noun clauses: between learners.

Null hypothesis	Estimate	Std. Error	<i>t</i> value	Sign.
Native: response - LT1	2.156	0.382	5.642	<0.001***
Native: exposition - LT1	1.448	0.382	3.790	0.008**
Native: narrative - LT1	1.901	0.382	4.975	<0.001***
Native: response - LT2	1.775	0.382	4.646	<0.001***
Native: narrative - LT2	1.520	0.382	3.978	0.004**
Native: exposition - LT7	-1.736	0.389	-4.466	<0.001***
Native: narrative - LT7	-1.283	0.389	-3.301	0.037*

P* < .05; *P* < .01; ****P* < .001.

Table 11. Significant group differences noun clauses: between learners and native speakers.

Null hypothesis	Estimate	Std. Error	<i>t</i> value	Sign.
Native: response - LT1	2.239	0.324	6.907	<0.001***
Native: exposition - LT1	2.496	0.324	7.703	<0.001***
Native: narrative - LT1	2.115	0.324	6.525	<0.001***
Native: response - LT2	1.390	0.330	4.218	0.00161**
Native: exposition - LT2	1.648	0.330	5.001	<0.001***
Native: narrative - LT2	1.266	0.330	3.842	0.007**
Native: response - LT3	1.390	0.330	4.218	0.002**
Native: exposition - LT3	1.648	0.330	5.001	<0.001***
Native: narrative - LT3	1.266	0.330	3.842	0.007**
Native: response - LT5	1.081	0.327	3.309	0.037*
Native: exposition - LT5	1.339	0.327	4.098	0.003**
Native: response - LT6	1.299	0.327	3.976	0.004**
Native: exposition - LT6	1.557	0.327	4.766	<0.001***
Native: narrative - LT6	1.175	0.327	3.597	0.015*

P* < .05; *P* < .01; ****P* < .001.

Table 12. Significant group differences between learners and native speakers: adjectival clauses.

Null hypothesis	Estimate	Std. Error	<i>t</i> value	Sign.
LT4 - LT1	1.934	0.270	7.168	<0.001***
LT5 - LT1	2.022	0.248	8.142	<0.001***
LT6 - LT1	1.900	0.248	7.649	<0.001***
LT7 - LT1	0.967	0.252	3.840	<0.002**
LT4 - LT2	1.406	0.276	5.099	<0.001***
LT5 - LT2	1.494	0.255	5.857	<0.001***
LT6 - LT2	1.372	0.255	5.377	<0.001***
LT4 - LT3	1.240	0.276	4.496	<0.001***
LT5 - LT3	1.328	0.255	5.205	<0.001***
LT6 - LT3	1.205	0.255	4.725	<0.001***
LT7 - LT4	-0.967	0.276	-3.499	0.008**
LT7 - LT5	-1.055	0.255	-4.142	<0.001***
LT7 - LT6	-0.932	0.255	-3.661	0.005**

P* < .05; *P* < .01; ****P* < .001.

Table 13. Significant group differences between learner texts: adverbial clauses.

Null hypothesis	Estimate	Std. Error	<i>t</i> value	Sign.
Native: response - LT1	2.239	0.324	6.907	<0.01***
Native: exposition - LT1	0.250	3.241	7.703	<0.01***
Native: narrative - LT1	2.115	0.324	6.525	<0.01***
Native: response - LT2	1.390	0.330	4.218	<0.01**
Native: exposition - LT2	1.648	0.330	5.001	<0.01***
Native: narrative - LT2	1.266	0.330	3.842	<0.01**
Native: response - LT3	1.390	0.330	4.218	<0.01**
Native: exposition - LT3	1.648	0.3295	5.001	<0.01***
Native: narrative - LT3	1.266	0.330	3.842	<0.01**
Native: exposition - LT5	1.081	0.327	3.309	<0.037*
Native: response - LT5	1.339	0.327	4.098	<0.01**
Native: response - LT6	1.299	0.3267	3.976	<0.01**
Native: exposition - LT6	1.557	0.327	4.766	<0.01***
Native: narrative - LT6	1.175	0.327	3.597	<0.015*

Table 14. Significant group differences between learner and native speakers, texts: adverbial clauses.

NOTES

1. An exception to this rule is that the finite auxiliary *har/hade* 'have/had' before the supine in the perfect tense can be omitted, leaving the clause with only a nonfinite verb form:

Han var inte full, men han märkte att han (hade) druckit.
 He be.PST not drunk, but he notice.PST that he have.PST drunk
 'He wasn't drunk, but he could feel that he had been drinking.'

2. An exception to this rule is for instance when a clear contrast and stress is involved:

(i) Att inte statsministern begriper bättre är underligt
 That not prime.minister.DEF understand.PRS better be.PRS strange
 'That the prime minister doesn't understand better is strange.'

Here, we have a contrast between the prime minister, on the one hand, and 'other people,' on the other hand. It implies that it is understandable that people do not understand the matter in question, but the fact that THE PRIME MINISTER does not understand it is strange.

3. I am grateful to Victoria Johansson from Lund University for letting me use her data in the current study.
4. I am grateful to Per Holmberg from Gothenburg University, Sweden, for letting me use his data for the current study.
5. It is important to mention here that the learner texts are momentary impressions. Although the texts are written at different points in time, they do not necessarily reflect developmental order.
6. Interestingly, in L1 acquisition of Swedish, children also produce adjectival clauses as their first subordinate clauses (Håkansson & Hansson 2000).

7. In Teleman et al. (1999), *när* is considered to be a relative adverb that often lacks a referent, but in traditional grammar, *när* is often analyzed as a subordinating conjunction instead (See Teleman 1999, 6 Anm. Ord.).
8. I would like to thank one of the anonymous peer reviewers for pointing this out to me.

REFERENCES

- Bardovi-Harlig, K. & T. Bofman. 1989. Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition* 11, 17–34.
- Baten, K. & G. Håkansson. 2015. The Development of subordinate clauses in German and Swedish As L2S. *Studies in Second Language Acquisition* 37(03), 517–547.
- Bulté, B. & A. Housen. 2012. Defining and operationalising L2 complexity. In A. Housen, F. Kuiken & I. Vedder (eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*, 21–46. Amsterdam: Benjamins.
- Cooper, T. 1976. Measuring written syntactic patterns of second language learners of German. *Journal of Educational Research* 69, 176–183.
- Culicover, P. W. 2014. Constructions, complexity, and word order variation. In F. J. Newmeyer & L. B. Preston (eds.), *Measuring Grammatical Complexity*, 148–178. Oxford: Oxford University Press.
- Ellis, R. 2005. *Planning and Task Performance in a Second Language*, volume 11. Amsterdam: John Benjamins.
- Ellis, R. 2009. The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics* 30(4), 474–509.
- Flahive, D. & B. Snow. 1980. Measures of syntactic complexity in evaluating ESL compositions. In J. W. O. & K. Perkins (eds.), *Research in Language Testing*, 171–176. New York: Newbury House.
- Ganuza, N. 2011. Syntactic variation in the Swedish of adolescents in multilingual urban settings—a thesis summary. In R. Källström & I. Lindberg (eds.), *Young Urban Swedish. Variation and Change in Multilingual Settings* (Göteborgss ed.), 89–103. Gothenburg: University of Gothenburg.
- Glahn, E., G. Håkansson, B. Hammarberg, A. Holmen, A. Hvenekilde & K. Lund. 2001. Processability in Scandinavian second language acquisition. *Studies in Second Language Acquisition* 23(3), 389–416.
- Håkansson, G. & K. Hansson. 2000. Comprehension and production of relative clauses: a comparison between Swedish impaired and unimpaired children. *Journal of Child Language* 27(2), 313–333.
- Håkansson, G. & U. Nettelbladt. 1993. Developmental sequences in L1 and L2 acquisition of Swedish syntax. *International Journal of Applied Linguistics* 3, 3–29.
- Håkansson, G. & C. Norrby. 2010. Environmental influence on language acquisition: Comparing second and foreign language acquisition of Swedish. *Language Learning* 60(3), 628–650.
- Hammarberg, B. & Å. Viberg. 1977. The place-holder constraint, language typology and the teaching of Swedish to immigrants. *Studia Linguistica* 31, 106–163.
- Hauser, M. D., N. Chomsky, & W. Tecumseh Fitch. 2002. The faculty of language: what is it, who has it, and how did it evolve? *Science* 298(5598), 1569–1579.
- Holmberg, P. & K. Wirdenäs. 2010. Skrivpedagogik i praktiken. *Språk och stil* 20(1), 105–131.

- Hooper, J. & S. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4, 465–497.
- Housen, A. & F. Kuiken. 2009. Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics* 30(4), 461–473.
- Hyltenstam, K. 1977. Implicational patterns in interlanguage syntax variation. *Language Learning* 27, 383–411.
- Johansson, V. 2009. *Developmental Aspects of Text Production in Writing and Speech*. Lund.
- Kusters, W. 2003. *Linguistic Complexity: The Influence of Social Change on Verbal Inflection*. PhD thesis, Universiteit Leiden.
- Lambert, C. & J. Kormos. 2014. Complexity, accuracy, and fluency in task-based L2 research: toward more developmentally based measures of second language acquisition. *Applied Linguistics* 35(5), 607–614.
- Larson-Hall, J. 2010. *A Guide to Doing Statistics in Second Language Research Using R* (web editio ed.), volume 20. New York: Routledge.
- Lord, C. 2002. Are subordinate clauses more difficult? In *Complex Sentences in Grammar and Discourse: Essays in honor of Sandra A. Thompson*, 223–233. Philadelphia, PA: John Benjamins.
- Menn, L., C. Duffield, F. Newmeyer & L. Preston 2014. Looking for a ‘Gold Standard’ to measure language complexity: what psycholinguistics and neurolinguistics can (and cannot) offer to formal linguistics. *Measuring Grammatical Complexity*, 281–302. Oxford: Oxford University Press.
- Miestamo, M. 2008. Grammatical complexity in a cross-linguistic perspective. In M. Miestamo, K. Sinnemäki & F. Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, 23–41. Amsterdam: John Benjamins.
- Miestamo, M., K. Sinnemäki & F. Karlsson. (eds.). 2008. *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins.
- Monroe, J. 1975. Measuring and enhancing syntactic fluency in French. *The French Review* 48, 48, 1023–1031.
- Newmeyer, F. J. & L. B. Preston (eds.). 2014. *Measuring Grammatical Complexity*. Oxford: Oxford University Press.
- Nippold, M. A., J. M. Ward-Lonerger & J. L. Fanning. 2005. Persuasive writing in children, adolescents and adults: A study of syntactic, semantic and pragmatic development. *Language, Speech and Hearing Services in Schools* 36(2), 125–138.
- Norris, J. M. & L. Ortega. 2009. Towards an organic approach to investigating CAF in instructed SLA: the case of complexity. *Applied Linguistics*, 30(4), 555–578.
- Pallotti, G. 2009. CAF: defining, refining and differentiating constructs. *Applied Linguistics* 30(4), 590–601.
- Pallotti, G. 2014. A simple view of linguistic complexity. *Second Language Research* 31(1), 117–134.
- Perkins, K. 1980. Using objective methods of attained writing proficiency to discriminate among holistic evaluations. *TESOL Quarterly* 14, 61–69.
- Petersson, D. 2014. *The Highest Force Hypothesis Subordination in Swedish*. Lund: Lund University.
- Sampson, G., D. Gil & Trudgill, P. (eds.) 2009. *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press.
- Sangaran, J. 2005. The effects of focusing on meaning and form in strategic planning. In R. Ellis (ed.), *Planning and Task Performance in a Second Language*, 111–142. Amsterdam: John Benjamins.

- Skehan, P. & P. Foster. 1999. The influence of task structure and processing conditions on narrative retellings. *Language Learning* 49(1), 93–120.
- Teleman, U., S. Hellberg & E. Andersson. 1999. *Svenska Akademiens Grammatik*. Stockholm: Svenska Akademien.
- Trotzke, A. & J. Bayer. (eds.). 2015. *Syntactic Complexity across Interfaces*. Berlin/Boston/Munich: Walter de Gruyter.
- Wijers, M. Forthcoming. *The development of Complex Clause Constructions in Swedish as a Second Language*. PhD dissertation, Ghent University.