Specific Syntactic Complexity: Developmental Profiling of Individuals Based on an Annotated Learner Corpus

NINA VYATKINA
University of Kansas
Department of Germanic Languages and Literature
1445 Jayhawk Blvd., Room 2080
Lawrence, Kansas 66045
Email: vyatkina@ku.edu

This study tracks the development of syntactic complexity in the writing of two beginning German as a second language learners with English as a first language over four semesters of collegiate language study by using developmental profiling techniques applied to an annotated learner corpus. The focus of the investigation is on individual developmental pathways and differences between learners who follow the same instructional sequences. The study explores variation in terms of frequencies of the selected complexity features (coordinate, nominal, and nonfinite verb structures) using corpus analysis techniques with semi-automatic corpus annotation. Two developmental profiles emerge from an in-depth contextual investigation of the target linguistic phenomena. The results show that the general developmental trend is for increasing frequency and range of syntactic complexity features with learners diverging more from one another in the second half of the observation period. This study addresses existing gaps in interlanguage complexity research by focusing on benchmarking development rather than gauging proficiency, addressing specific rather than global complexity measures, and targeting instructed learners at beginning rather than high-intermediate and advanced proficiency levels. Suggestions for future developmental second language acquisition research and foreign language pedagogy are made.

LANGUAGE ACQUISITION RESEARCH IS FUNdamentally concerned with language development in individuals over time. Great strides have been made with longitudinal methods in child language development (e.g., Brown, 1973; van Geert & van Dijk, 2002), language attrition (e.g., Kemper, Thompson, & Marquis, 2001) and naturalistic second language acquisition, or SLA (e.g., Meisel, Clahsen, & Pienemann, 1981). However, an overwhelming majority of instructed SLA studies have been contrastive and cross-sectional rather than developmental and longitudinal. Noting this situation, many prominent researchers have called for instructed SLA to turn to longitudinal methods and more dynamic descriptions accounting for variability and nonlinearity (e.g., Larsen–Freeman & Cameron, 2008; Norris & Ortega, 2009; Ortega & Byrnes, 2008; Ortega & Iberri–Shea, 2005; Verspoor, de Bot, & Lowie, 2011).

Another suggested direction that could enrich SLA research is Learner Corpus Research (LCR). Larsen–Freeman and Cameron (2008) propose that combining corpus analysis methods with a close examination of discourse patterns may help illuminate the development of learner language as a complex system because learner corpora "give us access to stabilized patterns and variability around

them" (p. 210). Still a "fairly recent phenomenon" (Nesselhauf, 2004, p. 12), LCR has yielded a number of publications in the last few years that test developmental hypotheses against large quantities of empirical data from corpora. Especially promising is research on annotated corpora because they allow researchers to move beyond word-based analyses to a more abstract level of linguistic patterns in learner language (Granger et al., 2007; Meurers & Müller, 2009). Currently available computational resources such as automated part-of-speech (POS) taggers (Santorini, 1990; Schmid, 1994) can considerably speed up the processing of large amounts of linguistic data. An example is a large-scale project recently taken up by the English Profile Programme to define linguistic *criterial features* based on the Cambridge Learner Corpus, which will correspond to proficiency levels in the Common European Framework of Reference, or CEFR (see Hawkins & McCarthy, 2010). However, despite a few promising advances, SLA and computational linguistics still remain in a "fickle alignment" as aptly characterized by Borin (2002). Ortega and Sinicrope (2008) attribute this phenomenon to the lack of a developmental component:

> Because corpus linguistics as an approach targets the discovery of frequency patterns and statistical properties of large quantities of language, applications to SLA are limited by a lack of developmental sophistication. In SLA, as in the field of child language acquisition, the unit of analysis is the individual acquirer, not usage statistics. (p. 11)

The current study addresses this disconnect by applying developmental profiling techniques (Ortega, 2012) to a POS-annotated learner corpus, thus contributing to the body of developmental LCR studies that have recently begun to emerge (e.g., Belz & Vyatkina, 2008; Byrnes & Sinicrope, 2008; Meunier, 2010; Myles, 2008; Vyatkina, 2012, in press). The main goal is to track closely the development of syntactic complexity in the writing of two beginning second language (L2) German learners with first language (L1) English over four semesters of collegiate language study.

Furthermore, by targeting the construct of complexity, the study addresses a number of research gaps identified by Ortega (2012) in her discussion of the state of the art and future directions of interlanguage complexity research (see also Bulté & Housen, 2012). More specifically, this study (a) focuses on benchmarking development with complexity measures, the area least researched, versus gauging proficiency and de-

scribing performance (Ortega, 2012); (b) addresses the problem of "construct reductionism" by considering "all three senses of complex (length, subordination, and selected form frequency)" (p. 139); and (c) investigates "incipient capacities of instructed foreign language learners before they reach the infamous and ubiquitous 'intermediate' level investigated most widely in the SLA literature" (p. 150).

The article is organized as follows: The next section introduces the concepts of general and specific syntactic complexity and reviews research literature on the latter. Then the design of the study is described, including its method, participants, instructional context, elicitation tasks, and measures, as well as corpus annotation and analysis tools and procedures. Next, study results are presented for the following measures: mean length of clause, coordinate structures, complex nominals, and nonfinite verb forms. The results are then summarized and discussed, and the final section presents the conclusion and implications of the study.

RESEARCH BACKGROUND

General and Specific Syntactic Complexity

Linguistic complexity has figured prominently in SLA research as one of three indexes of L2 development, along with accuracy and fluency (Skehan, 1989). Most widely cited definitions of linguistic complexity include the term *variety*, and the development of complexity in learner language is, logically, associated with an increase in lexical and syntactic variety. For example, Wolfe–Quintero, Inagaki, and Kim (1998) propose "a wide variety of both basic and sophisticated structures" (p. 107) as an indicator of developed complexity, following Foster and Skehan's (1996) definition of development in syntactic complexity as "progressively more elaborate language" and "a greater variety of syntactic patterning" (p. 303). In a recent study of syntactic complexity, Lu (2011) defines it as "syntactic variation and sophistication, or, more specifically, the range of syntactic structures that are produced and the degree of sophistication of such structures" (Lu, 2011, p. 36). However, as Polio (2001) rightly notes, "variety does not enter in the equations" (p. 96) when it comes to calculating syntactic complexity. Instead, it is typically measured in length of surface syntactic structures (e.g., number of words in a T-unit). Other popular measures of syntactic complexity such as frequency (e.g., number of clauses per time unit) and ratios (e.g., clauses per

T-unit) are also essentially length-based (see Norris & Ortega, 2009; Ortega, 2003). In contrast, studies of lexical complexity more often include variety measures such as *D*, the measure of lexical diversity (Malvern et al., 2004).

To summarize, explaining variety in terms of more specific forms beyond general measures has largely not yet been attempted in SLA syntactic complexity research. Norris and Ortega (2009), in their recent overview of the subject, encourage researchers "to engage in more form-specific and development-sensitive measurement of L2 production" (p. 567; see also Robinson, Cadierno, & Shirai, 2009; Robinson & Ellis, 2008). In addition to global measurement of syntactic complexity, Norris and Ortega (2009) propose to measure "complexity by subordination, [. . .] complexity via phrasal elaboration, as well as possibly coordination if early proficiency data are also included" (p. 574). Existing research on syntactic complexity in terms of variety of specific forms is reviewed in the next section.

Specific Syntactic Complexity Studies

Cooper (1976) was an early study that suggested and explored a series of specific syntactic structures in addition to length-based surface structures as L2 proficiency indices. In a cross-sectional study, Cooper compared raw tallies of selected structures in a variety of written assignments of L2 German college-level students at four different proficiency levels (measured in years of study) as well as of native speakers. The chosen structures included coordinate phrases, complex nominals, adverbial subordinate clauses, and infinitive constructions. Cooper found that the frequencies of all structures increased linearly from level to level (and to a statistically significant degree at every other level). Coordinate phrases presented an exception from this pattern: Third- and fourth-year students used fewer such structures than second-year students, although the frequencies increased again in the writing of graduate students and native speakers.

Coombs (1986) investigated the relationship between surface syntactic structure (simple and complex sentences, coordinate and subordinate clauses, infinitive and passive constructions) and the old/new information structure in a qualitative case study. Coombs constructed syntactic profiles of two intermediate L2 German writers in terms of variation and frequency of the target forms and then analyzed the information structure of the texts. The results showed that, although one learner demonstrated more variety in the use of syntactic forms, both participants failed to select appropriate forms for indicating new and old information in the text. The author concluded that more instruction in discourse structure, rather than only in syntactic structure, is needed in teaching L2 writing at the collegiate level.

Two studies explored syntactic variety with a focus on task effects in L2 English production. Ellis and Yuan (2005) compared learner written and oral performance on a narrative task under a planned and an unplanned condition. They used raw tallies of different morphosyntactic verb forms (tense, modality, voice) as a variety measure. The results showed more variety of forms in writing than in speaking and in the planned condition than in the unplanned condition. Robinson (2007) compared the effects of more complex and less complex speaking tasks on syntactic features considered sophisticated, such as wh-clauses, infinitival phrases, and clausal conjoinings. He found that L2 learners produced more complex syntax in more cognitively challenging tasks, thus being able to access simultaneously multiple attentional pools.[1]

With the development of Natural Language Processing (NLP) tools in general and corpus analysis in particular, a number of studies emerged that applied automatic POS tagging and syntactic parsing for investigating syntactic variety in large learner corpora. Granger and Rayson (1998) compared POS frequencies in two similar-sized annotated corpora of argumentative essays: a learner corpus collected from advanced English learners with L1 French and a comparison corpus collected from native speakers of English (both American and British). The results showed that learner academic essays contained high frequencies of indefinite articles, first and second person pronouns, auxiliaries, and infinitives, but low frequencies of definite articles and lexical verb participles. Clusters of features from the former group, however, are indicative of orality and involvement, whereas features of the second group lend themselves to expert academic writing (see Biber, 1988). Granger and Rayson conclude that "[t]he automatic profiling technique has highlighted the speech-like nature of learner writing. The essays produced by French learners display practically none of the features typical of academic writing and most of those typical of speech" (p. 129). This conclusion parallels findings from a number of other corpus-based studies investigating intermediate to advanced L2 English writing of learners with various L1 backgrounds (e.g., Aarts & Granger, 1998; Borin & Prütz, 2004; Reid, 1992).

The most recent study that combined a synthesis of research on syntactic complexity (including both general and specific measures) with a large-scale corpus study is Lu (2011). Lu explored the discriminatory power of 14 syntactic complexity measures as indices of L2 proficiency levels. The study was conducted on a large corpus of English as a Foreign Language writing of L1 Chinese college-level students that was automatically annotated for target categories. Among general complexity measures, Lu found the clause to be the most informative unit, more so than the T-unit. Moreover, following the only existing example by Cooper (1976), Lu employed the specific complexity measures of complex nominal structures and coordinate phrases and found that they had a stronger discriminative power than most other, more general, measures. Furthermore, these specific measures fared best in ratios per clause (and not per T-unit). The values for these measures (as well as the general measure of the mean length of clause) linearly increased from level to level and discriminated between several adjacent or non-adjacent levels. In addition, ratios of complex nominal and coordinate structures per clause had low to no correlation with each other, which showed that they measured different constructs and, therefore, captured different aspects of development. This finding prompted Lu to recommend using them as supplementary measures in future research. The study concludes with a call for closer attention to phrasal complexity expressed in specific complexity measures in future research.

Whereas these studies were cross-sectional or employed a single data collection wave, Byrnes and Sinicrope (2008) and Byrnes (2009) explored the development of specific measures of syntactic complexity over three to four curricular levels in the writing of college-level students of L2 German with L1 English. Both studies focused on features considered characteristic of advanced proficiency levels: relative clauses (Byrnes & Sinicrope, 2008), nominalization, and grammatical metaphor (Byrnes, 2009). Both studies ascertained dramatic increases in frequency of these sophisticated structures at the third and fourth curricular level due to the shift of the instructional focus toward public discourse types and more nominal, expository, text genres. At the same time, the first emergence of the target structures was attested in several learners' early writing, at the first or second curricular level. Byrnes and Sinicrope further showed that, whereas the overall frequency of relative clauses increased steadily from level to level, their range decreased due to the lower frequency of marked types. More specifically, learners used primarily subject and direct object relative clauses and rarely indirect object, genitive, and prepositional relative clauses. The study thus displayed a more complex developmental picture than previously assumed in SLA research (Keenan & Comrie, 1977). Importantly, along with cohort developmental data, both studies described cases of selected individuals and showed that learners whose use of target structures seemed average when measured with one measure (e.g., frequency or range of relative clauses) surpassed the class average vis-à-vis other measures (e.g., length of relative clauses). Respectively, learners who had comparable values when measured with one metric showed very different developmental profiles when measured with a different metric. These results highlighted the importance of employing a set of complexity indices that measure both global and specific constructs as well as of longitudinal designs for studying development.

This study contributes to this emerging line of research by exploring the development of specific syntactic complexity in the writing of two learners of L2 German, applying developmental profiling to an annotated learner corpus and using a longitudinal design with multiple and relatively dense data collection waves. The study method is described in detail in the next section.

DESIGN

Method

The "profile analysis" procedure was developed by Crystal, Fletcher, and Garman (1976) for identification of the most salient linguistic features of specific texts and was later applied to assessing the syntactic and morphological development of L2 learners (Clahsen, 1985; Pienemann, Johnston, & Brindley, 1988). In LCR, this procedure has been applied to large cross-sections of learner language at certain proficiency levels. Thus, Granger and Rayson (1998) proposed using fully annotated learner corpora for "automatic profiling" (p. 14). The authors drew on Krzeszowski's (1990) postulate that every interlanguage is characterized by a "unique matrix of frequencies of various linguistic forms" (p. 212, cited in Granger & Rayson, 1998, p. 119). Currently, this procedure is being applied in the large-scale English Profile Programme (Hawkins & McCarthy, 2010). In SLA, Byrnes, Maxim, and Norris (2010) expanded the notion of profiling to the context of instructed L2 acquisition. They describe both "idealized writing profiles" (p. 91) as expected

student learning outcomes at different curricular levels as well as actual learners' performance profiles vis-à-vis these expectations, not only for groups but also for individuals.

The current study pursues this latter approach by combining individual profiling with longitudinal LCR. The data for the study were extracted from a corpus of longitudinal data for dozens of L2 German learners (see Vyatkina, 2012). However, the focus here is on two particular individuals. As Barlow (2005) notes, the development of specific grammar features by individual learners has not been tracked even in longitudinal corpus studies, "hence it is not possible to assess the variability in the developmental sequences followed by individuals" (p. 347). This study thus responds to calls for using longitudinal corpora precisely for this purpose: tracking progress of individual students to account "for potentially considerable individual variation, *both* within the overall non-linear and unevenly paced trajectory *and* in learners' actual performance profiles at a particular time" (Ortega & Byrnes, 2008, p. 288, emphasis in original).

Moreover, this study advances longitudinal LCR by focusing on specific syntactic complexity measures, an L2 other than English (German), and beginning proficiency levels. These are precisely the directions suggested by Ortega and Sinicrope (2008) for developmental profiling in corpus-based research:

> [D]evelopmental profiling of specific areas of the L2 is more laborious [than if performed with global metrics], even when done with the aid of computerized searches and codings. However, they offer the advantage of not only being more informative of qualitative patterns of development than global metrics are, but also of being less sensitive to the overall length of the corpus, provided the structures of interest are sufficiently elicited by the particular elicitation tasks employed. Neither type of interlanguage measurement approach (whether global metrics or developmental profiling) has been fully applied to target languages besides English or to levels of proficiency below upper-intermediate levels. (p. 4)

The focus of the investigation is on individual developmental pathways and between-learner differences rather than on the definition of the proficiency levels of the two learners. Similar to large-scale corpus analyses, the study explores variation in terms of frequencies of the target linguistic features, semi-automatically annotated with corpus analysis techniques. However, the developmental profiles emerge from an in-depth investigation of the target linguistic phenomena in a specific context rather than from statistically grounded comparisons. The context is defined by the nature of the focal longitudinal learner corpus. It was collected with the goal of tracking the development of beginning language learners in a collegiate institutional setting. The corpus is supplemented with a variety of metadata about the learners, curricular conditions, and elicitation tasks.

In contrast to longitudinal studies with few data collection waves over long time periods, this study uses the microgenetic approach to study development with multiple data collection waves over dense time intervals, which allows exploration of the "path, rate, breadth, source, and variability" of development (Siegler & Svetina, 2002, p. 793; see also Belz & Kinginger, 2003; Belz & Vyatkina, 2008). More specifically, the developmental dynamic is explored as sets of "complexification strategies"[2] (Ortega, 2012, p. 139; see also van Geert & van Dijk, 2002), or "repertoire of choices" (Ortega & Byrnes, 2008, p. 287) of specific syntactic structures, used by learners at each of the 14 measurement occasions. This approach helps analyze not only differences in the levels of the measured variables but also appearance and disappearance of certain variables at different time points. In this way, learner development is analyzed in terms of multidimensional variability and nonlinear relationships between the instructional progression and individual developmental paths.

Participants and Previous Study

The impetus for this study came from a previous study (Vyatkina, 2012), which analyzed a number of more general syntactic and lexical complexity measures for a learner cohort and 2 individual learners. The 2 participants, given the pseudonyms "Braden" (a 19-year-old male) and "Cassie" (a 30-year-old female), had a number of similarities in their language learning histories: They had American English as their L1, lived all their lives in the Midwestern region of the United States, started their collegiate study of German as true beginners, and progressed through four subsequent semesters. Moreover, Braden and Cassie were selected as focal learners because both of them have been found to be representative of their learner cohort vis-à-vis the following general complexity measures. First, their sentence length linearly increased over time, similar to the class average. Second, the length of the clause-type unit (Finite Verb unit, or FV-unit) did not show any clear developmental trend either for the class or

for the 2 focal learners. However, their developmental profiles were found to be different from each other as well as from the class average on a number of counts: (a) Braden's sentence length increased primarily due to using progressively more FV-units. This increase was somewhat larger than the class average, whereas Cassie stayed slightly but consistently below average on this measure; (b) Cassie's data suggested a slight positive correlation of FV-unit length with time, which did not reach significance because of strong upward and downward oscillations; (c) although the class average use of coordinating conjunctions linearly declined, the focal learners' data did not follow this trend. (Braden did not decrease his use of coordinators over time, whereas Cassie's data suggest a slight decline, which nevertheless showed no statistically significant correlation with time.) Finally, the cohort, as well as both focal learners, increased their lexicogrammatical variety over time, but Braden was consistently and significantly below average on this measure, whereas Cassie was consistently and significantly above it.

These differences in developmental profiles in terms of more global complexity indicators prompted the present study, which set out to explore specific complexification strategies that both learners developed in their writing over time and, thus, to supplement and fine-tune previous results.

Instructional Context

The data came from a longitudinal corpus of learner German and were collected at time intervals ranging from 3 to 5 weeks during each semester. The study followed the "instruction-embedded total-sampling approach" (Byrnes et al., 2010, p. 165); that is, writing samples were rough drafts of essays written by the students in response to curricular tasks rather than to external experimental tasks. The participants were students who enrolled in a beginning German language program at a large Midwestern university over four sequential 16-week-long semesters. The classes met for 5 weekly contact hours in the first and second semester and for 3 contact hours in the third and fourth semester. All classes in this multisection program were taught by graduate teaching assistants who followed a uniform syllabus under the supervision of the researcher and used the same textbooks but had freedom to design their own lesson plans. Each course included a combination of spoken interaction, grammar explanation and practice, writing assignments, Internet-based activities, and creative culture projects. Writing was allocated considerable attention as a separate course component (ca. 20% of the syllabus) rather than being considered a mere supporting skill (following recommendations by Abrams, 2010; Byrnes et al., 2010).

Tasks

Although data for the original corpus were collected over 19 time points (see Vyatkina, 2012), only time points at which both Braden and Cassie submitted their essays are used in this study, for better comparability of the construction of their developmental profiles (namely 1–7, 9–10, 12–13, 15, 17, and 19; Figures 1–5). Each writing task except for the final task (see Appendix A and Vyatkina, 2012, for details) concluded a corresponding textbook chapter and reflected the book's instructional content, including the focus on selected grammar structures. The tasks at T1–T5 (first semester) and T6–T10 (second semester) requested learners to write personal narratives. The tasks at T11–T14 (third semester) also required students to write personal narratives and personal accounts but with added argumentative elements. The tasks at T15–T19 (fourth semester) prompted students to reflect on the content of books they had read before. These writing tasks are considered level-appropriate for first- and second-year college-level L2 learners (Byrnes et al., 2010; Maxim, 2011).

During the first three semesters (T1–T14), students typed each essay in class under timed conditions. They were required to write during the whole 50-minute-long class period and were allowed to use online dictionaries but not online translators, textbooks, or notes. During most of the fourth semester (T15–T18), they wrote essays at home under untimed conditions and were allowed to use reference materials. The very last essay (T19) was again timed and written under controlled conditions.

It must be noted that, although variation in tasks and topics considerably affects linguistic complexity (Ellis & Yuan, 2005; Lu, 2011; Ortega & Sinicrope, 2008; Robinson et al., 2009), this study does not focus on these specific effects. Rather, the focus is on how 2 different learners respond to one and the same task at each time point. However, it is reported whether each task prompt required the learners to use specific syntactic forms (see also Vyatkina, in press, for a study of the influence of the pedagogical input on the same learners' output).

FIGURE 1
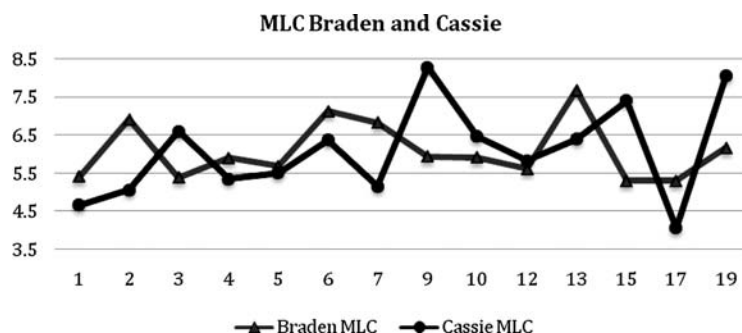Mean Length of Clause (MLC) by Participant and Time Point

MLC Braden and Cassie

—▲—Braden MLC   —●—Cassie MLC

FIGURE 2
Frequencies of Coordinate Structures per Clause (CS/C)

Braden CS/C   **Cassie CS/C**

▨ Predicates ■ Nominals   ▨ Predicates ■ Nominals
▨ Modifiers □ Clauses   ▨ Modifiers □ Clauses
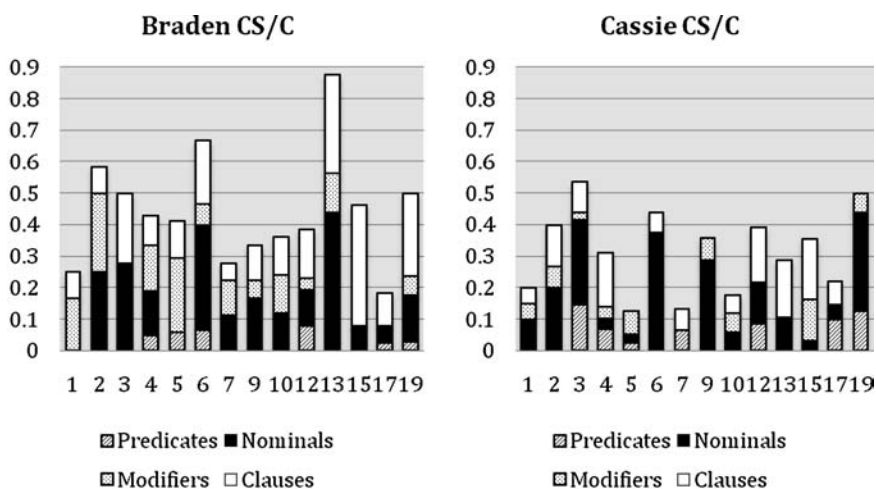
FIGURE 3
Frequencies of Complex Nominal Structures per Clause (CN/C)

Braden CN/C   **Cassie CN/C**

▨ Adjective phrases ■ Prep. phrases   ▨ Adjective phrases ■ Prep. phrases
▨ Noun clauses □ Relative clauses   ▨ Noun clauses □ Relative clauses
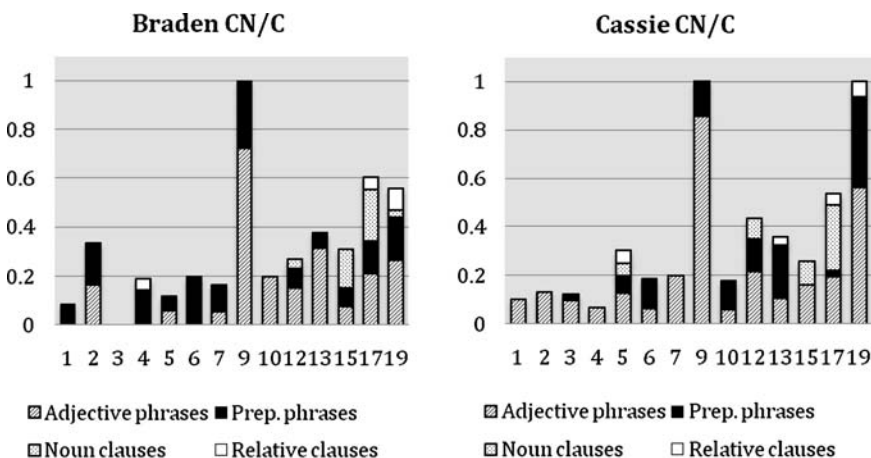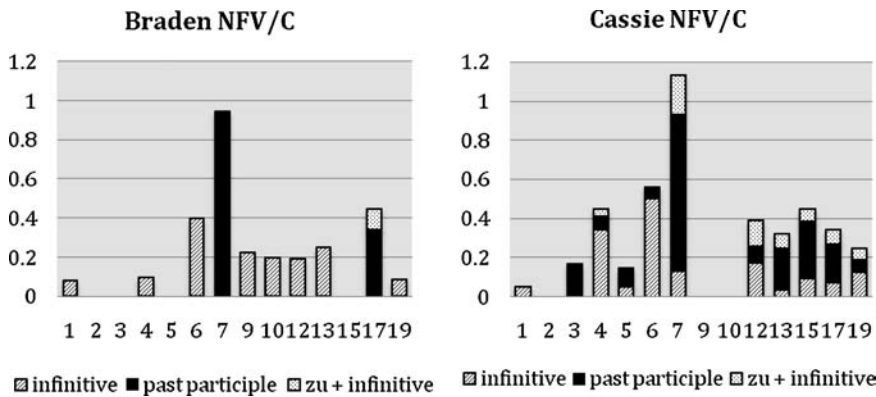
FIGURE 4
Frequencies of Nonfinite Verb Structures per Clause (NFV/C)

Measures

All measures in this study are ratios of selected syntactic structures per clause (following recommendations by Lu, 2011). The clause is a controversial unit lacking a uniform definition in complexity research (see Lu, 2011, and Polio, 2001, for discussion). Here, it is defined as a syntactic structure with a subject and a finite verb (Hunt, 1965), whereas nonfinite verb structures are considered phrases (following the majority of complexity studies).

Specific syntactic structures were selected to encompass different aspects of syntactic complexity: coordinate and subordinate clauses as well as verbal and nominal phrases (Norris & Ortega, 2009). These structures (the term *structure* is used here as an umbrella term for both clauses and phrases) were grouped based on the only available comprehensive classifications of specific L2 syntactic complexity features by Cooper (1976) and Lu (2010, 2011) and included (a) coordinate structures, (b) complex nominal structures, and (c) verb structures (Table 1). However, the specific constituency of these groups was slightly amended in comparison to Cooper and Lu to account for the context of this study.[3]

Corpus Annotation and Analysis: Tools and Procedure

This study aimed to benefit from available NLP resources and used automatic corpus tools for

FIGURE 5
Summative Frequencies of Coordinate Structures per Clause (CS/C), Complex Nominals per Clause (CN/C), and Nonfinite Verb Structures per Clause (NFV/C)
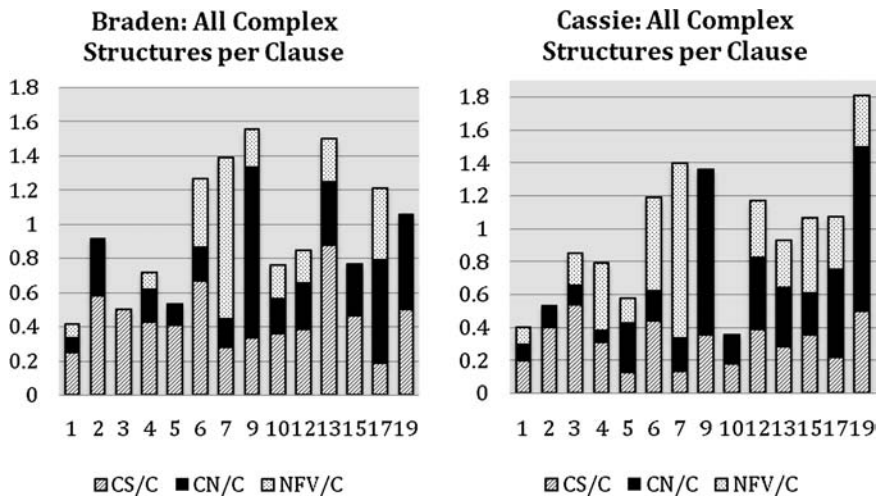
TABLE 1
Specific Syntactic Structures With Examples From the Learner Data (Errors Not Corrected)

Coordinate Structures	
nominal phrases	*zwei Kinder und zwei Enkel*
	'two children and two grandchildren'
modifier phrases	*Meine Jacke ist blau und weiß.*
	'My jacket is blue and white.'
predicate phrases	*Jan ist Single, aber hat einen Hund.*
	'Jan is single but has a dog.'
coordinate clauses	*Ich hasse Landmusik und ich höre nicht.*
	'I hate country music and I don't listen.'
Complex Nominal Structures	
attributive adjective phrases	*eine kleine menschliche Freundlichkeit*
	'a little human friendliness'
prepositional phrases extending nominal phrases	*Der Mann von das Bildnis*
	'The man from the picture'
nominal clauses	*Ich dachte, dass Frau Muschler den Wagen flickte.*
	'I thought that Mrs. Muschler repaired the carriage.'
relative clauses	*Dorian Gray ist ein reiche Mann, wer im viktorianisch England wohnt.*
	'Dorian Gray is a rich man, who lives in Victorian England.'
Nonfinite verb structures	
infinitive phrases governed by modal and auxiliary verbs	*Ich will eine Party machen.*
	'I want to make a party.'
"*zu* ('to') + infinitive" phrases	*Er muss seine Schuhe ausgezogen hatte, um niemanden zu stören.*
	'He must have taken off his shoes in order not to disturb anybody.'
past participle phrases	*Ich habe nich beschlossen.*
	'I have not decided.'

computing length-based measures as well as automatically assigning POS tags as proxy measures for surface syntactic structures (Aarts & Granger, 1998; Lu, 2010, 2011). However, since automatic POS-taggers never achieve 100% accuracy even on native speaker data (Schmid, 1994) and learner errors may affect the accuracy rate even more (Granger, 2009; van Rooy & Schäfer, 2002), this study adopted a semi-automated tagging procedure (Garretson & O'Connor, 2007).

For semi-automatic POS annotation, the learner corpus was first tagged automatically for 50 distinct word classes using the Tree Tagger for German (Schmid, 1994), and the output was manually checked (Meunier & de Mönnink, 2001). For evaluating POS annotations, the total tagger output on the writing of the 2 learners was checked manually by the researcher and independently by another annotator, using the guidelines for the tagset employed in the Tree Tagger (Schiller et al., 1999). The annotators evaluated tag accuracy based on the *intended context* of learner use (Byrnes & Sinicrope, 2008; Norris & Ortega, 2002), also called *target hypothesis*, or a

reconstruction of learner utterances in the target language (Ellis, 1994, p. 54; see also Lüdeling et al., 2005). In other words, the annotators confirmed the automatically assigned tag if it corresponded to the structure attempted by the learner (as recognized by the annotators), despite possible mistakes in its realization, or they corrected the tag in accordance with this target hypothesis. Inter-annotator agreement was calculated using the *F*-score (Lu, 2010, 2011), which ranged from 0.95 to 1 for different POSs. All discrepancies were resolved by agreement.

The final annotation of specific complex structures was performed as follows. Past participles of full verbs were captured by the initial semi-automatic POS annotation. For capturing coordinate structures, corpus analysis software WordSmith Tools (Scott, 2008) was used. First, automatic search was performed for coordinating conjunctions using the *Concord* function. This allowed retrieval of the so-called concordance lines (see Tribble, 2010): All text lines containing coordinating conjunctions were stacked one under the other with the search word (conjunction) in the middle. Then, different

coordinate structures were counted manually based on the context of each retrieved example. The concordancer was also used for capturing adjectival and prepositional phrases. First, the search retrieved all concordance lines containing nouns. Then, the results were automatically sorted by the left and right context using the *Sort* function. Adjectives appeared as adjacent lefthand neighbors and prepositions as adjacent righthand neighbors of nouns. All found concordance lines were analyzed manually by the researcher, and irrelevant structures (e.g., prepositional phrases extending verb phrases rather than noun phrases) were removed.

Finally, clause types and infinitive constructions were annotated manually by two independent annotators, following the CHILDES conventions (MacWhinney, 2000) and the guidelines for annotating learner German, developed by Norris and Ortega (2002) and fine-tuned by the research team at the Georgetown University German Department (e.g., Byrnes et al., 2010). For all clauses, the inter-annotator agreement was very high with the *F*-score of 0.98. For specific structures (noun and relative clauses as well as infinitive constructions), the *F*-score ranged from 0.8 to 0.9. All discrepancies were resolved by agreement. Finally, frequencies of infinitives governed by modal verbs were calculated by subtracting the number of infinitive constructions (annotated manually) from the number of total infinitives (annotated semi-automatically).

Frequency values for all complexity measures by participant and time point are summarized in Tables B1 and B2 in Appendix B.

RESULTS

Mean Length of Clause

As explained above, the clause was selected as the baseline unit for this study, and specific complexity measures were calculated as ratios to the number of clauses per time point (i.e., in each writing sample). Prior to analyzing these measures, it was decided to calculate the mean length of clause (MLC) to establish the general clausal complexity background. MLC is a unit that figured prominently in previous complexity research (e.g., Byrnes et al., 2010; Lu, 2011; Norris & Ortega, 2009) and is calculated by dividing the number of words in a writing sample by the number of clauses. Figure 1 presents the MLC dynamic over the observation period (14 time points) in the writing of the 2 learners.

Figure 1 illustrates that both line graphs exhibit considerable upward and downward oscillations. However, whereas Braden's MLC shows no discernible linear dynamic, Cassie's MLC suggests a trend to a linear increase. More specifically, Braden's MLC ranges from 5.3 to 7.7, and Cassie's from 4 to 8.3. The average MLC, however, is identical for both learners at six words per clause. This result corresponds with the findings from the previous study (Vyatkina, 2012), which showed a trend of a linear increase of a clause-type unit (FV-unit) in Cassie's data and no such trend in Braden's data, with no significant correlations with time in either case.

Coordinate Structures

The stacked column graphs in Figure 2 illustrate the frequency and distribution of the focal coordinate structures as used per clause by Braden and Cassie at each time point. Both graphs show upward and downward oscillations, but Braden's frequencies are almost invariably higher than Cassie's. More specifically, Braden's frequencies range from 0.18 to 0.88 with an average of 0.44, whereas Cassie's frequencies range from 0.13 to 0.54 with an average of 0.32.

The analysis of specific coordinate phrases, both in terms of frequencies per clause and context of use, revealed the following picture. Coordinate nominal phrases appear in almost all learners' essays. The analysis of concordance lines showed that they were used for the purpose of listing people or objects. For example, Cassie used a total of 11 coordinate nominals at T3 to list her relatives while describing her family tree (1a), and Braden used 7 coordinate nominals to describe various information sources and TV shows in response to the "mass media" topic at T13 (1b). Coordinate modifiers mostly appear as predicate adjectives, thus extending finite verb phrases (1c). Braden used more coordinate verb phrase modifiers per clause than Cassie, especially at T1–T5 (Figure 1), that is, during the first semester. In contrast, Cassie consistently used more coordinate predicates than Braden per clause, thus adding more finite (1d) or nonfinite (1e) verb forms.

Next, the analysis of coordinate clauses per clause shows that until T5, both learners used them with comparable frequencies. From T5 onward, although the dynamic for both learners continues to fluctuate, Braden's frequencies are consistently higher than Cassie's. Frequencies for both learners reach a peak of coordinate clauses per clause at T15, but for Braden, the peak value is

0.38 and for Cassie only 0.19. At the final time point T19, Cassie's frequencies go down to 0, whereas Braden's go up to 0.26.

Finally, a more in-depth contextual analysis of concordance lines showed that subjects in Braden's coordinate clauses were often repetitive. More specifically, he produced the total of 14 compound sentences with repeated subjects (1f), whereas Cassie used only five such sentences. In contrast, Cassie used 11 sentences with coordinate predicates (1d, 1e), whereas Braden used only three such sentences.

(1a) *Meine Tante Kay und ihr Mann Bill haben zwei Kinder und zwei Enkel.* (nominals, Cassie)
'My aunt Kay and her husband Bill have two children and two grandchildren.'

(1b) *In mein Auto höre ich Rockmusik und Metalmusik zum Beispiel Ted Leo und die Apotheker.* (nominals, Braden)
'In my car, I listen to rock music and metal music for example Ted Leo and the Pharmacists.'

(1c) *Meine Jacke ist blau und weiß.* (modifiers, Braden)
'My jacket is blue and white.'

(1d) *Jan ist Single, aber hat einen Hund.* (predicates, Cassie)
'Jan is single but has a dog.'

(1e) *Ich genieße Flugzeug reisen und meine eigene Kleidung nähen.* (predicates, Cassie)
'I enjoy traveling [by] plane and sewing my own clothes.'

(1f) *Ich hasse Landmusik und ich höre nicht.* (coordinate clauses with repeated subjects, Braden)
'I hate country music and I don't listen.'

In sum, Cassie's use of all coordinate structures shows upward and downward oscillations, which are fairly evenly distributed across types and the timeline. In contrast, Braden uses progressively fewer coordinate phrases (nominals, modifiers, and predicates) but more coordinate clauses per clause in the second third and especially the final third of the observation period.

Finally, to get additional information on lexical variety of specific syntactic structures, the number of unique conjunctions used by each learner to connect coordinate structures was counted

(Table 2). For this purpose, the analysis of concordance lines and frequency lists was carried out. Not surprisingly, the coordinator *und* ('and') was used by both learners most frequently (103 times by Braden and 75 times by Cassie).[4] However, both learners also used increasingly more contrastive conjunctions, especially from the beginning of the third semester (T12) onward, such as *aber* ('but,' used 18 times by Braden and 15 times by Cassie) and *oder* ('or,' used 2 and 9 times, respectively). Additionally, Cassie also used *sondern* ('but') once and *doch* ('but, however') once, and Braden used *bis* ('till') three times. In sum, Cassie showed more variety in coordinator types than Braden.

The overall column height in Figure 3 shows an early emergence and a stable use of complex nominals for both learners. The frequencies oscillate around 0.2 complex nominals per clause up until T9, when both graphs show a sharp spike at one complex nominal per clause. At T10, the frequencies drop back down to around 0.2 and then grow again at T12 to remain above 0.2 with a trend to a linear increase toward the end with upward and downward oscillations. Cassie's overall frequencies are higher than Braden's. Moreover, the dynamics of the column composition reveal that both learners gradually enriched their complex nominals repertoire and reached the greatest diversity in the second half of the observation period by using three to four different complex nominals strategies in their later essays.

As opposed to coordinate structures that have not been explicitly taught to learners, several components of complex nominal structures were associated with grammar instruction foci. Therefore, the emergence of these complex nominals in learner production can be traced in relation to their first exposure to each structure in the curriculum. Subordinate clauses (including noun clauses) were introduced immediately prior to T9, attributive adjectives prior to T10, and relative pronouns and clauses prior to T13. Finally, prepositional phrases had been progressively introduced first as chunks from the very beginning ("on Monday," "to go to the movies") and later as free constructions, although they had not been explicitly taught as a noun modification strategy.

TABLE 2
Frequencies of Coordinating Conjunctions

	und	*aber*	*Oder*	*sondern*	*doch*	*bis*	Total
Braden	103	18	2			3	126
Cassie	75	15	9	1	1		101

The sharp surge in complex nominals per clause at T9 happens due to a high number of attributive adjectives used as a pre-modifier strategy by both learners (2a). This can be directly attributed to the instructional focus, including the T9 writing prompt that asked learners to describe their home city and provided a few adjective phrase models. Although attributive adjectives appear in both learners' repertoires before the focused instruction, they begin using them in a larger proportion from T9 on. This is a desirable learning outcome because, as evidenced in the data, learners persist in employing this more advanced complex nominal strategy despite a notorious difficulty of the adjective inflection paradigm in German. Finally, the contextual analysis of concordance lines revealed that Cassie more frequently used strings of two non-coordinate adjective pre-modifiers, in other words, even more complex structures (2b). She used a total of four adjective pairs, whereas Braden used only one such pair in the final essay.

The analysis of post-modifier complex nominal structures (prepositional phrases and embedded clauses) reveals the following picture. Complex nominals in the form of prepositional phrases emerge in learner writing as early as at T1 (Braden) and T3 (Cassie) and are used consistently throughout the observation period (2c). The difference between the 2 learners is that there is no discernible dynamic change in the proportion of prepositional phrases to other complex nominals for Braden but a progressive increase for Cassie. Complement clauses (2d) emerge in Braden's writing after focused instruction (prior to T10), and in Cassie's writing even earlier (at T5). The use of this strategy increases for both learners in the second half of the timeline. Finally, relative clauses (usually considered the most advanced type of nominal post-modification) are first attempted by both learners early (at T4 by Braden and T5 by Cassie). Both return to this complex nominal strategy after focused instruction and use it at several time points toward the end of the observation period.

The contextual analysis of relative clauses showed that all of them were either subject relative clauses (2e) or direct object relative clauses (2f). Thus, the 2 focal learners used only those relative clause types that have been previously found to be the most frequent ones in the writing of collegiate students of German (Byrnes & Sinicrope, 2008).

(2a) *Es schneit im spä̲t̲e̲n̲ Winter und b̲a̲l̲d̲i̲g̲e̲n̲ Frühling.* (adjectives, Braden)
    'It snows in the late winter and early spring.'

(2b) *Eine k̲l̲e̲i̲n̲e̲ m̲e̲n̲s̲c̲h̲l̲i̲c̲h̲e̲ Freundlichkeit ist geteilt.* (adjectives, Cassie)
    'A little human friendliness is shared.'

(2c) *Der Mann v̲o̲n̲ d̲a̲s̲ B̲i̲l̲d̲n̲i̲s̲ alterte,* […] (prepositional phrase, Braden)
    'The man from the picture aged.'

(2d) *Ich dachte, d̲a̲s̲s̲ F̲r̲a̲u̲ M̲u̲s̲c̲h̲l̲e̲r̲ d̲e̲n̲ W̲a̲g̲e̲n̲ f̲l̲i̲c̲k̲t̲e̲.* (complement clause, Braden)
    'I thought that Mrs. Muschler repaired the carriage.'

(2e) *Dorian Gray ist ein reiche Mann, w̲e̲r̲ i̲m̲ v̲i̲k̲t̲o̲r̲i̲a̲n̲i̲s̲c̲h̲ E̲n̲g̲l̲a̲n̲d̲ w̲o̲h̲n̲t̲.* (subject relative clause, Braden)
    'Dorian Gray is a rich man, who lives in Victorian England.'

(2f) *Ich trage meinen Herzkette (d̲e̲n̲ i̲c̲h̲ j̲e̲n̲e̲n̲ g̲l̲e̲i̲c̲h̲e̲n̲ T̲a̲g̲ g̲e̲k̲a̲u̲f̲t̲), mehr oft.* (direct object relative clause, Cassie)
    'I wear my heart chain, which I bought on that same day, more often.'

Nonfinite Verb Structures

Comparison of the general dynamics of the focal nonfinite verb forms (Figure 4) shows that Cassie used more nonfinite verb forms per clause than Braden. Both graphs have a peak at T7. Furthermore, it is apparent that Cassie used a more balanced variety of nonfinite forms, especially in the second half of the time line (third and fourth semesters), with all three nonfinite types at all time points. In contrast, Braden used either no or only one nonfinite verb strategy at each particular time point except for T17, when he used two different nonfinite verb phrase types.

The selected nonfinite verb forms were introduced as instructional foci in the following sequence. Learners' exposure to infinitives began from the very beginning of instruction as vocabulary items in verb lists. However, the first syntactic construction containing an infinitive was introduced at T4 ("modal verb + infinitive" such as [I can go]). Similarly, a number of selected past participle forms were presented as unanalyzed vocabulary items early on (at T1–T6), for example, in phrases: *Ich bin geboren* 'I was born', *Das Zimmer ist möbliert* 'The apartment is furnished', and *Das Restaurant ist geöffnet/geschlossen* 'The restaurant is open/closed'. Then past participles became the focus of grammar instruction at T7 with the introduction of the present perfect tense (e.g., [I have gone]). It should be noted that, once introduced, infinitives and past participles consistently remained in the pedagogical input in the form of both old constructions and newly introduced constructions (such as future tense, introduced before T12, and passive voice,

introduced before T15). Finally, "*zu* [to] + infinitive" constructions were explicitly taught at T13.

Both learners began using infinitives at T1 (i.e., before the focused instruction at T4), kept using them at subsequent time points, and reached a peak of infinitives per clause at T6. The task context explains this fact. The prompt asked students to describe a planned party, which triggered many constructions of a modal auxiliary plus infinitive in response (3a). Additionally, although the prompt contained questions in the present tense (such as [When is the party?]), Cassie decided to use the future tense in response long before it became the focus of instruction, which entailed using the auxiliary *werden* [will] and an infinitive (3b). The more advanced infinitive constructions, *zu* + infinitive, were used differently by the 2 learners. Braden only used these forms at late observation points (3c). In contrast, Cassie used *zu* + infinitive first at an early time point, T4, and later used it consistently at each time point from T12 through T19 (3d).

The third focal NFV form, past participles, was formally introduced in conjunction with the present perfect tense at T7. However, Cassie began using it fairly extensively and consistently before the focused instruction, from T3 onward. Although four instances at T3 are accounted for by the form *geboren* [born] introduced as unanalyzed vocabulary items in chunks such as [I was born], she uses most past participles in free constructions with present perfect (3e). After focused instruction, Cassie resumed using past participles at T12 and consistently used them at each time point until the close of the observation period. In contrast, Braden used past participles for the first time at T7; that is, as a direct response to the focused instruction as well as to the writing prompt that contains specific present perfect tense forms. After T7, past participles disappeared from Braden's writing until T17, when he used them again in response to a prompt containing a direct request to use present perfect.

(3a) *Ich will eine Party machen.* (infinitive governed by a modal verb, Braden)[5]
  'I want to make a party.'
(3b) *Ich werde Abendessen kochen.* (infinitive governed by the future tense auxiliary, Cassie)
  'I will cook dinner.'
(3c) *Frau Muschler und Herr Muschler liebten lustig zu haben.* (*zu* + infinitive, Braden)
  'Mrs. Muschler and Mr. Muschler liked to have fun.'
(3d) *Er muss seine Schuhe ausgezogen hatte, um niemanden zu stören.* (*zu* + infinitive, Cassie)

  'He must have taken off his shoes in order not to disturb anybody.'
(3e) *Ich habe nich beschlossen.* (past participle with present perfect, Cassie)
  'I have not decided.'
(3f) *Ablagen mussten Farbe gekennzeichnet werden* (past participle with passive voice, Cassie)
  'Racks had to be color marked.'

Next, the use of past participle forms was explored in context using concordance lines and frequency lists. First, it was ascertained that both learners used past participle forms appropriately (although not always accurately) in present perfect (or, in a few instances, in past perfect) constructions. Moreover, Cassie used past participles appropriately in the passive voice at later time points (3f). Furthermore, the number of unique past participle forms was calculated for each learner. It turned out that Cassie used more unique verb types than Braden. In other words, Braden more frequently repeated the same forms when he used past participles, whereas Cassie more frequently built past participles from new verbs (see also Vyatkina, in press, for more details on this feature).

SUMMARY AND DISCUSSION

Figure 5 presents summative frequencies of all complexification strategies per clause for both learners. The comparison of the dynamic of the overall column height shows that both learners develop similarly in the first half of the observation period, starting at around 0.4 complex structures per clause and gradually increasing their frequency up to approximately 1.2 at T6 and approximately 1.4–1.5 at T7 and T9. Moreover, both learners use similar proportions of similar strategies at several time points, such as more nonfinite verb phrases at T7 and more complex nominals at T9. This similarity is explained by the instructional focus on respective syntactic structures at those time points. In contrast, the developmental paths of the 2 learners largely diverge in the final third of the time line. After two downward oscillations, Cassie's use of complex structures linearly increases after T13 to reach 1.8 structures per clause at the final time point. In contrast, Braden's complex structure frequencies show no stable dynamic, suggesting a declining trend. Furthermore, Cassie's use of complex structures is more balanced and evenly distributed in the final third of the observation period: She consistently uses all three strategies from T12 through T19, whereas Braden does not use NFV at T15 and T19. In terms of specific complex structures, Cassie uses more

complex nominals and nonfinite verbs per clause, whereas Braden uses more coordinate structures. More specifically, as shown earlier, Braden surpasses Cassie in his use of coordinate clauses but not coordinate phrases.

The previous study showed that the focal learners were similar vis-à-vis some general developmental measures. In particular, both of them used progressively longer sentences, and the mean clause-type unit length did not increase significantly for either of them. This study fine-tuned these results by demonstrating nonlinearity and variation within these general similarities and confirmed Lu's (2010, 2011) suggestion that rarely used specific features such as coordinate structures, complex nominals, and nonfinite verb forms (see Cooper, 1976) can reveal new facets of the development of syntactic complexity. Although specific syntactic complexity in the writing of both learners developed similarly during the first and second semester, Cassie developed to a much greater extent than Braden in the third and fourth semester. She surpassed Braden in her use of more advanced complex structures such as complex nominals and nonfinite verbs per clause, while the proportion of less advanced structures, such as coordinate clauses, was higher in Braden's writing. It can be concluded that Braden made his sentences progressively longer by using more coordinate and simplex clauses and Cassie by using more complex clausal structures. Although Cassie's data only suggest a trend toward increasing length of clause (see Figure 1), the developmental profiling method employed in this study revealed emerging complexity in Cassie's writing in terms of specific advanced syntactic features. This finding highlights the importance of combining general and specific measures in developmental research.

The results of the study have also shed more light on the heterogeneous nature of coordination and its use by developing writers. As the previous study showed, neither Braden nor Cassie significantly reduced their use of coordinating conjunctions over time. However, this study shows that they used coordination for different purposes. Cassie used more complex coordinate structures per clause than Braden in the second half of the observation period and, in particular, more coordinate predicates. In contrast, Braden reduced his use of such structures but increased his use of coordinate clauses in the second half of the timeline and used almost twice as many such structures per clause as Cassie. Additionally, as the analysis of the concordance lines has shown, subjects in Braden's coordinate clauses were often

repetitive. Following Halliday and Martin (1993), Byrnes et al. (2010, p. 167) explain that the ability to combine clauses within sentences (a "dynamic" writing style) is associated with lower proficiency levels, whereas complexity through phrasal elaboration (a "synoptic" writing style) characterizes more advanced writing. In line with this argument, Cassie goes a longer way toward a more economical writing style typical of expert discourses, whereas Braden's style with piled up coordinate clauses and redundant subjects remains at a less advanced stage, typical of everyday oral discourses.

Moreover, Cassie's writing became progressively more balanced, more so than Braden's. While both learners exhibited spikes in specific structures after focused instruction sessions, different complexity strategies became more evenly distributed in Cassie's writing at later time points, whereas in Braden's writing, some strategies disappeared for several observation points and other strategies dominated. Higher variability in Braden's use of complex structures may indicate that his syntactic complexity system is still developing, whereas Cassie's is closer to a stabilization period (cf. Verspoor et al., 2011). Cassie's writing proved to be more balanced also in terms of a richer lexical repertoire, which is evidenced by more variety of coordinating conjunctions and past participles. This finding confirmed the result from the previous study, which showed that Cassie consistently and significantly surpassed Braden (as well as the class average) on the general lexical variety measure (adjusted type–token ratio, or CTTR).

CONCLUSION AND IMPLICATIONS

This study contributes to the longitudinal study of L2 development by combining developmental profiling and corpus analysis methods. It supplements the currently small body of "concrete empirical evidence … that interlanguage complexity in second language production increases as individual interlanguage grammars develop, and that the increases are meaningful vis-à-vis the mapped course of second language development" (Ortega, 2012, p. 133). Importantly, the study contributes to benchmarking interlanguage development at lower proficiency levels, which have been largely neglected to date. It supplements and fine-tunes results from a previous study (Vyatkina, 2012), providing in-depth developmental profiles of two focal learners in terms of specific syntactic complexity, an underexplored area to date. These successive levels of analysis have led to greater insight into similarities and differences between

individual developmental paths taken in the same instructional context.

The study used a semi-automatic corpus annotation method, which allowed the researcher to look at corpus evidence "in a more linguistically informed way" (Hunston, 2010, p. 159). Additionally, the data collection in dense observation waves "allowed insight into the change process that would have been impossible without it" (Siegler & Svetina, 2002, p. 804). Because the study tracked the 2 learners from the incipient level of proficiency, it pinpointed the "timing of emergence," a criterion never considered before in conjunction with syntactic complexity features, according to Ortega (2012, p. 139). Specifically, it showed at what point the target features emerged in the writing of the 2 learners, whether they stayed or disappeared after focused instruction, and how different strategies were distributed at each time point. The study showed that all specific syntactic complexity measures applied in previous research for higher proficiency levels (Cooper, 1976; Lu, 2010, 2011) emerge at the very beginning levels. It confirmed suggestions from previous research (Bardovi–Harlig, 1992; Norris & Ortega, 2009) that coordination is an important syntactic complexity feature to consider for beginning learners. It also supplemented the inventory of nonfinite verb measures with past participles. The study shows that the general developmental trend is toward increasing frequency and range of syntactic complexity features, with instructional foci and task prompts initially causing surges in the use of targeted features. This finding is in line with the trend found by Ortega and Sinicrope (2008) for the repertoire of word classes in beginners' oral performance. However, whereas some features show an almost linear increase in the writing of both learners (e.g., adjective phrases), others only hint at emerging patterns at later time points (e.g., "*zu* + infinitive" constructions). Moreover, whereas both learners follow similar developmental paths in the first two semesters, they diverge considerably in the second half of the observation period. While one student continued relying more on coordination, the other student favored more diverse and more complex clausal and phrasal complexification strategies. This latter finding confirms the trade-off effect between different complexity types found in previous research (Bardovi–Harlig, 1992; Biber, Gray, & Poonpon, 2011; Norris & Ortega, 2009). These observations will serve as hypotheses to be tested in a study of cross-student pattern identification (currently underway).

Moreover, it was especially interesting to observe how one participant readily responded to instruction but abandoned some syntactic features when progressing to the next task, while the other participant balanced both previously learned and new features in her writing (see also Vyatkina, in press). This latter result entails an important pedagogical implication that, given the great influence of instruction on learner performance, learners should be explicitly taught how to use level-appropriate and task-appropriate linguistic features. Instruction should go beyond presenting grammar as rules and paradigms and aim at explaining the meaning-making value of these features for achieving specific communicative purposes (Byrnes et al., 2010; Chavez, 2011; Coombs, 1986). One of the instructional innovations implemented in later iterations of the focal language program was the design of rubrics listing specific lexical and syntactic features (with examples) associated with each level-appropriate writing task. In this way, the learners can see the "idealized writing profiles" (Byrnes et al., 2010, p. 91) expected from them before engaging with the task. For example, the rubric for an intermediate-level task involving hypothetical narration about a future event solicits more general, previously learned complexity features (subordinate clauses) as well as more specific, newly introduced features (subjunctive mood verb phrases). Prior to writing, it is also advisable to provide learners with model texts with the desired features and discuss their contextual use and functions (see Dykstra–Pruim & Redmann, 2011, for sample pedagogical materials). Although it will still depend on each learner's agency to what extent they will use the learned features in their own writing, such pedagogical activities will raise learners' awareness about what their expected developmental targets are.

Future developmental corpus research should continue combining general and specific measures as well as tracking L2 development both in groups and in individuals with the goal of achieving as multifaceted a picture of longitudinal SLA as possible. In particular, researchers should elicit and explore multidimensional developmental profiles from corpora combining lexical and syntactic complexity measures as well as accuracy and fluency measures. This study has only tapped into the area of lexical variety by counting specific lexemes used in specific syntactic constructions. A larger study is now underway that investigates the development of an interrelated system of clausal and phrasal syntactic features, word classes, and lexicogrammatical constructions (Biber et al., 2011) in the focal learner corpus. Such an integrated usage-based approach has a high potential

for advancing developmental SLA research in all three areas highlighted by Ortega (2012): developmental benchmarking, gauging proficiency, and assessing performance.

Finally, learner corpus research should take advantage of developments in the NLP field, which provide increasingly reliable parsing and annotation tools. Such a marriage of LCR and NLP can speed up the processing of large amounts of language data, which can lead to more rapid advances in longitudinal SLA research. Moreover, with more corpus analysis resources made accessible to language teachers, they will be able to compile their own small learner corpora (Brown, 2007), quickly retrieve learner developmental profiles in the form of charts, and share those with their students for the purposes of setting learning goals, feedback, and assessment.

NOTES

[1] See Housen and Kuiken (2009) for discussion of two alternative hypotheses: Robinson's Cognition Hypothesis and Skehan's Trade-off Hypothesis.

[2] The term *strategy* is used to refer to syntactic constructions (also called *features* and *structures*) used by learners in their writing. More specifically, the term *strategy* highlights that various forms may be used to express the same syntactic function. For instance, prepositional phrases and adjective phrases are different strategies for nominal complexification. In contrast, the term *measure* refers to analytical constructs expressed in relative frequencies of these structures (see section "Measures").

[3] This classification differs from Cooper's (1976) and Lu's (2010, 2011) in the following ways. First, adverbial clauses are not considered here because they were captured under the analysis of subordination in Vyatkina (2012) and were shown to have similar frequencies in the writing of the 2 focal learners. Second, coordinate clauses were added to encompass the full range of possible coordinate structures both below and at the clausal level. Third, possessive pronouns (such as "my," "your") were found in abundance in learner writing throughout the developmental course, similarly to determiners (such as "this," "many"). Previous research includes possessives as markers of syntactic complexity but excludes determiners. To avoid this inconsistency as well as inflation of frequencies due to the high numbers of possessives in the data, this feature was excluded because it can hardly serve as a complexity indicator. Finally, as far as verb structures are concerned, only nonfinite verb forms were selected, following Cooper. Lu counted all verb phrases (both finite and nonfinite); however, he found this complexity measure not informative. In addition to Cooper's infinitives governed by modal verbs and "*zu* ('to') + infinitive" constructions, past participles were also added to nonfinite verb constructions. Infinitives in

subject position were not included because they were not found in the data during preliminary screening.

[4] The conjunction "and" and its equivalents in other languages is one of the most frequent words in both native and learner corpora (Ortega & Sinicrope, 2008; Tschirner, 2005).

[5] The German modal verb *wollen* is used without the infinitive particle *zu*, unlike the English 'want to.'

REFERENCES

Aarts, J., & Granger, S. (1998). Tag sequences in learner corpora: A key to interlanguage grammar and discourse. In S. Granger (Ed.), *Learner English on computer* (pp. 132–141). New York: Longman.

Abrams, Z. (2010). Writing. In C. Blyth (Ed.), *Foreign language teaching methods*. Texas Language Technology Center, University of Texas at Austin. Retrieved November 27, 2012, from http://coerll.utexas.edu/methods/modules/writing/

Bardovi–Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly, 26*, 390–395.

Barlow, M. (2005). Computer-based analyses of learner language. In R. Ellis & G. Barkhuizen (Eds.), *Analyzing learner language* (pp. 335–357). Oxford: Oxford University Press.

Belz, J. A., & Kinginger, C. (2003). Discourse options and the development of pragmatic competence by classroom learners of German: The case of address forms. *Language Learning, 53*, 591–647.

Belz, J. A., & Vyatkina, N. (2008). The pedagogical mediation of a developmental learner corpus for classroom-based language instruction. *Language Learning and Technology, 12*, 33–52.

Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*, 5–35.

Borin, L. (2002, August). *What have you done for me lately? The fickle alignment of NLP and CALL*. Paper presented at the Annual Conference of the European Association for Computer-Assisted Language Learning, Jyväskylä, Finland.

Borin, L., & Prütz, K. (2004). New wine in old skin? A corpus investigation of L1 syntactic transfer in learner language. In G. Aston, S. Bernardini, & D. Stewart (Eds.), *Corpora and language learners* (pp. 67–87). Philadelphia/Amsterdam: John Benjamins.

Briggs, J., Di Donato, R., Clyde, M., & Vansant, J. (2008). *Workbook to accompany Deutsch, Na Klar!: An introductory German course.* New York: McGraw-Hill.

Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Brown, S. (2007). Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora. *ReCALL, 19,* 307–328.

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, V. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21–46). Philadelphia/Amsterdam: John Benjamins.

Byrnes, H. (2009). Emergent L2 German writing ability in a curricular context: A longitudinal study of grammatical metaphor. *Linguistics and Education, 20,* 50–66.

Byrnes, H., Maxim, H., & Norris, J. M. (2010). Realizing advanced foreign language writing development in collegiate education: Curricular design, pedagogy, assessment [Monograph]. *Modern Language Journal, 94* (s1).

Byrnes, H., & Sinicrope, C. (2008). Advancedness and the development of relativization in L2 German: A curriculum-based longitudinal study. In L. Ortega & H. Byrnes (Eds.), *The longitudinal study of advanced L2 capacities* (pp. 109–138). New York: Routledge/Taylor & Francis.

Chavez, M. (2011). German grammar in the students' words: The *essentialization* of German grammar by American college-level learners. *Die Unterrichtspraxis [Teaching German], 44,* 83–97.

Clahsen, H. (1985). Profiling second language development: A procedure for assessing L2 proficiency. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 283–331). Clevedon, UK: Multilingual Matters.

Coombs, V. M. (1986). Syntax and communicative strategies in intermediate German composition. *Modern Language Journal, 70,* 114–124.

Cooper, T. C. (1976). Measuring written syntactic patterns of second language learners of German. *Journal of Educational Research, 69,* 176–183.

Crystal, D., Fletcher, P., & Garman, M. (1976). *The grammatical analysis of language disability: A procedure for assessment and remediation.* London: Edward Arnold.

Dykstra–Pruim, P., & Redmann, J. (2011). *Schreiben lernen: A writing guide for learners of German.* New Haven, CT: Yale University Press.

Ellis, R. (1994). *The study of second language acquisition.* Oxford: Oxford University Press.

Ellis, R., & Yuan, F. (2005). The effects of careful within-task planning on oral and written task perfor-mance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 167–192). Philadelphia/Amsterdam: John Benjamins.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18,* 299–323.

Garretson, G., & O'Connor, M. C. (2007). Between the Humanist and Modernist: Semi-automated analysis of linguistic corpora. In E. Fitzpatrick (Ed.), *Corpus linguistics beyond the word: Corpus research from phrase to discourse* (pp. 87–106). New York: Rodopi.

Granger, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation. In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 13–32). Philadelphia/Amsterdam: John Benjamins.

Granger, S., Kraif, O., Ponton, C., Antoniadis, G., & Zampa, V. (2007). Integrating learner corpora and natural language processing: A crucial step towards reconciling technological sophistication and pedagogical effectiveness. *ReCALL, 19,* 252–268.

Granger, S., & Rayson, P. (1998). Automatic profiling of learner texts. In S. Granger (Ed.), *Learner English on computer* (pp. 119–131). New York: Longman.

Halliday, M. A. K., & Martin, J. R. (1993). *Writing science: Literacy and discursive power.* Bristol, PA: The Falmer Press.

Hawkins, J., & McCarthy, M. (2010). Introduction to *English Profile Journal. English Profile Journal, 1,* 1.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics, 30,* 461–473.

Hunston, S. (2010). How can a corpus be used to explore patterns? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 152–166). New York: Routledge.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels* (NCTE research report No. 3). Champaign, IL: National Council of Teachers of English.

Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry, 5,* 117–136.

Kemper, S., Thompson, M., & Marquis, J. (2001). Longitudinal change in language production: Effects of aging and dementia on grammatical complexity and propositional content. *Psychology and Aging, 16,* 600–614.

Krzeszowski, T. (1990). *Contrasting languages: The scope of contrastive linguistics.* Berlin: Mouton de Gruyter.

Larsen–Freeman, D., & Cameron, L. (2008). Research methodology on language development from a complex systems perspective. *Modern Language Journal, 92,* 200–213.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics, 15,* 474–496.

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly, 45,* 36–62.

Lüdeling, A., Walter, M., Kroymann, E., & Adolphs, P. (2005). Multi-level error annotation in learner corpora. *Proceedings of Corpus Linguistics 2005*, Birmingham, UK.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. In Transcription format and programs (3rd ed., Vol. I). Mahwah, NJ: Lawrence Erlbaum.

Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development: Quantification and assessment*. Houndmills, UK: Palgrave Macmillan.

Maxim, H. (2011, March). *Establishing a curricular trajectory: A socio-semiotic perspective on text selection and sequencing*. Paper presented at the Annual Conference of the American Association for Applied Linguistics, Chicago.

Meisel, J. M., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition, 3*, 109–135.

Meunier, F. (2010). Learner corpora and English language teaching: Checkup time. *Anglistik: International Journal of English Studies, 21*, 209–220.

Meunier, F., & de Mönnink, I. (2001, May). *Assessing the success rate of EFL learner corpus tagging*. Paper presented at the Annual Conference of International Computer Archive of Modern and Medieval English, Louvain-la-Neuve, Belgium.

Meurers, W. D., & Müller, S. (2009). Corpora and syntax. In A. Lüdeling & M. Kytö (Eds.), *Corpus linguistics* (pp. 920–933). Berlin: Mouton de Gruyter.

Myles, F. (2008). Investigating learner language development with electronic longitudinal corpora: Theoretical and methodological issues. In L. Ortega & H. Byrnes (Eds.), *The longitudinal study of advanced L2 capacities* (pp. 58–72). New York: Routledge/ Taylor & Francis.

Nesselhauf, N. (2004). Learner corpora: Learner corpora and their potential for language teaching. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 125–152). Philadelphia/Amsterdam: John Benjamins.

Norris, J., & Ortega, L. (2002, September). *Using CHILDES for the analysis of L2 German data from the GUGD Multiple Literacies FL curriculum*. Workshop presented at the German Department, Georgetown University, Washington, DC.

Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30*, 555–578.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*, 492–518.

Ortega, L. (2012). Interlanguage complexity: A construct in search of theoretical renewal. In B. Szmrecsanyi & B. Kortmann (Eds.), *Linguistic complexity: Second language acquisition, indigenization, contact* (pp. 127–155). Berlin: Mouton de Gruyter.

Ortega, L., & Byrnes, H. (2008). Theorizing advancedness, setting up the longitudinal research agenda. In L. Ortega & H. Byrnes (Eds.), *The longitudinal study of advanced L2 capacities* (pp. 281–300). New York: Routledge/Taylor & Francis.

Ortega, L., & Iberri–Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics, 25*, 26–45.

Ortega, L., & Sinicrope, C. (2008). *Novice proficiency in a foreign language: A study of task-based performance profiling on the STAMP test*. (Technical report). University of Oregon, Center for Applied Second Language Studies.

Pienemann, M., Johnston, M., & Brindley, G. (1988). Constructing an acquisition-based procedure for second language assessment. *Studies in Second Language Acquisition, 10*, 217–243.

Polio, C. (2001). Research methodology in second language writing research: The case of text-based studies. In T. Silva & P. K. Matsuda (Eds.), *On second language writing* (pp. 91–115). Mahwah, NJ: Lawrence Erlbaum.

Reid, J. (1992). A computer text analysis of four cohesion devices in English discourse by native and nonnative writers. *Journal of Second Language Writing, 1*, 79–107.

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics, 45*, 237–257.

Robinson, P., Cadierno, T., & Shirai, Y. (2009). Time and motion: Measuring the effects of the conceptual demands of tasks on second language speech production. *Applied Linguistics, 30*, 533–554.

Robinson, P., & Ellis, N. C. (2008). Conclusion: Cognitive linguistics, second language acquisition and instruction: Issues for research. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 489–528). New York: Routledge/Taylor & Francis.

Santorini, B. (1990). *Part-of-speech tagging guidelines for the Penn Treebank Project* (Technical report, 3rd Revision, 2nd Printing). Technical Report, Department of Computer and Information Science, University of Pennsylvania.

Schiller, A., Teufel, S., Stöckert, C., & Thielen, C. (1999). *Guidelines für das Tagging deutscher Textcorpora mit STTS [Guidelines for tagging German corpora of written language with STTS]*. Technical Report. Stuttgart, Germany: Institut für maschinelle Sprachverarbeitung [Institute for Machine Language Processing].

Schmid, H. (1994). Probabilistic part-of-speech tagging using decision trees. In *Proceedings of the international conference on new methods in language processing* (pp. 44–49). Manchester, UK.

Scott, M. (2008). *WordSmith Tools version 5*. Liverpool, UK: Lexical Analysis Software.

Siegler, R. S., & Svetina, M. (2002). A microgenetic/ cross-sectional study of matrix completion: Comparing short-term and long-term change. *Child Development, 73*, 793–809.

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

Teichert, H., & Teichert, L. (2005). *Allerlei zum Lesen [Everything about reading].* Boston: Heinle/ Cengage Learning.

Tribble, C. (2010). What are concordances and how are they used? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 167–183). New York/London: Routledge.

Tschirner, E. (2005). Korpora, Häufigkeitslisten, Wortschatzerwerb [Corpora, frequency lists, acquisition of vocabulary]. In A. Heine, M. Hennig, & E. Tschirner (Eds.), *Deutsch als Fremdsprache— Konturen und Perspektiven eines Fachs [German as a foreign language—Contours and perspectives of the field]* (pp. 133–149). München, Germany: Iudicium.

van Geert, P., & van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development, 25*, 340–375.

van Rooy, B., & Schäfer, L. (2002). The effect of learner errors on POS tag errors during automatic POS tagging. *Southern African Linguistics and Applied Language Studies, 20*, 325–335.

Verspoor, M. H., de Bot, K., & Lowie, W. (Eds.). (2011). *A dynamic approach to second language development. Methods and techniques.* Philadelphia/Amsterdam: John Benjamins.

Vyatkina, N., (2012). The development of second language writing complexity in groups and individuals: A longitudinal learner corpus study. *Modern Language Journal, 96*, 572–594.

Vyatkina, N. (in press). Analyzing part-of-speech variability in longitudinal learner corpus and pedagogic corpus. *Proceedings of the 2011 Learner Corpus Research conference.* Louvain-la-Neuve, Belgium: Press universitaires de Louvain.

Wolfe–Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity.* Honolulu: Second Language Teaching & Curriculum Center, University of Hawai'i at Manoa.

APPENDIX A

Writing Tasks

TABLE A1
Writing Tasks (Adopted from Vyatkina, 2012, p. 594)

T	Semester	Origin of Tasks	Writing Task
1	1st	Briggs, Di Donato, Clyde, & Vansant (2008)	Who are you?
2			Your apartment, your friends
3			Your family
4			Your daily routine
5			Your favorite clothes OR Your luck charm OR A shopping day
6	2nd		Planning a party
7			Your last weekend
9			Your town
10			Comparison of two trips
12	3rd		Describe a person
13			You and the media OR You and technology
15	4th	Teichert & Teichert (2005)	Interpret a short story
17			Interpret a short story
19		Baseline Writing Task, adopted from Byrnes et al. (2010, p. 165)	Book review

APPENDIX B

Complexity Values by Time Point

Tables B1 and B2 present frequency values for all complexity measures used in the study. The first three lines in each table list values of general measures (sample length in words, number of clauses, mean length of clause), and the remaining lines list values for the focal-specific syntactic complexity measures. The latter are given in absolute numbers rather than ratios per clause, which were calculated by dividing these values by "total clauses" as presented in Figures 2, 3, 4, and 5.

TABLE B1
Braden: Complexity Values by Time Point (Absolute Frequencies)

	1	2	3	4	5	6	7	9	10	12	13	15	17	19
General Measures														
Total words	65	83	97	130	108	107	123	107	148	146	123	69	228	216
Total clauses	12	12	18	21	17	15	18	18	25	26	16	13	38	34
MLC	5.4	6.9	5.4	5.9	5.7	7.1	6.8	5.9	5.9	5.6	7.7	5.3	5.3	6.2
Coordinate Structures (CS)														
Predicate				1	1	1				2			1	1
Nominal		3	5	3		5	2	3	3	3	7	1	2	5
Modifier	2	3		3	4	1	2	1	3	1	2			2
Coord. clause	0	1	4	2	2	3	1	2	3	4	5	4	5	8
Total CS	2	7	9	9	7	10	5	6	9	10	14	5	8	16
Complex Nominals (CN)														
Adjective	0	2	0	0	1	0	1	13	5	4	5	1	8	9
Prep. phrase	1	2	0	3	1	3	2	5	0	2	1	1	5	6
Noun clause										1		2	8	1
Relative clause				1									2	3
Total CN	1	4	0	4	2	3	3	18	5	7	6	4	23	19
Nonfinite Verb Forms (NFV)														
Infinitive	1			2		6		4	5	5	4			3
Past Participle							17						13	
zu + Infinitive													4	
Total NFV	1	0	0	2	0	6	17	4	5	5	4	0	17	3

TABLE B2
Cassie: Complexity Values by Time Point (Absolute Frequencies)

	1	2	3	4	5	6	7	9	10	12	13	15	17	19
General Measures														
Total words	98	76	271	166	226	102	98	116	110	146	192	237	195	145
Total clauses	203	15	41	29	40	16	15	14	17	23	28	31	41	16
MLC	4.7	5.1	6.6	5.4	5.5	6.4	5.2	8.3	6.5	5.8	6.4	7.4	4.1	8.1
Coordinate Structures (CS)														
Predicate			6	2	1		1			2			4	2
Nominal	2	3	11	1	1	6		4	1	3	3	1	2	5
Modifier	1	1	1	1	3			1	1			4		1
Coord. clause	1	2	5	6	1	1	0	0	1	1	4	7	1	0
Total CS	4	6	23	10	6	7	1	5	3	6	7	12	7	8
Complex Nominals (CN)														
Adjective	2	2	4	2	5	1	3	12	1	5	3	5	8	9
Prep. phrase			1	0	3	2	0	2	2	3	6	0	1	6
Noun clause					2					2		3	11	
Relative clause					2						1		2	1
Total CN	2	2	5	2	12	3	3	14	3	10	10	8	22	16
Nonfinite Verb Forms (NFV)														
Infinitive	1			10	2	8	2			4	1	3	3	2
Past Participle			7	2	4	1	12			2	6	9	8	1
zu + Infinitive				1			3			3	2	2	3	1
Total NFV	1	0	7	13	6	9	17	0	0	9	9	14	14	4