Review Article

second language research

Second Language Research
2015, Vol. 31(1) 117–134

DOI: 10.1177/0267658314536435
slr.sagepub.com
SAGE

A simple view of linguistic complexity

Gabriele Pallotti
University of Modena and Reggio Emilia, Italy

Abstract

Although a growing number of second language acquisition (SLA) studies take linguistic complexity as a dependent variable, the term is still poorly defined and often used with different meanings, thus posing serious problems for research synthesis and knowledge accumulation. This article proposes a simple, coherent view of the construct, which is defined in a purely structural way, i.e. the complexity directly arising from the number of linguistic elements and their interrelationships. Issues of cognitive cost (difficulty) or developmental dynamics (acquisition) are explicitly excluded from this theoretical definition and its operationalization. The article discusses how the complexity of an interlanguage system can be assessed based on the limited samples with which SLA researchers usually work. For the areas of morphology, syntax and the lexicon, some measures are proposed that are coherent with the purely structural view advocated, and issues related to their operationalization are critically scrutinized.

Keywords

interlanguage analysis, lexicon, linguistic complexity, linguistic typology, morphology, quantitative linguistics, research methods, syntax

1 Introduction

Complexity is a notion that has received considerable attention in second language acquisition (SLA) studies over the last years. It has been employed both as an independent variable, referring to features making a communicative task more or less complex, or as a dependent variable, to describe aspects of linguistic production. This article will be concerned with the second meaning only, leaving aside issues of how task complexity can be defined and assessed.

Corresponding author:
Gabriele Pallotti, Department of Education and Human Sciences, University of Modena and Reggio Emilia, via Allegri 9, 42121 Reggio Emilia, Italy.
Email: pallotti.gabriele@unimore.it

Although several operationalizations of linguistic complexity have been proposed, they have often been applied with little or no reflection about their theoretical underpinnings and issues of construct validity. These shortcomings have been addressed in a number of recent publications, which attempt to clarify the theoretical status of complexity and discuss how it can be measured and operationalized (Bulté and Housen, 2012; Norris and Ortega, 2009; Pallotti, 2009). Despite these advances, the construct still poses a number of theoretical and methodological problems, mainly due to its polysemy. This article proposes a simple view of complexity, restricting its meaning to structural, formal aspects of texts and linguistic systems, and avoiding any unnecessary assumptions, such as the idea that complexity grows over time or that different aspects of complexity develop in parallel. Some measures will be proposed that are coherent with this narrow definition and that will, it is hoped, increase comparability across studies, thus favouring meta-analysis and research synthesis.

II The many meanings of complexity

Dictionaries normally list two main meanings of 'complexity':

> 1. 'composed of two or more parts' and 2. 'hard to separate, analyze, or solve' (Merriam-Webster)

In the typological discussion about the complexity of different languages, this polysemy has been acknowledged by employing terms such as 'objective vs. agent-related' (Dahl, 2004) or 'absolute vs. relative' (Miestamo, 2008) complexity. The first term refers to formal properties of the linguistic system, while the second has to do with issues of cost, difficulty, level of demand for a language user/learner. In the field of SLA, Bulté and Housen (2012: 23–24) similarly differentiate 'absolute, inherent complexity, or complexity for short' and 'cognitive complexity or simply difficulty' (see also Pallotti, 2009).

There is also a third meaning of complexity in linguistic research, which does not appear in dictionary definitions. It is closely related to difficulty and it has to do with how a linguistic structure is acquired by a first or second language learner. According to Trudgill (2001: 371), 'linguistic complexity … equates with "difficulty of learning for adults" '; others speak of 'L2 acquisition difficulty' (Szmrecsanyi and Kortmann, 2009) or 'outsider complexity' (Kusters, 2003). In SLA studies, complexity has been similarly identified with 'the capacity to use more advanced language' (Ellis, 2009: 475), thus equating complex structures with structures appearing late in L2 development.

There are thus three main meanings of 'complexity' in linguistic research:

> 1. Structural complexity, a formal property of texts and linguistic systems having to do with the number of their elements and their relational patterns; 2. Cognitive complexity, having to do with the processing costs associated with linguistic structures; 3. Developmental complexity, i.e. the order in which linguistic structures emerge and are mastered in second (and, possibly, first) language acquisition.

This polysemy poses a number of problems, as the three meanings of 'complexity' clearly refer to different constructs, analytically separated. Whether and to what extent

they are related is an empirical issue, and even if they were found to correlate very strongly, this would not mean they are three facets or names of one single construct. That they are different constructs can be seen by the asymmetry of the relationship: 'cognitive difficulty reflects rather than creates complexity' (Rescher, 1998: 17). But there might be exceptions: a Sudoku with 18 digits is structurally less complex but cognitively more complex, or difficult, than one with 25. Similarly, it has been noted that linguistic structures that are structurally more complex may be easier to produce than others, so that 'sometimes efficiency results in greater complexity' (Hawkins, 2009: 253).

The point made here is not that these relationships should not be studied, let alone that they do not exist. The point is mainly terminological: in order to assess the relationships between two or three constructs it is advisable to call them by different names. Otherwise one ends up with statements such as 'complex$_1$ structures are often more complex$_2$ and complex$_3$' instead of the much more perspicuous 'complex structures are often more difficult and acquired late'; or 'this structure is complex$_3$ because it is complex$_1$ and complex$_2$' instead of 'this structure is acquired late because it is complex and difficult'. The use of subscripts, as I use here, adds a bit of clarity, but it is still an unnecessary complication when three different terms are available.

In order to avoid this polysemy, this article advocates a simple view of complexity, treating it as a purely descriptive category, limiting its use to structural complexity and excluding from its definition any theoretical assumption about when, how and why it increases or remains constant. The term 'structural' is preferred to 'objective' and 'absolute' because these seem to imply the existence of an objective, theory-free description of linguistic facts (Kusters, 2008: 8). This is clearly not the case: even the most factual descriptions always contain a theoretical dimension; there are probably very few, if any, descriptions of a linguistic fact on which all linguists would agree; and what is complex according to one theory may not be so according to another (Bulté and Housen, 2012: 26). Still, some descriptions appear to be more theory-laden than others, using labels that are specific to a certain school, approach or author. In this article we will employ terms from 'Basic Linguistic Theory' (Dryer, 2006), which does not imply that such terms are objective or neutral, but simply that they should be understood by most readers.

There are two main ways of defining structural complexity (Dahl, 2009: 51). One is called Kolmogorov complexity and it is represented by the length of the shortest description that is needed to represent a string of symbols. From this point of view, these three expressions display increasing complexity: *hahaha* (= 3 × *ha*; 4 symbols) < *byebye* (= 2 × *bye*; 5 symbols) < *pardon* (= *pardon*; 6 symbols). The last string cannot be compressed in any way; it is like a sequence of random characters, which represents the highest complexity. However, the idea that complete lack of order has the highest complexity is somewhat counterintuitive. To address this issue, Gell–Mann proposed a measure called Effective Complexity, which corresponds to the length of description required to specify the set of regularities in a string. From this point of view, the expressions above would be ranked differently: *pardon* (0 complexity) < *hahaha* (= 3 × ha) < *byebye* (= 2 × *bye*). In other words, the complexity of a monotonous string and that of a completely random string of characters are the same, i.e. zero.

A second terminological note concerns the level at which the notion of complexity is applied. I propose to distinguish between 'system complexity', i.e. the complexity of a

whole linguistic system, with all of its elements and rules (Saussurean *langue*), and 'text complexity', i.e. the complexity of a given piece of discourse (*parole*).

I will also speak of 'grammatical complexity' and 'stylistic complexity'. Grammatical complexity has to do with the complexity of linguistic rules that must be followed to produce grammatical sentences. For example, the rules for constructing a subordinate clause in German are more complex than those for producing one in English, and in order to produce a grammatical German subordinate clause one must follow these rules. On the other hand, the degree of subordination in a text is not prescribed by any specific rule: it might at most be influenced by some culture-specific rhetorical patterns, but these allow for ample individual stylistic variation and even the most extreme deviations will never be deemed ungrammatical, but rather awkward or uncommon. Hence stylistic complexity is always, at least to some extent, a matter of speaker's or writer's choice, while grammatical complexity is not, with the exception of a few cases of variable rules.

III Operationalizing the simple view in SLA research

Any operational definition of a construct must be grounded on a conceptual definition. As regards complexity in general, we will follow Rescher (1998: 1), who defines it as 'a matter of the number and variety of an item's constituent elements and of the elaborateness of their interrelational structure'.

Applying such a purely structural definition to linguistic facts yields statements like the following: 'A definition of grammatical complexity can be based on the usual understanding of a complex system as one consisting of many different elements each with a number of degrees of freedom' (Nichols, 2009: 111); 'complexity should … be defined, to put it in the most general terms, as the number of parts in a system or the length of its description' (Miestamo, 2008: 26); 'the number of discrete components that a language feature or a language system consists of, and the number of connections between the different components' (Bulté and Housen, 2012: 24).

All these definitions explicitly mention a 'system'. It is hard enough, even for native languages with published grammars and large corpora produced by millions of people, to exactly define what a system's boundaries are and how its overall complexity can be assessed (Deutscher, 2009). This becomes virtually impossible for interlanguages, whose inherent variability and instability do not allow one to provide an exhaustive list of clear-cut features of such 'systems'. Hence, SLA researchers have to limit themselves to assessing the complexity of learners' texts, and whatever conclusion about interlanguage systems would be a more or less warranted inference from the few acts of *parole* observed in such texts.

In this article, a text's structural complexity will be defined in general as the number of different elements and their interconnections (i.e. their systematic, organized relationships), which both produce a longer description of the text's structure. Description length thus works as an additional check, as it should automatically increase in the presence of many different and highly interconnected elements.

The problem is operationalizing this general definition with specific measures. In what follows, some proposals will be made that are compatible with the simple view advocated here. Space constraints allow us to provide just a sketch of these measures, for

each of which an entire paper would be required to tackle all issues of construct definition and operationalization. The discussion will be as language-neutral as possible, and some details for operationalizing each measure will have to be worked out specifically for each target language. Appendix 1 (published online on the journal's and the author's websites; slr.sagepub.com; www.gabrielepallotti.it) provides a concrete illustration of how the proposed measures can be applied to an interlanguage text.

1 Morphological complexity

Morphological complexity has not been investigated in many SLA studies. Some of the proposed measures include frequency of tensed forms, number of different verb forms, variety of past tense forms (Bulté and Housen, 2012), all of which deal with verbal morphology. Here a wider-ranging definition will be provided, which, after some emic fine-tuning, can be applied to a variety of typologically diverse languages. However, our discussion, too, will be restricted to the inflectional morphology of individual words, deliberately ignoring derivational morphology and the complexity produced by agreement phenomena.

Inflectional morphology concerns the relationships between the forms that lexemes can take (i.e. their 'exponents') and semantic or syntactic features such as gender, number, case, person and tense. The notion of morpheme may be adequate to describe concatenative morphological processes, but it runs into various problems when the morphologically relevant form is produced by reduplication, stem modification or even by no change at all. Many morphologists thus prefer to speak more generally of morphological operations or processes, only some of which can be described in terms of 'adding grammatical morphemes to a lexical base', and this is what will be done here, too.

a Number of exponents. In order to assess a text's morphological complexity one may count, for each word class (nouns, verbs, adjectives, etc), its exponents, i.e. the forms taken by lexemes to express grammatical categories and functions. To do so, one must be able to identify abstract schemas of the relationships between stems and the morphological processes modifying them, which implies identifying, for each inflected form, the underlying stem. With a concatenative morphological process, the distinction between base and affix is relatively easy, as in English *book-s*. With non-concatenative morphological processes things get more complicated. In such cases, either the same lexeme is found in different syntactico-semantic contexts (which can occur with highly structured forms of data elicitation), or one will have to resort to what is known of the L2, so that a form like *Bücher* in German will be seen as an operation on the base *Buch* involving umlaut on the stem's vowel plus a suffix *-er*.[1]

Periphrastic morphemes, like *be V-ing* or *have V-ed* in English, will be counted as one single operation. Allomorphs, such as different ways of forming the past tense in English (*arriv-ed, went, took*) will be counted separately, as they increase the number of elements in the system and thus its global complexity.

One should also try to distinguish between morphological and phonological variations – i.e. between allomorphs and allophones – although this may not always be easy.

For example, if a learner produces two plural nouns like *toy-s* and *boy-sh*, it may be sensible to consider these as two tokens of the same morphological exponent, rather than two different exponents. Evidence for such a conclusion may come from an analysis of the learner's phonology, which may show that [s] and [ʃ] tend to be realized as allophones in free, non morphologically-conditioned, variation, for example in words such as *moush* (instead of *mouse*) or *Rolls Roysh*.

This index of morphological variation is a type-based measure, and is thus sensitive to text length in the same way as lexical variety indices are. A solution to this problem is to count the number of different exponents for subsamples of 10 inflected forms[2] each, and then calculate their mean value. However, the same exponents may either be repeated again and again in various subsamples, or different samples may contain different sets of exponents each. In the latter case, the text's global morphological complexity would be higher, as there is variation both within and across subsamples. This notion can be operationalized by computing the dissimilarity between each pair of subsamples, to arrive at an average value. This value multiplied by the average number of exponents per subsample gives an index of morphological complexity, whose theoretical values range from zero (all subsamples exhibit only one and the same form, hence $1 \times 0 = 0$) to 200 (each subsample contains 10 different forms, and all pairs of subsamples are completely different from one another, with an average dissimilarity index of 20, hence $10 \times 20 = 200$).

b Questionable measures. An alternative definition of morphological complexity may be the number of form–function relationships, or 'morphological patterns' (Haspelmath and Sims, 2010). These amount to the cells in a paradigm, displaying all the systematic relationships between some semantico-syntactic features and their exponents. In principle, one may say that a text is more or less morphologically complex not only because it contains a certain number of exponents, but also because these encode a more or less numerous set of features. For example, in a language like German nouns are inflected for gender, number and case, while in one like English they are inflected for number only. Thus the basic paradigm for German nouns contains eight cells, while that for English nouns only has two. Similarly, verbs in different languages may be inflected for person, tense, aspect, mood, honorificity etc.

In a native language, with a stable and well-documented grammar, it would be relatively easy to describe, for each inflected form, its grammatical function(s), as is routinely done in morphemic transcriptions. In this case, a morphological complexity index could be computed along the lines just illustrated; for example, for any 10-noun subsample in German, one would count not just the number of different forms (*-es, -e, -en*), but also the number of different form–function strings like *es*:gen.sg, *e*:nom.pl, *e*:acc.pl, *en*:dat.pl, *e*:gen.pl, etc.

Interlanguages, however, are much more unstable and variable systems, so that it is quite hard to say exactly what grammatical functions are expressed by a given form. For example, can one be sure that a sentence like *John playing in park* encodes progressive aspect, and not just a generic present tense (or not even that, in certain initial varieties where *-ing* forms may be used as default in all temporal and aspectual contexts)? And can one say that it encodes indicative mood, in a system where there seem to be no traces

of subjunctive, conditional or other moods? For these reasons, it is very difficult, especially for initial learner varieties, to state exactly what functions are expressed by a given form, so that using them as a base for quantifying a complexity score would be highly problematic.

c Measures incompatible with the simple view. Definitions of morphological complexity in terms of processing cost or developmental stages are incompatible with the purely formal view advocated here. Hence, treating 3.sg *-s* as more complex than *-ing* on English verbs because the former tends to appear later in interlanguage development would amount to conflating complexity with difficulty of acquisition, which is at odds with the simple view presented here.

2 Syntactic complexity

Syntactic structures in a text can be more or less complex depending on the number of constituents and the number of combinations they may take. While the former aspect has been widely used in SLA research, the latter poses several problems. The first is defining what exactly syntactic combinations are and whether the category should include, conservatively, only categorical rules or also tendencies, preferences, pragmatically-motivated choices. Second, the characterization of some syntactic patterns as more complex than others will often be theory laden and may not be shared by other researchers. In sum, studying grammatical syntactic complexity in interlanguage productions may prove to be quite difficult, though not unfeasible in principle.

Stylistic syntactic complexity, on the other hand, can be assessed by looking at the number of interconnected constituents in a structure, which is the principle behind three measures such as length of phrase, number of phrases per clause and number of clauses per unit.

a Length of phrase. Our operational definition (largely based on Van Valin and LaPolla, 1997), sees the verb (or, more generally, the predicating element) as the nucleus of the clause, so that the phrases to be considered are only those surrounding it, i.e. arguments and adjuncts, with no reference to a 'verb phrase' consisting of V+NP. Furthermore, counts only refer to immediate constituents of the clause, which means that phrases embedded in other phrases will not be considered. Hence, in a clause like *The man saw the boy with the red hat after dinner* we will count three phrases (*the man; the boy with the red had; after dinner*), consisting of 2, 6 and 2 words, respectively.

The first obvious problem is the definition of 'word', which implies, for instance, deciding whether units such as *sports car* (or *sportscar*) should be counted as one or two words. This issue cannot be tackled here, and there is a vast literature on the subject (for a review see Booij, 2012). In any case, segmentation principles should be made explicit and followed coherently.

The second issue concerns coordinated phrases. In a clause like *The man saw the girl and the boy* is one to count two or three NPs? In our operationalization, it is suggested that *the girl and the boy* should be treated as a single constituent, occupying one slot in the clause's argument structure.

b Number of phrases per clause. Having decided how to define phrases, counting their number in a clause should prove relatively straightforward, but it implies defining clauses, which is the next point.

c Number of clauses per unit. This is one of the most employed measures of syntactic complexity in SLA research, but it requires a clear definition of clauses and superordinate units. At the lower end, one needs to draw a line between multiple clauses and complex predicative constructions, such as *keep trying, make stop, begins to rain*. The criterion proposed here is to treat as a complex predication all the cases where (1) there is strong syntactic integration between the two predicates, evidenced for example by their obligatorily sharing grammatical features or arguments, or by one of them losing some sentential properties such as finiteness; (2) the two predicates do not denote different states of affairs. Hence, in the three examples above, *trying, stop* and *to rain* are all non-finite verbs, and one verb semantically modifies the other without introducing another action or state: *keep* and *begins* indicate aspectual meanings, *make* produces a causative reading. Even a sentence like *Sam asked Fred to leave* should be considered a single clause with two 'cores'; Van Valin and LaPolla (1997: 447ff) indicate a number of syntactic reasons for this, but also from a semantic point of view *ask to leave* does not denote two different actions, but rather specifies the propositional content of the act of asking. On the contrary, *Fred studies law to become a lawyer* clearly indicates two different events.

At the higher end, the decision must be made whether several clauses belong to a single overarching unit or whether they are independent of each other. Both the T-Unit (Hunt, 1965) and the AS-Unit (Foster et al., 2000) comprise a main clause plus all of its dependent clauses, which means that coordinate clauses are interpreted as new units. This stands to reason, as coordination is the weakest form of clause linkage, consisting in the mere sequencing of clauses retaining their full structural independence. From a cross-linguistic perspective it is not always easy to draw a neat line between coordination and subordination, and not just one, but several, continua for a variety of 'clause linkage' strategies have been identified (Lehmann, 1988). Again, there is no space here to tackle this complex issue of syntactic typology, but for most European languages traditional criteria still seem to be valid and it is not normally an issue whether two clauses are in a coordination or subordination relationship.

d Number of word-order patterns. As we said, establishing what counts as a 'pattern' in an interlanguage sample is quite challenging, and even establishing the basic word order of a native language is often not so easy (Dryer, 2007). If one opts for a Kolmogorov definition of complexity, then a text exhibiting ever-varying syntactic patterns would be seen as extremely complex. Following a Gell–Mann definition of complexity, one should instead count only the number of *regular* patterns, which requires criteria for setting patterns apart from random variation, and also for identifying elements in a pattern (e.g. 'subject', 'negative clause', 'adverb'). If these methodological challenges are met, then a text showing a variety of regular syntactic patterns would be considered more complex than one with all sentences and clauses following one or two patterns or where no pattern is discernible at all. However, this measure will not be addressed in the analysis in Appendix 1.

e *Questionable measures.* Mean length of clause could be taken as a rough indicator of clausal and phrasal complexity, but it is a hybrid measure depending on the number of phrases in a clause and their length and thus becomes redundant if these two measures are computed separately. Likewise, subordination ratio is unnecessary if one knows the number of clauses per unit, provided that these macro-units are defined to comprise only the main and subordinate clauses. An index of coordination may indicate how many clauses that were excluded from the clauses/unit index actually form units of another kind, i.e. 'sentences' or 'chains'. However, operationalizing coordination in order to distinguish it from the almost unlimited range of textual cohesive devices means essentially identifying it with clause linkage through coordinating conjunctions, which may be appropriate for most European languages but may run into problems for other linguistic systems.

f *Measures incompatible with the simple view.* All measures representing syntactic complexity in terms of structures' difficulty, sophistication and acquisitional timing are at odds with the simple view proposed here. Ortega (2012), for example, notes that long noun phrases containing nominalizations, modifiers and prepositional phrases are characteristic of the prose of advanced learners, and thus emerge after complexification via subordination. If this empirical remark were translated into a definition of complexity whereby a complex phrase should be considered to be 'more complex' than a complex T-Unit, it would be incompatible with the purely structural view advocated here.

3 Lexical complexity

Complexity at the lexical level can be basically seen as a matter of the number of components of the lexical system. Since it is virtually impossible to produce an exhaustive count of the lexemes known by an individual, analysis will once again be limited to text realizations. It will also focus on lexical forms, leaving aside considerations about semantic complexity. A word may be semantically complex because of its high polysemy, or because of the constraints on its co-occurrence with other words. While these aspects are worth noting from a theoretical point of view, they seem to be impervious to practical operationalization in production data.

The result is that, under the simple view advocated here, lexical complexity can be operationalized essentially in terms of diversity. A text with a wide variety of lexemes will be said to be more complex than one where the same few words are repeated over and over.

a *Lexical diversity measures.* Lexical diversity at the text level can be gauged basically by looking at type/token ratios, with subsequent refinements proposed to overcome the effects of text length, such as the Guiraud index and D. None of them is without problems and there is currently no consensus as regards the best index of lexical diversity (McCarthy and Jarvis, 2010). In the example given in Appendix 1, lexical diversity has been computed with the D index, using the vocd programme contained in CLAN (Malvern et al., 2004).

b Questionable measures. Another possible operational definition of lexical complexity could refer to the complexity of individual words. A compound or derived lexeme appears to be inherently more complex than a primitive one. One could thus count the number of lexemes in a word (De Groot, 2008: 210) or the number of derivational affixes (which, however, Bulté and Housen, 2012, consider a measure of grammatical complexity). While this approach appears to be both theoretically motivated and practically feasible, it is certainly time-consuming, and one wonders whether, in the end, its outcome would be significantly different from that of a simple diversity analysis; this is an empirical question needing further investigation.

c Measures incompatible with the simple view. Lexical density – the proportion of lexical words to function words or to all words in a text – has been proposed as an indicator of complexity, although it is not clear whether a higher rate of lexical words should denote more or less complexity. Whatever the interpretation, it rests on the idea that a certain subset of the lexicon is more complex because it is used by more advanced learners, as there are no clear reasons why, from a purely structural point of view, lexical words should be more or less complex than function words. Similarly, indices of lexical sophistication, like the percentage of rare or difficult words, may be valid indicators of development, but they do not directly tap structural complexity; from a structural point of view, a rare word like *tar* is not in itself more complex than a common one like *car*.

IV Conclusions

This article has presented a small set of measures of linguistic complexity that can be seen as coherent with one another, as they all tap a single construct, i.e. structural complexity. This small and coherent inventory can form the base for future investigations, facilitating meta-analysis and research synthesis in an area that has seen too many inconclusive results due to the plethora of measures employed in different studies. The proposed measures can also be employed in first language acquisition research and in comparative linguistics. The various measures can be applied individually, as one might be interested in studying a certain type of complexity only, e.g. lexical or syntactic, or they might be used together, to provide a global estimate of a text's complexity.

However, the step from individual measures to a single complexity index is not straightforward. It is certainly not possible to simply add up the different scores, as they come from different scales with different magnitudes – a value of '6' may be very high on one scale (e.g. words per phrase) and quite low on another (e.g. number of verb forms in a language with rich verbal paradigms). A possible solution would be ranking the scores for each scale in a given corpus (e.g. all the texts produced in a study on complexity variation across tasks or developmental levels), and then assign a score of 4 to values in the highest quartile, 3 for the second, 2 for the third and 1 for the lowest quartile. In this way, a text with various measures of complexity all falling in the first and second quartile would receive a global complexity score higher than that of a text with all or most measures falling in the third or fourth quartile. Surely, this implies that quartiles may be defined differently in different studies, depending on the score range of each sample. Meta-analysis would still be possible, as effect sizes can be computed even if the

underlying absolute scores vary. A more general scale might be computed by taking into consideration several samples from a number of native speakers and from learners with varied backgrounds and competence levels, so that such an extended data sample could be used to calculate a distribution on which language-specific quartile cut-off points can be identified. This might be a direction for future research in this area.

A final question concerns the validity of the proposed measures. The perspective advocated here explicitly excludes from the complexity construct notions such as difficulty or development, so that showing that complexity measures increase over time or with a higher or lower cognitive load cannot be taken as an indicator of their validity. Under this purely formal, structural definition, complexity is to be seen as an 'observable attribute' rather than as a 'theoretical construct' (Kane, 2001), and its validity should be assessed mainly in terms of internal consistency and reliability of observations. Hence a measure that keeps varying across observers and observations, even when applied to the very same text, would be invalid in the same sense as an elastic ruler is not a valid instrument for measuring objects, at least in our understanding of measurement (Wittgenstein, 1956: I.5). If any reliable correlation were to be found between structural linguistic complexity and development, task conditions or anything else, this should be considered more a validation of the theory postulating such a relationship rather than a validation of the complexity measure itself.

Acknowledgements

I wish to thank Alex Housen, Cornelia Lahmann, Adriano Colombo and two anonymous reviewers for their comments on previous versions of this article; all remaining errors are my own.

Declaration of Conflicting Interest

The author declares that there is no conflict of interest.

Funding

This research was partially funded by the Italian Ministry of Education and Research under the PRIN 2009 programme.

Notes

1. Identifying what counts as the stem is not always straightforward even in native languages, and the exponent of a morphological operation is sometimes itself the alternation of different stems.
2. The denominator must be inflected forms (e.g. verbs or nouns) and not number of words, as the number of verbs or nouns may vary in different N-word samples.

References

Booij G (2012) *The grammar of words*. Oxford: Oxford University Press.

Bulté B and Housen A (2012) Defining and operationalising L2 complexity. In: Housen A, Kuiken F, and Vedder I (eds) *Dimensions of L2 performance and proficiency: Investigating complexity, accuracy and fluency in SLA*. Amsterdam/Philadelphia: Benjamins, 21–46.

Dahl Ö (2004) *The growth and maintenance of linguistic complexity*. Amsterdam: Benjamins.

Dahl Ö (2009) Testing the assumption of complexity invariance: The case of Elfdalian and Swedish. In: Sampson G, Gil D and Trudgill P (eds) *Language complexity as an evolving variable*. Oxford: Oxford University Press, 50–63.

De Groot C (2008) Morphological complexity as a parameter of linguistic typology: Hungarian as a contact language. In: Miestamo M, Sinnemäki K, and Karlsson F (eds) *Language complexity: Typology, contact, change*. Amsterdam: Benjamins, 191–216.

Deutscher G (2009) Nominalization and the origin of subordination. In: Sampson G, Gil D, and Trudgill P (eds) *Language complexity as an evolving variable*. Oxford: Oxford University Press, 199–214.

Dryer M (2006) Descriptive theories, explanatory theories, and Basic Linguistic Theory. In: Ameka F, Dench A, and Evans N (eds) *Catching language: The standing challenge of grammar writing*. Berlin: de Gruyter, 207–34.

Dryer M (2007) Word order. In: Shopen T (ed.) *Language typology and syntactic description: Volume 1: Clause structure*. Cambridge: Cambridge University Press, 61–131.

Ellis R (2009) The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics* 30: 474–509.

Foster P, Tonkyn A, and Wigglesworth G (2000) Measuring spoken language: A unit for all reasons. *Applied Linguistics* 21: 354–75.

Haspelmath M and Sims A (2010) *Understanding morphology*. 2nd edition. London: Hodder Education.

Hawkins J (2009) An efficiency theory of complexity and related phenomena. In: Sampson G, Gil D, and Trudgill P (eds) *Language complexity as an evolving variable*. Oxford: Oxford University Press, 252–68

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

Kane MT (2001) Current concerns in validity theory. *Journal of Educational Measurement* 38: 319–42.

Kusters W (2003) *Linguistic complexity: The influence of social change on verbal inflection*. Leiden: LOT.

Kusters W (2008) Complexity in linguistic theory, language learning and language change. In: Miestamo M, Sinnemäki K, and Karlsson F (eds) *Language complexity: Typology, contact, change*. Amsterdam: Benjamins, 3–22.

Lehmann C (1988) Towards a typology of clause linkage. In: Haiman J and Thompson SA (eds) *Clause combining in grammar and discourse*. Amsterdam: Benjamins, 181–225.

Malvern D, Richards B, Chipere N, and Durán P (2004) *Lexical diversity and language development*. Houndmills: Palgrave MacMillan.

McCarthy PM and Jarvis S (2010) MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods* 42: 381–92.

Miestamo M (2008) Grammatical complexity in a cross-linguistic perspective. In: Miestamo M, Sinnemäki K, and Karlsson F (eds) *Language complexity: Typology, contact, change*. Amsterdam: Benjamins, 23–41.

Nichols J (2009) Linguistic complexity: A comprehensive definition and survey. In: Sampson G, Gil D, and Trudgill P (eds) *Language complexity as an evolving variable*. Oxford: Oxford University Press, 110–25.

Norris J and Ortega L (2009) Measurement for understanding: An organic approach to investigating complexity, accuracy, and fluency in SLA. *Applied Linguistics* 30: 555–78.

Ortega L (2012) Interlanguage complexity: A construct in search of theoretical renewal. In: Kortmann B and Szmrecsanyi B (eds) *Linguistic complexity: Second language acquisition, indigenization, contact*. Berlin: de Gruyter, 127–55.

Pallotti G (2009) CAF: Defining, refining and differentiating constructs. *Applied Linguistics* 30: 590–601.

Rescher N (1998) *Complexity: A philosophical overview*. London: Transaction Publishers.

Szmrecsanyi B and Kortmann B (2009) Between simplification and complexification: Non-standard varieties of English around the world. In: Sampson G, Gil D, and Trudgill P (eds) *Language complexity as an evolving variable*. Oxford: Oxford University Press, 64–79.

Trudgill P (2001) Contact and simplification. *Linguistic Typology* 5: 371–74.

Van Valin RD and LaPolla RJ (1997) *Syntax: Structure, meaning, and function*. Cambridge: Cambridge University Press.

Wittgenstein L (1956 [1978]) *Remarks on the foundations of mathematics*. Revised edition. von Wright GH, Rhees R and Anscombe GEM (eds); translated by GEM Anscombe. Oxford: Blackwell.

Appendix 1. Analysis of an interlanguage sample.

		w/ phrase	w/ phrase	w/ phrase	phr/cl	phr/cl	cl/T-U	V infl	V infl	N infl	N infl	N infl
*CHI:	a boy and a dog found a frog.	5	2		2		1	ound[a]		Ø	Ø	Ø
*INV:	very good .											
*CHI:	at night (…) the frog go away.	2	2	1	3		1	Ø		Ø	Ø	
*CHI:	and[b] :: the [/] the [/] (.) the they [?] [/] (.) uh uh uh they [?]											
	[/] they [?] other day when [/] when she go to see the frog :: there	3	1	2	3		2	Ø	Ø	Ø	Ø	
	(i)s no frog .	2			1			is		Ø		
*CHI:	and he go <to> [//] (…) uh for find it .	1	1		2		1	Ø	Ø			
*CHI:	and the dog (.) uh push his head in (…) a jar .	2	2	3	3			Ø		Ø	Ø	Ø
*CHI:	and then :: when the boy go to the window :: (…) and called the	2	3	2	2	1	3	Ø	ed	Ø	Ø	
	frog :: the dog fell down . ::	2	1		2			fell		Ø	Ø	
*CHI:	and [/] and then + …											
*INV:	was he dead ?											
*CHI:	no .											
*CHI:	and he was save .	1	1		2		1	was				
*INV:	uhuh .											
*CHI:	and the jar it was broken .	2	1	1	3		1	was		Ø		
*CHI:	and then the boy and the [/] the dog go to a forest .	5	3		2		1	Ø		Ø	Ø	Ø
*CHI:	and the dog found (…) <a house> [/] a house of bees .	2	4		2		1	ound		Ø	Ø	s
*CHI:	and the boy found a hole .	2	2		2		1	ound		Ø	Ø	
*CHI:	in [/] in the hole it was uh a mouse	3	1	2			1	was		Ø	Ø	
*CHI:	and he uh uh uh uh (…) + …	1										
*INV:	what did the mouse do?											
*CHI:	uh (..) uh uh he uh uh + …	1										
*INV:	was he angry (..) the mouse ?											
*CHI:	no (…)											
*CHI:	yes yes uh + …											
*INV:	did he tell him where the frog was ?											
*CHI:	no .											
*CHI:	uh uh uh then the boy climb up a tree.	2	3		2		1	Ø		Ø	Ø	

Appendix 1. (Continued)

		w/ phrase	w/ phrase	w/ phrase	phr/cl	phr/cl	cl/T-U	V infl	V infl	N infl	N infl	N infl
*CHI:	and (…) he found a hole in the tree.	1	5		2		1	ound		Ø	Ø	
*INV:	uhuh.											
*CHI:	the dog uh has made uh fell the [/] the house of bees .	2	4		2		1	made	fell	Ø	Ø	s
*CHI:	and the bees are <very> [/] uh very bad with the dog .	2	5		2		1	are		s	Ø	
*CHI:	and in the hole of the tree come out a owl .	6	1	2	3		1	Ø		Ø	Ø	Ø
*CHI:	and the boy fell down the [/] the tree .	2	3		2		1	fell		Ø	Ø	
*CHI:	then the boy climb up a rock .	2	3		2		1	Ø		Ø	Ø	
*CHI:	and he called the frog .	1	2		2		1	ed		Ø		
*CHI:	and then out of the [/] the rock came out a deer .	4	1	2	3		1	came		Ø	Ø	
*INV:	a deer?											
*CHI:	and yes											
*CHI:	and the boy he was in [//] up his head .	2	1	3	3		1	was		Ø	Ø	
*INV:	uhuh (…) of the deer?											
*CHI:	yes											
*CHI:	and then the deer uh ran[c] [?] away .	2	1		2		1			Ø		
%com:	run?											
*CHI:	and then the boy and the dog fell [/] fell down a rock .	5	3		2		1	fell		Ø	Ø	Ø
*CHI:	and down the rock it was uh (.) water .	3	1	1	3		1	was		Ø	Ø	
*CHI:	then the dog and the boy found a [/] a dead tree .	5	3		2		1	ound		Ø	Ø	Ø
*CHI:	and then and [/] the boy and the dog look on the other part of the dead tree .	5	8		2		1	Ø		Ø	Ø	Ø
*CHI:	and there was a family of frogs .	4			1		1	was		Ø	s	
*CHI:	then uh the boy take uh uh one of the babies .	2	4		2		1	Ø		Ø	s	
*CHI:	and [/] and then he go away .	1	1		2		1	Ø				
	Mean values (for all columns)	*2.42*			*2.16*		*1.10*					

Notes. [a] *found* belongs to a small inflection class together with other forms such as *bound, ground* etc. In these cases, only one exponent (*-ound*) should be counted, not one for each verb. [b] Textual connectives, such as *and, then, and then*, do not enter in word/phrase count as they are not considered parts of clauses, but elements joining them. [c] Not counted in V morphology analysis as it isn't clear whether it is /rʌn/ or /ræn/.

General notes

The transcript follows the Chat transcription system (childes.psy.cmu.edu); CHI = Child learning L2 English; INV = Interviewer; Lines represent AS-units; clauses within AS-units are separated by :: (Foster et al 2000); Columns:

- w/phrase: number of words per phrase. Phrases are indicated by underlining. Retracings are not included in word counts.
- phr/cl: number of phrases per clause.
- cl/T-U: number of clauses per T-Unit. In spoken language, subclausal units are often produced, either as the result of planning issues, e.g. in truncated clauses like *and then he…*, or as perfectly functional pieces of discourse occurring on their own (*goodbye, thank you*), or integrated with previous talk (*Where are you going? To the pub*). Foster et al. (2000), in their definition of the AS-unit, group these sub-clausal units together with simple clauses and multi-clausal sentences. In our operationalization they will instead be treated differently and will not be considered in computing the phrases/clause and the clauses/unit ratios. In the first case, quite obviously, as they are not clauses; in the second, because a subclausal unit would give a score of 0 in the clause/unit ratio (being subclausal, such a unit would contain 0 clauses). This would produce the same score, i.e. 10, both for a text with 10 simple clauses and for one with five subclausal units and five complex sentences of two clauses each. However, the integration of clauses into hierarchically higher syntactic structures and the (dis)integration of subclausal units in discourse seem to be rather different phenomena, and it is questionable whether they can simply be added up together. Thus we take the T-unit as the superordinate unit for the clause/unit ratio, which has the further advantage of making oral and written data more comparable.
- V/infl: verb inflectional forms. Highlighting indicates that 10 verbs have been reached.
- N/infl: noun inflectional forms. Highlighting indicates that 10 nouns have been reached.

Computing the scores

Syntactic complexity

Words / phrase = 2.42
Phrases / clause = 2.16
Clauses / t-unit = 1.10

Lexical complexity

D = 16.42. The following CLAN command was used on a lemmatized version of the file, where stems and roots were separated from inflections:

vocd +t”CHI” +r6 +s”*–%%” +s”*&%%” file_name.cha –s”uh”

Morphological complexity

V1, V2, V3 … and N1, N2, N2 … stand for the first, second, third … set of 10 verbs or nouns (the last verbs and nouns, not reaching a set of 10, were discarded). For each set, the different morphological exponents are counted and averaged. Then each list is compared to all the others, to calculate a set of dissimilarity scores, which are also averaged. These two averages are multiplied, to arrive at the morphological complexity score.

Verb forms:

V1	V2	V3
ed	fell	are
is	made	came
ound	ound	ed
Ø	was	fell
	Ø	was
		Ø

Note. Mean number of different verb forms per 10 verbs: (4 + 5 + 6) / 3 = 5.

Number of unique forms in two-set comparisons:

V1V2	V1V3
5	6
ed, is, fell, made, was	is, ound, are, came, fell, was
	V2V3
	5
	made, ound, are, came, ed

Note. Mean number of unique forms in two-set comparisons: (5 + 6 + 5) / 3 = 5.33.

Verb morphological complexity: 5 * 5.33 = 26.65.

Noun forms:

N1	N2	N3	N4	N5	N6
Ø	Ø	Ø s	Ø s	Ø	Ø

Note. Mean number of different noun forms per 10 nouns: (1 + 1 + 2 + 2 + 1 + 1) / 6 = 1.33.

Number of unique forms in two-set comparisons:

N1N2	N1N3	N1N4	N1N5	N1N6
0	1 N2N3 1	1 N2N4 1 N3N4 0	0 N2N5 0 N3N5 1 N4N5 1	0 N2N6 0 N3N6 1 N4N6 1 N5N6 0

Note. Mean number of unique forms in two-set comparisons: (0 + 1 + 1 + 0 + 0 + 1 + 1 + 0 + 0 + 0 + 1 + 1 + 1 + 1 + 0) / 15 = 0.53.

Noun morphological complexity: 0.53 * 1.33 = 0.7.