



UvA-DARE (Digital Academic Repository)

Syntactic complexity across proficiency and languages: L2 and L1 writing in Dutch, Italian and Spanish

Kuiken, F.; Vedder, I.

DOI

[10.1111/ijal.12256](https://doi.org/10.1111/ijal.12256)

Publication date

2019

Document Version

Final published version

Published in

International Journal of Applied Linguistics

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Kuiken, F., & Vedder, I. (2019). Syntactic complexity across proficiency and languages: L2 and L1 writing in Dutch, Italian and Spanish. *International Journal of Applied Linguistics*, 29(2), 192-210. <https://doi.org/10.1111/ijal.12256>

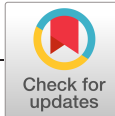
General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)



Syntactic complexity across proficiency and languages: L2 and L1 writing in Dutch, Italian and Spanish

Folkert Kuiken | Ineke Vedder

Amsterdam Center for Language and Communication, University of Amsterdam, VB, Amsterdam, The Netherlands

Correspondence

Folkert Kuiken, University of Amsterdam, Amsterdam Center for Language and Communication, Spuistraat 134, 1012 VB Amsterdam, The Netherlands.
Email: f.kuiken@uva.nl

This study explores to what extent syntactic complexity as assessed by four types of complexity measures may vary in written, argumentative texts of L2 learners (A2-B1) and native writers of Dutch, Italian and Spanish. All texts were assessed by calculating both overall complexity measures and more fine-grained measures that focus on the type and number of coordinate and subordinate structures, and the use of post-modifiers within the NP. The results of the study indicate that there is variation in the process of gradual complexification in written L2 production across proficiency levels, across languages, and between L2 and L1. In line with findings from earlier research, this study shows the importance of employing both general and more specific complexity measures for assessing syntactic growth in L2.

KEYWORDS

complexity by coordination, complexity by subordination, L2 writing, phrasal complexity, syntactic complexity

Questo studio esamina la variazione della complessità sintattica, misurata con quattro indici di complessità diversi, in un corpus di testi argomentativi prodotti da apprendenti L2 (A2-B1) e parlanti nativi di olandese, italiano e spagnolo. La valutazione della complessità sintattica è avvenuta in

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. International Journal of Applied Linguistics published by John Wiley & Sons Ltd.

base ad indici di complessità globali e più specifici, focalizzati sul tipo e sul numero di strutture sintattiche coordinate e subordinate e sull'impiego di post-modificatori nel sintagma nominale. Lo studio evidenzia come la complessità sintattica, intesa come il risultato del processo di graduale complessificazione della produzione scritta in L2, può variare attraverso livelli di competenza linguistica e lingue target differenti, e tra L2 e L1. In linea con i risultati emersi da ricerche precedenti, lo studio dimostra l'importanza di utilizzare sia indici di complessità generali che misure più specifiche per valutare la crescita sintattica in L2.

PAROLE CHIAVE

complessità per coordinazione, complessità per subordinazione, scrittura L2, complessità frasale, complessità sintattica

1 | INTRODUCTION

Studies of L2 writing have frequently employed general length-based performance measures (e.g. length of T-unit) that capture overall syntactic complexity in order to assess complexity growth across proficiency levels. As argued by Norris & Ortega (2009), it is, however, necessary to distinguish between various dimensions of syntactic complexity (i.e., complexity by subordination, complexity by coordination, and phrasal complexity), as these may be reflected in different ways by learners of different proficiency levels. In addition to general indices of syntactic complexity, researchers need to employ more developmentally sensitive complexity measures targeting different aspects of complexity that are thought to develop at different stages of learning.

As pointed out by Gyllstadt, Granfeldt, Bernardini, and Källkvist (2014), syntactic complexity may also vary across languages. Typologically different languages, such as Germanic and Romance languages, may differ in verbal and nominal complexity on a number of morphosyntactic constructions, both with respect to the VP (e.g., use of the past tense, or the conjunctive) or the NP (e.g., pre-modification versus post-modification). Languages may also differ in stylistic preferences towards particular morphosyntactic structures, such as the type and number of subordinate clauses, non-finite clauses, and the use of the gerund or determiners. In order to account for structural variation across languages, syntactic complexity should thus be assessed by particular measures, which focus on language-specific features (De Clercq, 2015; De Clercq & Housen, 2016; Gyllstadt et al., 2014; Hulstijn, Alderson, & Schoonen, 2010).

The overarching research question addressed in the present exploratory study is to what extent syntactic complexity, as a multi-layered construct, may vary in L2 writing across proficiency levels, between L2 and L1, and across different languages. To answer this question, we investigated syntactic complexity in a corpus of multiple argumentative essays written by L2 writers of Dutch, Italian and Spanish (two different essays written by each participant). The proficiency level of the L2 learners varied from A2 to B1 on the CEFR scale. The analysis of these three different interlanguages may be revelatory with regard to the influence of cross-linguistic differences on syntactic complexity.

In order to examine variation and regularities in L2 syntactic growth, we compared the L2 data of different proficiency levels with data of native speakers by submitting the same tasks to native writers of Dutch, Italian and Spanish. The comparison of learner production with native-speaker data may also shed light on potential trade-off

effects, both in L2 and L1, between complexity by subordination, coordination, and phrasal complexity (De Clercq, 2015; De Clercq & Housen, 2016; Gyllstadt et al., 2014).

In what follows we will first address the issue of the definition and operationalisation of syntactic complexity as a construct, and the ways in which complexity may vary across proficiency levels, languages, and between L2 and L1. We will then discuss the complexity indices that have often been employed in complexity research for assessing syntactic complexity by coordination, subordination, and at the phrasal level. Thereafter, we will describe the design and methodology of our study. Finally, we will present the main results of the study and discuss its implications and future perspectives.

2 | SYNTACTIC COMPLEXITY: CONCEPTUALISATION, OPERATIONALISATION AND MEASUREMENT

As has been emphasised in various studies, complexity is a multi-faceted construct that is conceptually different from related notions such as difficulty, maturity, development, and proficiency (Housen & Kuiken, 2009; Housen, Kuiken, & Vedder, 2012; Norris & Ortega, 2009; Pallotti, 2009). Pallotti (2014: 120) proposes a "simple" definition of complexity as a dependent variable reflecting language development, rather than as an independent variable influencing the difficulty or cognitive complexity of a task. In the present study, our definition of syntactic complexity follows the definition of "system complexity" by Bulté and Housen (2012, 2014), who define complexity as the number of discrete components that a language feature consists of and the number of connections (i.e., their systematic, organised relationships) between different components. Syntactic complexity, thus considered in objective and quantitative terms, is distinct from the broader and more subjective notion of difficulty, referring to the mental ease or difficulty with which particular linguistic items are learned and processed (Bulté & Housen, 2012: 23–24).

In many studies, syntactic complexity was assessed by means of general ratio measures employed for measuring overall complexity or clausal subordination. This was also the main outcome of a meta-study by Norris and Ortega (2000), and Bulté and Housen (2012), based on 16 and 40 task-based studies respectively. Although recently a number of studies have approached syntactic complexity in a more nuanced way (Biber & Gray, 2011; Biber, Gray, & Poonpon, 2011; Lu, 2011), length-based complexity measures or subordination measures have been adopted in the majority of complexity studies.

In the present study, we follow the recommendations of Norris and Ortega (2009) that syntactic complexity should be assessed by distinct and non-redundant measures that capture different dimensions of the construct. In order to account for the growth of syntactic complexity in L2 performance as a multi-layered and multi-dimensional construct, Norris and Ortega propose an organic and dynamic approach to complexity. In line with this conceptualisation, they suggest operationalising syntactic complexity as: (i) overall or general complexity; (ii) complexity by coordination; (iii) complexity by subordination; and (iv) complexity by phrasal elaboration.

3 | THE DEVELOPMENT OF SYNTACTIC COMPLEXITY: PATTERNS AND VARIATION

Ortega (2012) distinguishes three areas in which research into syntactic complexity has been conducted. The first area examines complexity for its value as an index of global L2 proficiency. Research that has been conducted within this framework focuses on the question of to what extent syntactic complexity increases as L2 learners become more proficient and on the type of indices that should be employed for the assessment of complexity. In this line of work, researchers sample second language users representing heterogeneous levels of linguistic ability and elicit from them extended discourse samples from which complexity measures are extracted. The second type of research, collocated within the framework of task-based language learning, attempts to describe L2 performance, rather than gauging L2 proficiency. The main issue at stake is whether certain manipulations of the cognitive demands of tasks can have

predictable effects on the three strands of language performance, i.e. complexity, accuracy, and fluency. The third area aims to measure the ability to benchmark: does the interlanguage complexity of production increase with growing grammatical development? Assuming the answer is affirmative, researchers can then search for the best complexity indices that can be used as benchmarks of broad interlanguage developmental stage. Our study is primarily situated in the first strand of research, focusing on patterns and variation in the process of gradual complexification of L2 production across proficiency levels, across languages, and between L2 and L1.

3.1 | Development across proficiency levels

In the SLA literature, syntactic complexity has been shown to relate strongly to the overall level of L2 proficiency (Ortega, 2003). As has been argued in various studies (Bardovi-Harlig, 1992; Norris & Ortega, 2009; Wolfe-Quintero, Inagaki, & Kim, 1998), syntactic complexity is thought to develop in three stages, in relation to the global increase of L2 proficiency (see Table 1). First, sentences or clauses are produced as independent, uncoordinated utterances (e.g., "I have a son. He is 12 years old"; stage 0). These utterances are then linked by coordination (stage 1), then by subordination (stage 2), and finally by complexification of the noun phrase (stage 3). In this final stage, the pre-modifying phrase "12- year-old" is embedded into the main clause (Bardovi-Harlig, 1992; Norris & Ortega, 2009). Syntactic complexity in L2 is thus thought to expand from coordination to subordination to phrasal elaboration, as learners gain proficiency (Wolfe-Quintero et al., 1998).

At beginner and low-intermediate proficiency levels (see Table 1, stage 1), syntactic growth may show an increase of coordination (Bardovi-Harlig, 1992; Vyatkina, 2012). Upper-intermediate levels are thought to display an increase in subordinate structures (stage 2). More advanced proficiency levels may be characterised by subclausal complexification at the phrasal level, which is supposed to be characteristic of academic discourse and written prose (Biber et al., 2011; Biber & Gray, 2011; Norris & Ortega, 2009).

Previous research has emphasised that the three stages may partially overlap and that syntactic complexity may decrease again at higher proficiency levels, as a result of task and genre effects, (Lu, 2011; Ortega, 2003; Pallotti, 2009; Polio & Yoon, this issue). Wolfe-Quintero et al. (1998) suggested that complexity measures might exhibit "omega-shaped" patterns, with an increase in complexity, followed by a decline at the higher proficiency levels. As argued by Ortega (2003), "more complex" does not necessarily mean "better," since a higher or lower complexity rate may also be determined by personal and stylistic choices, providing an indication of higher L2 proficiency. Many studies have also highlighted the high degree of variability between learners and the non-linearity of the individual developmental trajectories that learners follow (Bulté & Housen, this issue; Larsen-Freeman, 2006; Spoelman & Verspoor, 2010; Verspoor, Lowie, & Van Dijk, 2008).

3.2 | Variation across languages

Contrary to L1 where typologically oriented complexity research has been a major theme (Dahl, 2004; Miestamo, Sinnemäki, & Karlsson, 2008), research on syntactic development in L2 has rarely been considered from a cross-linguistic perspective (De Clercq & Housen, 2016; Gyllstadt et al., 2014). As pointed out by Granfeldt and

TABLE 1 Examples of stages of syntactic development in L2

Stage 0: I have a son. He is 12 years old.

Stage 1: I have a son and he is 12 years old.

Stage 2: I have a son who is 12 years old.

Stage 3: My 12-year-old son.

Bernardini (this issue), cross-linguistic variation in syntactic complexity may be caused by structural differences between target languages or by different stylistic preferences. In comparative studies on L1 Swedish, L2 English, L3 French, and L4 Italian, it was observed that closely related languages display morphosyntactic variation that should be taken into consideration in cross-linguistic comparisons (Gyllstadt et al., 2014; Granfeldt & Bernardini, this issue). Differences with respect to basic structural properties (i.e., complementiser omission in English and Swedish; null-subject, use of pre-possessive determiner articles in Italian) may have an effect on measures of syntactic complexity and be a potential source of cross-linguistic variation. De Clercq (2015) and De Clercq and Housen (2016) examined the validity of a number of morphosyntactic and lexical complexity measures as cross-linguistic indicators of proficiency. In order to do so, they analysed 100 L2 English and 100 L2 French oral narratives from adolescent L1 speakers of Dutch, who were at four proficiency levels (beginner-advanced), as well as L1 benchmark data from each language. The results revealed a gradual process of structural elaboration and syntactic diversification in both learner groups, while, especially in French, considerable differences were found between learners and native speakers with respect to the distribution of specific clause types.

The study at hand aims to examine syntactic variation in three different target languages: a Germanic language (Dutch L2) and two Romance languages (Italian L2 and Spanish L2). Let's consider some examples in which Germanic and Romance languages differ with respect to morphosyntactic structures, both in the verbal group and in the nominal group. In the verbal group, Romance languages seem to have a preference for particular types of subordinate clauses introduced by a gerund or past participle in the verbal group; in Germanic languages, these structures are less common and therefore less frequent. In the nominal group, Germanic languages prefer pre-modification where post-modification is used in Romance languages (cf. Dutch "een geel boek" and English "a yellow book" versus Italian "un libro giallo" and Spanish "un libro amarillo." These differences may be reflected particularly by stylistic preferences in written prose and may lead to variation in overall complexity (i.e., mean length of T-unit), complexity by coordination (i.e., number and type of coordinate clauses per T-unit), complexity by subordination (i.e., number and type of subordinate clauses per T-unit), and phrasal complexity (i.e., number and type of pre or post-modifying NP's; mean length of clause). Given the paucity of studies that have been conducted for L2, it is, however, unclear to what extent variation in complexity due to differences in stylistic preferences can be observed between L2 writers of Dutch and L2 writers of Italian and Spanish.

Despite a high degree of typological similarity of the Romance languages, closely-related languages such as Italian and Spanish may also vary in syntactic complexity. As demonstrated by Della Putta (2016a, 2016b), the two languages differ with respect to particular syntactic structures, both in the VP and the NP. Structural differences leading to differences in overall syntactic complexity can be detected in the frequency of the planned future periphrasis and the iterative periphrasis in Spanish ("ir a," go to, + infinitive; "volver a," return to, + infinitive), which in Italian are usually expressed by affixation, lexical means, and/or verbal morphology ("rivedere"/"vedere di nuovo," to meet again). Examples within the NP of typologically different structures concern the occurrence of the pre-possessive determiner article in Italian contrary to Spanish ("la mia macchina," my car, in Italian versus "mi coche" in Spanish), or the use of the prepositional accusative in Spanish in SVO sentences containing an animate and specific direct object, which in standard Italian are ungrammatical (Spanish: "espero a Franco," I am waiting for Franco versus Italian: "*aspetto a Franco"). It remains to be seen, however, to what degree these structural differences between Italian and Spanish are reflected by differences in syntactic complexity between the two languages, as established by the complexity indices employed in the study (see the section on the methodology of the study, below).

3.3 | Variation between L2 and L1

Compared to native speakers, the performance of L2 learners at beginner and intermediate proficiency levels is generally characterised by lower overall syntactic complexity (Wolfe-Quintero et al., 1998). As observed by Véronique (2004) in a study on the acquisition of discourse in oral French, native speakers employed a higher number

of subordinate structures compared to beginner and intermediate Moroccan L2 learners. Similar differences in syntactic complexity were demonstrated between L1 and L2 speakers by Bartning and Schlyter (2004) in a study on the development of syntax in oral speech. Compared to L2 French, speakers in L1 Swedish used more complement clauses, adverbial clauses, and, especially, relative clauses.

As already mentioned in the previous section, a decline of syntactic complexity has often been observed at more advanced levels of L2 proficiency, as higher proficient L2 speakers may adapt and fine-tune the complexity of their utterances to the pragmatic requirements of task or genre, which is similar to native speakers. This was one of the outcomes of a study by Pallotti (2009), where the non-native participants—who undertook a film retelling task and a phone task to a shopkeeper or a travel agency—tended to produce long and complex syntactic units. In the phone task, the L2 learners sounded pragmatically unnatural compared to the native speakers, who decomposed their utterances into syntactically simpler and shorter units, adapting the complexity of their speech to the specific task requirements. Over time, the L2 participants approached native speakers' behaviour by increasing syntactic complexity in the retelling task and decreasing it in phone calls (Ferrari & Pallotti, 2008; Pallotti, 2009).

Task and genre effects in relation to syntactic complexity in L2 and L1 were also observed by Lu (2011) and Polio and Yoon (this issue). Both native and non-native writers exhibited greater syntactic complexity in argumentative essays than in narratives, but the differences between the two genres were less pronounced for the native writers. As hypothesised by Polio and Yoon, the larger genre differences among the non-native participants may be related to more focused attention to language in the argumentative condition or to more explicit genre awareness among the L2-learners.

Kormos (2011) found little difference in subordination rates between native and non-native speakers. According to Lambert and Kormos (2014), expert and more advanced L2 speakers and writers, similar to native speakers, seem to express complex ideas more simply than novices. Simplification of syntax can thus be due to higher L2 proficiency—rather than a lack of linguistic resources—and to a mastery of linguistic and pragmatic skills in efficient and effective message formation (Lambert & Kormos, 2014; Lambert & Nakamura, this issue).

4 | STUDIES ON MEASURES OF SYNTACTIC COMPLEXITY: COORDINATION, SUBORDINATION, AND PHRASAL COMPLEXITY

In this section, we review some of the complexity indices that have been employed in the literature for measuring complexity by coordination, subordination, and phrasal complexity.

4.1 | Coordination

In studies which assess complexity by coordination, the Coordination Index (CI) proposed by Bardovi-Harlig (1992) has often been employed. The CI is calculated by dividing the total number of independent clauses in a language sample by the total number of combined clauses (i.e., both coordinate and subordinate clauses). As pointed out by Bulté and Housen (2012: 38), a potential problem of the CI is that it is a hybrid measure of coordination, because the coordination rate measured by the CI depends not only on the number of coordinate structures but also on the amount of subordination that is produced. For that reason, inspired by Hudson (1976), Progovac (1998) and Osborne (2006), we measured coordination in an alternative way, focusing on the position of the coordination conjunction (see for more details the upcoming section on Measures and data analysis).

The results obtained by studies on L2 English in which the CI was employed, so far, have been contradictory. Bardovi-Harlig (1992) observed a decrease in CI with an increase in proficiency level. Lu (2011) and Ai and Lu (2013), instead, found no significant differences between L2 proficiency levels. In line with the latter findings, Neary-Sundquist (2016) found no significant differences in the rate of coordination between intermediate,

advanced and highly advanced L2 learners. Despite this lack of significant differences between proficiency levels, the outcomes of the study by Neary-Sundquist appeared to confirm the developmental “omega-shaped” pattern suggested by Wolfe-Quintero et al. (1998), which is to say a decrease of coordination at higher proficiency levels in favour of subordination and, similarly, a decline of subordination in favour of phrasal complexity at the superior levels.

4.2 | Subordination

As pointed out by Norris and Ortega (2009), complexity by subordination is usually measured as the number of clauses or subclauses per T-unit (Hunt, 1965), AS-unit (Foster, Tonkyn, & Wigglesworth, 2000), or C-unit (Loban, 1976). Although some studies have indicated a rise of subordination rates at intermediate levels and a levelling-off at higher proficiency level (Bardovi-Harlig & Bofman, 1989), the majority of studies on complexity by subordination have shown mixed outcomes. Hirano (1991) found that the ratio of dependent clauses to total clauses increased with proficiency level, but no relation could be established between the number of dependent clauses per clause and the proficiency level of the learners. Contrary to Hirano's results, Neary-Sundquist (2016) detected significant differences in complexity by subordination between the intermediate, advanced, and highly advanced L2 participants of the study.

Bulté and Housen (2012) have argued that subordination measures that address the sentential level but not the clausal and phrasal level may have their limitations as indices of overall syntactic development, due to their fairly narrow linguistic scope. Moreover, as demonstrated by Biber and Gray (2011) and Biber et al. (2011), clausal subordination is much more common in oral discourse compared to written discourse, particularly academic writing. For this reason, some researchers have questioned the suitability of subordination ratios as indicators of syntactic growth (Bulté & Housen, 2012; Ellis & Barkhuizen, 2005; Verspoor et al., 2008). In the present study we have adopted the measures suggested by Wolfe-Quintero et al. (1998), Norris and Ortega (2009), and Bulté and Housen (2012).

Finally, an acquisitional sequence concerning the development of particular types of subordinate clauses has been proposed in some studies, such as the one developed for oral L2 French by Bartning and Schlyter (2004). Although our study is not longitudinal in nature, it may be interesting to examine tentatively whether such a sequence in which adverbial clauses are used first, followed by complement clauses, and then relative clauses can also be detected in the texts of our learners with different proficiency levels.

4.3 | Phrasal complexity

For the assessment of phrasal complexity, Norris and Ortega (2009) suggest measuring the mean number of words per clause (mean length of clause, MLC). The assumption underlying this recommendation is that increases in clause length (by means of pre or post-modification of the head) will reflect increases in phrase length. Their suggestion to adopt the MLC as a measure of phrasal complexity has been followed in a number of studies (Gyllstadt et al., 2014; Neary-Sundquist, 2016). As pointed out by Bulté and Housen (2012: 38–39), clause length as the indicator of phrasal complexity is problematic for various reasons. Since clause length may increase through expansion at the phrasal level as well as at the clausal level (by adding adjuncts of time, manner, place), the MLC cannot be considered a “pure” measure of phrasal complexity. Secondly, clause length crucially depends on the definition and operationalisation of clauses. Across studies there has been a lack of agreement with respect to the definition of a clause, leading to different findings and interpretations. Verb clusters such as “start to look” and “begin shouting,” for example, are either analysed as consisting of two clauses (a main clause plus a nonfinite subordinate clause), or as one single main clause, resulting in a lower or higher MLC score (Bulté & Housen, 2012: 40). In addition to the MLC, alternative indices of noun phrase complexity have been employed in a number of studies to address various features that may determine phrasal complexity. An example of an alternative indice of noun phrase complexity has been

proposed by Ravid and Berman (2010), which takes into account length of noun phrases, number of noun phrase modifier tokens, number of modifier types, number of subordinated lexical nouns, and abstractness of lexical head nouns (see also Lambert & Nakamura, this issue).

5 | DESIGN AND METHODOLOGY OF THE STUDY

5.1 | Research questions and hypotheses

The main research focus of this study concerns the extent to which syntactic complexity, operationalised as overall syntactic complexity, complexity by coordination, subordination, and phrasal complexity may vary: (i) across proficiency levels; and (ii) across languages and between L2 and L1. In order to answer this overarching question, the following research questions have been formulated:

- RQ1: To what extent does syntactic complexity as assessed by four types of complexity measures vary in written texts of L2 learners of Dutch, Italian and Spanish in relation to their proficiency level?
- RQ2: Which differences can be detected between L2 and L1 writers of Dutch, Italian and Spanish respectively with regard to the syntactic complexity of their texts?

On the basis of the findings of the studies that have been discussed above (e.g., Bardovi-Harlig, 1992; Bartning & Schlyter, 2004; Neary-Sundquist, 2016; Norris & Ortega, 2009; Ortega, 2003; Pallotti, 2009; Véronique, 2004; Wolfe-Quintero et al., 1998), regarding our first research question (RQ1) we expect to find significant correlations between overall syntactic complexity and L2 proficiency. As hypothesised by Norris and Ortega (2009), an increase of L2 proficiency should result in learners using more coordinate and subordinate clauses at intermediate levels of L2 proficiency, but less at more advanced levels. Following Bartning and Schlyter (2004) in their study on syntactic development in oral French, we may find the following acquisitional sequence in the use of subordinate clauses: first adverbials, followed by complements, and then relatives. The more advanced proficiency levels, where phrasal complexity has been found to develop, will probably be characterised by an increase in the number and length of post-modifying NPs (Norris & Ortega, 2009).

Our second research question concerns the comparison of syntactic complexity in written L2 and L1 production. Based on Ferrari and Pallotti (2008), Pallotti (2009), Kormos (2011), Lambert and Kormos (2014), and Neary-Sundquist (2016), our hypothesis is that L2 learners will use more coordination, but less subordination and fewer post-modifying noun phrases than native speakers.

It is uncertain to what extent it will be possible to observe cross-linguistic variation in syntactic complexity between Dutch, Italian and Spanish. In the texts written by the native speakers, a number of differences in overall complexity, complexity by coordination, subordination, and phrasal complexity between Dutch versus Italian and Spanish may, however, be detected. Possible differences might, for instance, be observed in the use of particular types of subordinate clauses.

5.2 | Participants and tasks

The analyses were carried out on the basis of data from 32 young adult L2 learners of Dutch (average age 22.2 years), 39 of Italian (average age 21.8 years), and 23 of Spanish (average age 23.9 years). All learners were students at a Dutch university, either preparing for university (in the case of the L2 learners of Dutch) or already registered as a bachelor student (in the case of the Italian L2 and Spanish L2 learners). The proficiency level of the participants ranged from A2 to B1 level of the CEFR; all learners were enrolled in a course preparing for B2 level. The students of Italian and Spanish all had Dutch as their mother tongue, whereas the Dutch L2 learners came

from 27 different cultural and language backgrounds. All participants were submitted on the same day to two writing tasks during class, comprising a short argumentative text. We opted for two argumentative tasks with similar task requirements, as research has shown that one single task may not yield a representative picture of a learner's writing proficiency (Schoonen, 1991). The time available for each task was 35 minutes, which turned out to be sufficient to complete the task. In the first task, learners were required to make a decision about which of three non-governmental organisations to choose as a candidate for receiving a grant, whereas in the second task they had to decide which topic they would like to see published in the monthly magazine of their favourite newspaper (for an example of the prompt of the latter task, see the Appendix). In order to assess the general proficiency level in the target language of the participants in a similar way, all participants were submitted to a C-test (for a theoretical account of the C-test, see Klein-Braley, 2002). In the C-test, learners were asked to complete 100 words in five short texts in which the second part of every other word had been replaced by blanks. The test versions for Dutch, Italian and Spanish differed for the three languages, but were based on the same type of texts, i.e., short news items from newspapers and magazines. Test scores varied for Dutch L2 from 66 to 95 (mean 79.1, SD 12.3), for Italian L2 from 46 to 93 (mean 69.9, SD 23.2), and for Spanish L2 from 71 to 97 (mean 83.3, SD 14.3).

In order to examine regularities and variation in syntactic growth in L2 compared to L1, as a benchmark the same writing tasks were administered to a group of 17 native speakers of Dutch (average age 23.9 years), 18 of Italian (average age 23.8 years) and 10 of Spanish (average age 24.0 years). As expected, they all obtained (very) high mean scores on the C-test (Dutch L1: 97.9; Italian L1: 96.2; Spanish L1: 96.9), implying a ceiling effect.

5.3 | Measures and data analysis

In our study, syntactic complexity was assessed by means of overall ratio measures and more specific measures for assessing complexity by coordination, subordination, and phrasal complexity. As measures of overall complexity, we employed mean number of clauses per T-unit and number of dependent clauses per clause. In doing so, we follow Wolfe-Quintero et al. (1998) and Norris and Ortega (2009), although we are aware that the number of dependent clauses per clause in particular has also been interpreted as an index of subordination by others (e.g., De Clercq, 2015; De Clercq & Housen, 2016).

For the assessment of complexity by coordination, we chose to adopt a specific coordination measure focusing on the number and type of coordinate structures instead of the Coordination Index (CI). Following Hudson (1976), Progovac (1998) and Osborne (2006), we distinguished three types of coordinated structures: (i) coordination between T-units ("Paul sings in a choir *and* Mark plays the guitar"); (ii) coordination within a T-unit ("Children should eat less *and* exercise more"); and (iii) coordination between constituents ("Physical exercise can reduce feelings of loneliness *and* stress"). For these three measures, mean scores were calculated, i.e., number of coordination type per 100 words.

Regarding subordination, we distinguished three types of subordinate structures: (i) complement clauses; (ii) adverbial clauses; and (iii) relative clauses. A further distinction was made within the three types of subordinate clauses. Based on Noonan (1985), complement clauses were subdivided into implicit complement clauses containing a non-finite verb ("you may laugh") and explicit complement clauses containing a finite verb ("I hope (that) he will come"). Adverbial clauses were subdivided into temporal, causal, consecutive, thematic, hypothetical, and concessive clauses (Hengeveld, 1998). Relative clauses were classified as related to subject, direct object, indirect object, oblique, and genitive (Keenan & Comrie, 1977). Again, for all these three measures mean scores were calculated, i.e., number of coordination type per 100 words. As the occurrences of these subtypes within the three types of subordinate clauses turned out to be very low, we decided not to take them into account in our further analyses.

In line with Bulté and Housen (2012), for phrasal complexity, we decided to use a specific measure focusing on post-modification, particularly the number and mean length of post-modifying noun phrases, instead of Mean Length

of Clause (MLC). Phrasal complexity was thus calculated by counting both the number of post-modifying noun phrases per 100 words and the mean length of post-modifying noun phrases. In a sentence like “Also the 1500 delegates from different countries all over the world, who have come together in Chicago [...],” “delegates”—preceded by the pre-modifiers “the” and “1500”—is followed by the post-modifying 13-word phrase, “from different countries all over the world, who have come together in Chicago,” referring to the head of the nominal group “delegates” (cf. Van de Velde, 2009). In this particular example, “from different countries all over the world” and “who have come together in Chicago” would be counted as two separate post-modifying noun phrases.

Data were scored by three native speakers of Dutch, Italian and Spanish respectively. They received information on the scoring rubrics and for each of the three languages, raters started with a sample of five texts, compared and discussed their outcomes until a percentage of agreement of at least 90 was reached. In order to answer our research questions, tasks 1 and 2 were aggregated by averaging the scores of both tasks. Examination of histograms, QQ-plots, and Shapiro–Wilk test statistics indicated that nearly all outcome variables violated the assumptions of Pearson correlation, ANOVA, and t-tests, particularly the assumption of normality; there was a high degree of right skewness. Transformations and alternative model fitting were unsuccessful in correcting assumption violations. Therefore, all outcome variables were examined non-parametrically. A Kruskal–Wallis test was used to evaluate differences between all groups and Mann–Whitney U post-hoc tests to compare L2 and L1, both overall and within languages. Associations between the outcome domains and the C-test score were evaluated by calculating Spearman's Rho. To account for the multiple comparisons being made in the post-hoc comparisons, a Bonferroni correction was applied to alpha (0.05/40), adopting a significance threshold of 0.00125.

6 | RESULTS

In Table 2, the descriptive statistics for the four types of syntactic measures are presented. The measures on overall complexity indicate that the three groups of L2 learners produce T-units which consist generally of two clauses (range: 1.840–1.923), while roughly every second clause is a dependent clause (range: 0.447–0.458). Coordinating devices are used by all three groups of L2 learners, although the mean score for coordination between T-units is higher for Dutch L2 learners than for Italian L2 and Spanish L2 learners, while Italian L2 learners obtain a higher score for coordination within T-units and Spanish L2 learners for coordination between constituents. With respect to complexity by subordination, Dutch L2 learners score higher on complement clauses, while this is the case for Spanish L2 learners on adverbial and relative clauses. Regarding phrasal complexity, the highest number of post-modifying NP's is obtained by Dutch L2 learners, whereas Spanish L2 learners attain the highest score for the length of post-modifying NP's.

Our first research question regards both the relationship between syntactic complexity and proficiency level and cross-linguistic differences in complexity between the three languages. In order to answer this question, scores on the four types of syntactic measures were correlated by means of Spearman's Rho with the language proficiency of the L2 learners as measured by a C-test (see Table 3). On the whole, most outcome domains were moderately associated positively with C-test scores, although only a few correlations turned out to be significant. For Italian L2 speakers, significant correlations are observed between the C-test on the one hand and coordination within T-units, relative clauses, and length of post-modifying noun phrases per 100 words on the other hand. No significant associations were found for Dutch and Spanish.

Our second research question concerned differences between scores obtained on the four types of syntactic measures by L2 learners as compared to native speakers. Differences between L1 and L2 were detected by means of an independent-samples Mann–Whitney U test per language (see Table 4). For Dutch, L1 and L2 did not differ significantly on any of the outcome variables. For Italian, L1 and L2 speakers differed significantly on number of clauses per T-unit, dependent clauses per clause, and mean length of post-modifying noun phrases. For Spanish, there was a significant difference between L1 and L2 for clauses per T-unit, coordination within T-units, and relative clauses.

TABLE 2 Mean scores (and standard deviations) on four types of syntactic measures

L1/L2	Overall complexity		Complexity by coordination			Complexity by subordination			Phrasal complexity	
	Clauses/T-unit	Dep. clauses/ clause	Betw. T-units	Within T-unit	Betw. constituents	Complement	Adverbial	Relative	Postm. NP's numb.	Postm. NP's length
Dutch L1 (n = 17)	1.938 (.208)	0.472 (0.058)	0.941 (0.827)	3.059 (1.391)	1.000 (0.685)	8.794 (3.575)	4.294 (2.194)	2.147 (1.057)	5.1364 (1.053)	5.030 (0.988)
Dutch L2 (n = 32)	1.923 (0.317)	0.458 (0.086)	1.438 (1.176)	2.313 (1.839)	1.391 (0.914)	10.203 (3.786)	3.188 (1.496)	2.078 (1.327)	4.887 (1.352)	4.606 (1.219)
Italian L1 (n = 18)	2.984 (0.857)	0.624 (0.078)	0.694 (0.645)	4.889 (2.447)	1.361 (0.921)	6.333 (2.093)	2.222 (1.374)	3.333 (1.404)	5.645 (1.455)	5.492 (1.586)
Italian L2 (n = 39)	1.878 (0.354)	0.451 (0.097)	0.680 (0.877)	3.795 (2.397)	1.051 (1.063)	7.346 (2.033)	3.577 (1.567)	2.218 (1.464)	4.142 (1.206)	3.927 (0.923)
Spanish L1 (n = 10)	2.3975 (0.558)	0.557 (0.091)	1.550 (1.343)	9.600 (5.317)	3.100 (1.745)	10.050 (4.038)	5.900 (2.980)	5.350 (2.148)	4.745 (0.673)	5.965 (0.916)
Spanish L2 (n = 23)	1.840 (0.200)	0.447 (0.059)	0.522 (0.574)	2.696 (1.579)	1.587 (.961)	9.196 (3.579)	3.652 (1.480)	3.130 (1.200)	4.553 (1.195)	4.941 (1.111)

TABLE 3 Spearman's Rho correlations (and significance level) of four types of syntactic measures with language proficiency (C-test)

L2	Overall complexity		Complexity by coordination			Complexity by subordination			Phrasal complexity	
	Clauses/ T-unit	Dep. clauses/ clause	Betw. T-units	Within T-unit	Betw. constituents	Complement	Adverbial	Relative	Postm. NP's numb.	Postm. NP's length
Dutch L2 (n = 32)	0.334 (0.062)	0.337 (0.060)	0.008 (0.963)	0.271 (0.133)	0.115 (0.531)	0.095 (0.607)	0.115 (0.529)	0.310 (0.084)	-0.023 (0.900)	0.239 (0.188)
Italian L2 (n = 39)	0.275 (0.090)	0.293 (0.070)	-0.110 (0.506)	0.472** (0.002)	0.087 (0.597)	0.310 (0.055)	-0.096 (0.561)	0.559** (0.000)	0.264 (0.105)	0.397* (0.012)
Spanish L2 (n = 23)	-0.038 (0.864)	-0.044 (0.841)	-0.022 (0.920)	0.119 (0.588)	0.359 (0.092)	0.070 (0.752)	0.173 (0.430)	-0.096 (0.662)	0.025 (0.909)	0.131 (0.552)

*p < 0.05, **p < 0.01.

TABLE 4 Mann–Whitney U test results for differences between L1 and L2

	Overall complex.		Complexity by coordination			Complexity by subordination			Phrasal complexity	
	Clauses/T-unit	Depend. clauses/clause	Betw. T-units	Within T-unit	Betw. constituents	Complement	Adverbial	Relative	Postm. NP's numb.	Postm. NP's length
Mann–Whitney U test results for differences between L1 and L2, Dutch										
U	264.5	245.5	332.0	175.5	342.5	344.0	181.0	266.5	242.0	200.5
SE	47.6	47.6	46.9	46.1	46.4	47.5	47.4	47.3	47.6	47.6
<i>p</i>	0.875	0.577	0.201	0.036	0.129	0.130	0.055	0.907	0.529	0.133
Mann–Whitney U test results for differences between L1 and L2, Italian										
U	52.2	60.5	307.5	255.5	271.0	470.5	523.5	190.0	154.0	123.0
SE	58.2	58.2	55.6	57.6	57.1	58.0	57.9	57.9	58.2	58.5
<i>p</i>	<0.001*	<0.001*	0.434	0.097	0.162	0.039	0.003	0.005	0.001	<0.001*
Mann–Whitney U test results for differences between L1 and L2, Spanish										
U	33.0	37.0	58.0	5.5	58.5	101.5	57.0	32.5	101.0	51.5
SE	25.5	25.5	24.7	25.3	25.1	25.5	25.4	25.3	25.5	25.5
<i>p</i>	0.001*	0.002	0.025	<0.001*	0.025	0.603	0.022	0.001*	0.603	0.011

*Significant at alpha = 0.00125.

TABLE 5 Summary: correlations measures of complexity with language proficiency (C-test)

L2	Overall complex.		Complexity by coordination			Complexity by subordination			Phrasal complexity	
	Clauses/T-unit	Depend. clauses/ clause	Betw. T- units	Within T-unit	Betw. constituents	Complement	Adverbial	Relative	Postm. NP's numb.	Postm. NP's length
Dutch										
Italian				X				X		X
Spanish										

X = significant correlation

TABLE 6 Summary: comparison L2-L1

L2	Overall complexity		Complexity by coordination			Complexity by subordination			Phrasal complexity	
	Clauses/T-unit	Depend. clauses/ Clause	Betw. T- units	Within T-unit	Betw. constituents	Complement	Adverbial	Relative	Postm. NP's numb.	Postm. NP's length
Dutch										
Italian	X									X
Spanish	X			X				X		

X = significant difference

7 | CONCLUSION AND DISCUSSION

Table 5 summarises the correlations of the four types of syntactic measures employed in the study, with language proficiency as measured by a C-test. With regard to our first research question concerning the relationship between syntactic complexity and L2 proficiency level, we found significant correlations for learners of Italian L2 for coordination within T-units, use of relative clauses, and length of post-modifying NP's. These findings (partly) confirm our hypothesis that with an increase of L2, proficiency learners will use more coordinate and subordinate clauses. One should keep in mind, however, that the Italian L2 group was the largest group with, therefore, the largest statistical power to detect associations. Significant correlations for Dutch L2 and Spanish L2 were not found. A possible—albeit, because of the cross-sectional nature of our data, tentative—acquisitional sequence of subordinating clauses as hypothesised by Bartning and Schlyter (2004)—first adverbials, then complement clauses, followed by relatives—could not be confirmed. Only for Italian we found that more proficient L2 learners used more relatives (and not more adverbial or complement clauses). At the phrasal level, the more proficient learners of Italian L2 produced longer post-modifying NPs, as expected. This was not the case, however, for Dutch and Spanish L2 learners.

This different result for Dutch (and Spanish) compared to Italian may have to do with differences in language background of our participants: all Italian (and Spanish) L2 learners were native speakers of Dutch, whereas the Dutch L2 learners had 27 different mother tongues. As some of these 27 languages were typologically very different from Dutch, this may have influenced the texts of the L2 writers. It might also be due to typological differences between Romance and Germanic languages: whereas in Italian elaboration of the noun phrase is mainly achieved by means of post-modification (e.g., post-positive adjective following the noun), in Dutch pre-modification is generally required (adjective placed in front of the noun). However, given the typological similarity of Italian and Spanish, the question remains why, contrary to Italian, no significant correlation between subordination and phrasal complexity on the one hand and proficiency level on the other could be established for the Spanish L2 learners. Further research is needed in order to shed more light on this issue.

With respect to our second research question regarding possible differences between L2 and L1 writers, we hypothesised that L2 learners would score higher on coordinate structures, but lower on subordinate structures and post-modifying phrases. This hypothesis was confirmed for Italian concerning the length of post-modifying phrases and for Spanish with respect to the use of relative clauses (see Table 6 for an overview of significant differences between L1 and L2). In these cases, L1 writers outperformed the L2 learners. Contrary to our expectation, Spanish L1 writers used more coordinating devices within T-units than L2 learners. No significant differences between L1 and L2 were found for Dutch. Due to this non-normal distribution of scores obtained, it was necessary to use nonparametric testing, which may have led to less precise estimates of the differences between L1 and L2 speakers.

Given these different findings for the three target languages involved in the study, further research is necessary in order to unravel whether different syntactic-rhetorical patterns or pragmatic and stylistic choices might lay at the base of these cross-linguistics differences, both between Germanic and Romance languages (i.e., Dutch versus Italian and Spanish, and between different Romance languages (i.e., Italian versus Spanish).

Concerning the main issue discussed in this chapter—that is, the extent to which syntactic complexity may vary—our study seems to indicate that there is indeed variation in the process of gradual complexification in written L2 production across proficiency levels (A2–B1), across languages (Dutch, Italian, Spanish), and between L2 and L1. What our study does ascertain is the importance of distinguishing different types of complexity, as more fine-tuned measures appear to lead to specific findings that cannot be demonstrated by the use of more general, overall measures. Had we not distinguished various types of coordination, subordination, and phrasal complexity, we would not have been able to assess that more proficient learners of Italian L2 used more coordination within T-units, relative clauses, and longer post-modifying phrases than their less proficient peers. The results of our study also underscore the importance of employing both general and more specific measures for the investigation of syntactic complexity across proficiency levels and across languages.

Our exploratory study has, however, a number of limitations. It should be kept in mind that the data have been collected cross-sectionally, that the number of learners participating in the study was limited, and that their proficiency levels in L2 covered a limited range. The latter limitation was evident in the distribution of scores, which showed a high degree of right skewness. Due to the (lack of) correlations found between various measures of complexity and language proficiency as measured by a C-test, it is not possible to discern a clear developmental pattern based on our results, just as no conclusive answer can be given as yet to the question of to what extent syntactic complexity may act as an indicator of L2 proficiency or the other way around. A longitudinal design could possibly shed more light on this issue. For these reasons, it may be too early to conclude what our findings imply for the practice of language teaching, including how different teaching approaches might affect complexity scores in different ways.

Further research with more participants and a higher range of proficiency levels might result in more and/or other types of variation as demonstrated by the present study. A more elaborate, large scale study might also allow for a more detailed analysis into subtypes of, for instance, complement, adverbial, and relative clauses. Other issues to be investigated in future research might be: (i) whether and to what extent trade-offs exist between complexity by coordination, subordination, and phrasal complexity; (ii) whether differences might be observed in the structure of the noun phrase (e.g., in Dutch, attributive adjectives are usually employed as pre-modifiers, whereas in Italian and Spanish attributive adjectives are more often placed after the noun); (iii) whether it is possible to characterise stages of L2 acquisition in terms of type and number of syntactic structures; and (iv) the extent to which syntactic complexity can be regarded as a (partly?) language specific construct.

ORCID

Folkert Kuiken  <https://orcid.org/0000-0002-1667-2501>

Ineke Vedder  <https://orcid.org/0000-0002-2677-0228>

REFERENCES

- Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students writing. In N. Ballier, A. Díaz-Negrillo, & P. Thompson (Eds.), *Automatic treatment and analysis of learner corpus data* (pp. 249–264). Amsterdam: John Benjamins. <https://doi.org/10.1075/scl.59.15ai>
- Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly*, 26, 390–395. <https://doi.org/10.2307/3587016>
- Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition*, 11(1), 17–34. <https://doi.org/10.1017/S0272263100007816>
- Bartning, I., & Schlyter, S. (2004). Procédés de grammaticalisation et son acquisition d'une langue étrangère. *Acquisition et Interaction en Langue Etrangère (AILE)*, 9, 9–50.
- Biber, D., & Gray, B. (2011). Grammatical change in the noun phrase: The influence of written language use. *English Language and Linguistics*, 15(2), 223–250. <https://doi.org/10.1017/S1360674311000025>
- Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5–35. <https://doi.org/10.5054/tq.2011.244483>
- Bulté, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency. Complexity, accuracy and fluency in SLA* (pp. 21–46). Amsterdam: John Benjamins.
- Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26(4), 42–65. <https://doi.org/10.1016/j.jslw.2014.09.005>
- Dahl, Ö. (2004). *The growth and maintenance of linguistic complexity. Studies in Language Companion Series*. Amsterdam: John Benjamins.
- De Clercq, B. (2015). The development of lexical complexity in second language acquisition: A cross-linguistic study of French and English. In L. Roberts, K. McManus, N. Vanek, & D. Trenkic (Eds.), *Eurosla yearbook* (Vol. 15) (pp. 69–94). Amsterdam: John Benjamins.

- De Clercq, B., & Housen, A. (2016). The development of morphological complexity: a comparative study on L2 French and English. *Second Language Research*, 35, 71–97. <https://doi.org/10.1177/0267658316674506>
- Della Putta, P. (2016a). The effects of textual enhancement on the acquisition of two nonparallel grammatical features by Spanish-speaking learners of Italian. *Studies in Second Language Acquisition*, 38(2), 217–238. <https://doi.org/10.1017/S0272263116000073>
- Della Putta, P. (2016b). Do we also need to unlearn constructions? The case of constructional negative transfer from Spanish to Italian and its pedagogical implications. In S. De Knop, & G. Gilquin (Eds.), *Applied construction grammar* (pp. 237–267). Berlin: Mouton de Gruyter.
- Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.
- Ferrari, S., & Pallotti, G. (2008). La variabilità situazionale dell'interlingua: Implicazioni per la ricerca acquisizionale e il testing linguistico. In G. Bernini, L. Spreafico, & A. Valentini (Eds.), *Competenze lessicali e discorsive nell'acquisizione di lingue seconde* (pp. 437–462). Perugia: Guerra.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: a unit for all reasons. *Applied Linguistics*, 21(3), 354–375. <https://doi.org/10.1093/applin/21.3.354>
- Gyllstadt, H., Granfeldt, J., Bernardini, P., & Källkvist, M. (2014). Linguistic correlates to communicative proficiency levels of the CEFR: The case of syntactic complexity in written L2 English, L3 French and L4 Italian. In L. Roberts, I. Vedder, & J. Hulstijn (Eds.), *Eurosla yearbook* (Vol. 14) (pp. 1–30). Amsterdam: John Benjamins.
- Hengeveld, K. (1998). Adverbial clauses in the languages of Europe. In J. van der Auwera, & D. P. Ó. Baidó (Eds.), *Adverbial constructions in the languages of Europe* (pp. 335–419). Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110802610.335>
- Hirano, K. (1991). The effect of audience on the efficacy of objective measures of EFL proficiency in Japanese university students. *Annual Review of English Language Education in Japan*, 2, 21–30.
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461–473. <https://doi.org/10.1093/applin/amp048>
- Housen, A., Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 performance and proficiency. Complexity, accuracy and fluency in SLA*. Amsterdam: John Benjamins.
- Hudson, B. (1976). Conjunction reduction, gapping, and right-node raising. *Language*, 52(3), 535–562. <https://doi.org/10.2307/412719>
- Hulstijn, J. H., Alderson, J. C., & Schoonen, R. (2010). Developmental stages in second-language acquisition and levels of second-language proficiency: Are there links between them? In I. Bartning, M. Martin, & I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research. Eurosla monographs series 1*. (pp. 11–20).
- Hunt, K. (1965). *Grammatical structures written at three grade levels*. NCTE Research report No. 3. Champaign, IL: NCTE.
- Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8(1), 63–99.
- Klein-Braley, C. (2002). Psycholinguistics of C-Test taking. In J. A. Coleman, R. Grotjahn, & U. Raatz (Eds.), *University language testing and the C-test* (pp. 131–142). Bochum: AKS-Verlag.
- Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing*, 20(2), 148–161.
- Lambert, C., & Kormos, J. (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. <https://doi.org/10.1093/applin/amu047>
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590–619. <https://doi.org/10.1093/applin/aml029>
- Loban, W. (1976). *Language development: Kindergarten through grade twelve*. NCTE Committee on Research Report No. 18.
- Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, 45(1), 36–62. <https://doi.org/10.5054/tq.2011.240859>
- Miestamo, M., Sinnemäki, K., & Karlsson, F. (2008). *Language complexity: Typology, contact, change*. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.94>
- Neary-Sundquist, C. A. (2016). Syntactic complexity at multiple proficiency levels of L2 German speech. *International Journal of Applied Linguistics*, 27(1), 242–262. <https://doi.org/10.1111/ijal-12128>.
- Noonan, M. (1985). Complementation. In T. Shopen (Ed.), *Language typology and syntactic description* (Vol. II) (pp. 42–140). Cambridge: Cambridge University Press.

- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417–528. <https://doi.org/10.1111/0023-8333.00136>
- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555–578. <https://doi.org/10.1093/applin/amp044>
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24(4), 492–518. <https://doi.org/10.1093/applin/24.4.492>
- Ortega, L. (2012). Interlanguage complexity: A construct in search of theoretical renewal. In B. Kortmann, & B. Szmrecsanyi (Eds.), *Linguistic complexity: Second language acquisition, indigenization, contact* (pp. 127–155). Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110229226.127>
- Osborne, T. (2006). Gapping vs. non-gapping coordination. *Linguistische Berichte*, 207, 307–308.
- Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, 30(4), 590–601. <https://doi.org/10.1093/applin/amp045>
- Pallotti, G. (2014). A simple view of linguistic complexity. *Second Language Research*, 31(1), 117–134.
- Progovac, L. (1998). Structure for coordination (Part 1). *GLOT International*, 3(7), 3–6.
- Ravid, D., & Berman, R. (2010). Developing noun phrase complexity at school age: A text-embedded cross-linguistic analysis. *First Language*, 30(1), 3–26. <https://doi.org/10.1177/0142723709350531>
- Schoonen, R. (1991). *De evaluatie van schrijfvaardigheidsmetingen. Een empirische studie naar betrouwbaarheid, validiteit en bruikbaarheid van schrijfvaardigheidsmetingen in de achtste groep van het basisonderwijs*. Dissertation, Universiteit van Amsterdam.
- Spoelman, M., & Verspoor, M. (2010). Dynamic patterns in development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish. *Applied Linguistics*, 31(4), 532–553. <https://doi.org/10.1093/applin/amq001>
- Van de Velde, F. (2009). *De nominale constituent. Structuur en geschiedenis*. Leuven: Universitaire Pers Leuven.
- Véronique, D. (2004). The development of referential activities and clause-combining as aspects of the acquisition of discourse in French as L2. *French Language Studies*, 14, 257–280. <https://doi.org/10.1017/S0959269504001796>
- Verspoor, M., Lowie, W., & Van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. *Modern Language Journal*, 92(2), 214–231. <https://doi.org/10.1111/j.1540-4781.2008.00715.x>
- Vyatkina, N. (2012). The development of second language writing complexity in groups and individuals: A longitudinal learner corpus study. *Modern Language Journal*, 96(4), 572–594.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. Honolulu, HI: University of Hawai'i Press.

How to cite this article: Kuiken F, Vedder I. Syntactic complexity across proficiency and languages: L2 and L1 writing in Dutch, Italian and Spanish. *Int J Appl Linguist*. 2019;29:192–210. <https://doi.org/10.1111/ijal.12256>

APPENDIX: PROMPT TASK 2

(English translation)

Every month your favourite newspaper invites its readers to have a say in what will be the leading article for their monthly supplement. This time the Editorial Board has come up with three suggestions:

1. the effects of global warming;
2. the importance of physical education;
3. pros and cons of animal experiments.

Out of these three suggestions one has to be selected. The selection is made by a Readers' Committee. Every member of the committee has to write a report to the editors in which s/he states which article should be selected

and why. On the basis of the arguments given by the committee members the Editorial Board will decide which article will be placed on the front page.

This month you have been invited to be a member of the Readers Committee. Read the brief descriptions of the suggestions for articles below. Determine which article should be on the front page and why. Write a report in which you give at least three arguments for your choice. Try to be as clear as possible and include the following points in your report:

- which article should be selected;
- what the importance of the article is;
- which readers will be interested in the article; and
- why the Editorial Board should place this article on the front page of the monthly supplement (give three arguments).

You have 35 minutes to write your text and you need to write at least 150 words (about 15 lines). The use of a dictionary is not allowed.