

Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners

Scott A. Crossley^{a,*}, Danielle S. McNamara^{b,1}

^a Department of Applied Linguistics/ESL, Georgia State University, 34 Peachtree Street, Suite 1200, One Park Tower Building, Atlanta, GA 30303, United States

^b Department of Psychology, Arizona State University, Tempe, AZ 85287, United States

Abstract

This study examines second language (L2) syntactic development in conjunction with the effects such development has on human judgments of writing quality (i.e., judgments of both overall writing proficiency and more fine-grained judgments of syntactic proficiency). Essays collected from 57 L2 learners in a longitudinal study were analyzed for growth and scoring patterns using syntactic complexity indices calculated by the computational tool Coh-Metrix. The analyses demonstrate that significant growth in syntactic complexity occurred in the L2 writers as a function of time spent studying English. However, only one of the syntactic features that demonstrated growth in the L2 learners was also predictive of human judgments of L2 writing quality. Interpretation of the findings suggest that over the course of a semester, L2 writers produced texts that were increasingly aligned with academic writing (i.e., texts that contain more nouns and phrasal complexity), but that human raters assessed text quality based on structures aligned with spoken discourse (i.e., clausal complexity). Thus, this study finds that the syntactic features that develop in L2 learners may not be the same syntactic features that will assist them in receiving higher evaluations of essay quality.

© 2014 Elsevier Inc. All rights reserved.

Keywords: Computational linguistics; L2 writing; Writing development; Writing quality; Syntactic complexity

Introduction

Syntactic development is an important component of second language (L2) acquisition and one that has received considerable attention in previous research (Hawkins, 2001; Lu, 2010) in both longitudinal and cross-sectional studies. Researchers have focused on L2 syntactic development under the notion that the ability to arrange words syntactically into phrases and phrases into clauses demonstrates the capacity to manipulate a language's combinatorial properties, which is argued to be a strong indicator of general language acquisition. One of the primary questions addressed by syntactic research is how syntactic knowledge develops over time and, more specifically, what syntactic features develop early and which develop later for L2 learners (Hawkins, 2001).

* Corresponding author. Tel.: +1 404 413 5179.

E-mail addresses: sacrossley@gmail.com (S.A. Crossley), dsmcnamara1@gmail.com (D.S. McNamara).

¹ Tel.: +1 480 727 5690.

Examinations into the development of syntactic features often focus on the variation and sophistication of the phrases and clauses produced by L2 learners. The basic premise underlying such examinations is that syntactic complexity can be used to directly measure L2 learner proficiency (Foster & Skehan, 1996; Lu, 2011; Ortega, 2003; Wolfe-Quintero, Inagaki, & Kim, 1998).

While a number of studies have examined longitudinal growth in L2 learners using both spoken and written corpora, few studies have examined L2 syntactic development in conjunction with the relationships such developments have with human judgments of writing quality (both judgments of overall writing proficiency and more fine-grained judgments of syntactic proficiency). That is to say, while past research has focused on L2 learner development, it has rarely linked the effects of such development to assessments of language proficiency. However, such an approach is important because it can afford an opportunity to examine not only syntactic growth, but also the relations of such growth with the judgments of expert raters. To address this research gap, this study examines L2 writing samples using computational indices of syntactic complexity to understand how syntactic complexity changes over time in L2 writers (i.e., longitudinal growth) and to understand how changes in syntactic complexity are related to human ratings of language use in L2 writing.

Syntactic complexity

As mentioned earlier, syntactic complexity refers to the sophistication of syntactic forms produced by a speaker or writer and the range or variety of syntactic forms produced (Lu, 2011; Ortega, 2003). Analysis of L2 output in terms of its syntactic complexity is a common means to L2 growth because language development in L2 learners is argued to entail the acquisition and production of less frequent syntactic features along with the use of a greater variety of syntactic features. Many features related to syntactic complexity are relatively easy to investigate using both hand- and automated-coding of texts which allows for the sampling of a variety, but by no means all, of available syntactic features.

The traditional method of measuring syntactic complexity is with T-units (Biber, Gray, & Poonpon, 2011), which can be defined as the shortest allowable grammatical units that can be punctuated at the sentence level (i.e., the main clause plus additional, embedded subordinated clauses; Street, 1971 as cited in Larsen-Freeman, 1978, p. 441). T-units were initially used to assess writing development in first language (L1) writers (Hunt, 1965) and were later adopted for use by the L2 research community (Casanave, 1994; Henry, 1996; Lu, 2011; Ortega, 2003; Stockwell & Harrington, 2003). The use of T-units as measures of syntactic complexity for L2 learners has provided mixed results, with some studies demonstrating no links between classic T-unit measures such as mean length of T-unit and measures of L2 syntactic growth (Bardovi-Harlig, 1992; Casanave, 1994; Ishikawa, 1995) and other studies finding strong links (Ortega, 2003; Stockwell & Harrington, 2003).

The most promising T-unit indices are error-free T-units (Larsen-Freeman, 1978), but such indices are not strictly syntactic and focus more on accuracy than T-units. Additionally, such indices are difficult, if not impossible, to implement computationally and require expert hand coding, which is prone to subjectivity and error. The use of T-units to investigate L2 writing has also been called into question recently by Biber et al. (2011). They found that the clausal subordination measured by T-unit indices is more common in conversation whereas academic writing is characterized syntactically by the use of noun phrase constituents and complex phrases.

Other measures of syntactic complexity that are not specifically based on T-units but are commonly used in L2 writing studies include indices that measure the length of syntactic structures, the types and incidence of embeddings, the types and number of coordinations between clauses, the range and types of phrasal units produced, and the frequency of clauses and phrases used (Ortega, 2003). Such indices can be accessed in computational tools such as the Biber tagger (Biber, 1988) and Coh-Matrix (Graesser, McNamara, Louwerse, & Cai, 2004; McNamara, Graesser, McCarthy, & Cai, 2014).

Syntactic development in L2 learners

Previous research into L2 syntactic acquisition has focused on syntactic development in both spoken and written L2 language samples and has demonstrated that L2 learners follow general patterns of syntactic development that occur in identifiable stages. For instance, English speakers learning French must acquire the rule that direct and indirect object pronouns come before the verb (as compared to after the verb in English). When learning such a rule, L2 learners

Download English Version:

<https://daneshyari.com/en/article/363996>

Download Persian Version:

<https://daneshyari.com/article/363996>

[Daneshyari.com](https://daneshyari.com)