Dimensionality of early writing in English and Spanish

Salas, Naymé; Caravolas, Marketa

Journal of Literacy Research

Published: 01/09/2019

Peer reviewed version

Cyswllt i'r cyhoeddiad / Link to publication

10. Apr. 2024

Dimensionality of Early Writing in English and Spanish

Naymé Salas[a,b] & Markéta Caravolas[a]

Bangor University[a]; Universitat Autònoma de Barcelona[b]

Corresponding author:

Naymé Salas

Dept. Didàctica de la Llengua i la Literatura, i de les Cs. Socials

Facultat d'Educació, Universitat Autònoma de Barcelona

Campus de Bellaterra, Edifici G-5, despatx 131

08193 – Bellaterra, Cerdanyola del Vallès, Barcelona, Spain

Email: nayme.salas@uab.cat

Tel.: +34 935 811678

Abstract

Writing development is understood to be a multidimensional task, heavily constrained by spelling in its early stages. However, most available evidence comes from studies with learners of the inconsistent English orthography, so our understanding of the nature of early writing could be highly biased. We explored writing dimensions in each language by comparing a series of text-based features in the texts of British ($n = 188$) and Spanish ($n = 190$) children between mid-Grade 1 to mid-Grade 2. Results revealed that two constructs, writing conventions and productivity, emerged in both languages, but the influence of orthographic consistency started to be evident in the later time points. Other constructs of text generation seemed to emerge later and were less stable over time. The article thus highlights the language-general underpinnings of early text writing development and the impact of orthographic consistency; furthermore, it strengthens the view that some writing components develop before others. We discuss implications for the assessment of early written products.

Keywords: writing development, cross-linguistic, dimensionality, longitudinal, text-based

Dimensionality of Early Writing in English and Spanish

Writing is a multidimensional skill that involves multiple factors and processes (e.g., Alamargot & Chanquoy, 2001; Kim, Al Otaiba, Folsom, Greulich, & Puranik, 2014; Kim, Al Otaiba, Wanzek, & Gatlin, 2015; Puranik, Lombardino, & Altmann, 2008; Wagner et al., 2011). In particular, the execution of cognitive processes—that is, planning, translating, and revising (Hayes & Flower, 1980)—varies with time and writing experience (Berninger & Swanson, 1994). The translating process involves two subprocesses: text generation, during which ideas become linguistic propositions, and transcription, which comprises both spelling and handwriting (Berninger & Swanson, 1994). During transcription, the linguistic units created during text generation are put into written form (McCutchen, 1996). It is precisely this subprocess, transcription, that makes the greatest contribution to text quality in children's early productions (e.g., Graham, Berninger, Abbott, Abbott, & Whitaker, 1997; Juel, Griffith, & Gough, 1986).

The translation process has received particular attention in developmental research because its two subprocesses are assumed to compete for the available cognitive resources in novice writers (Berninger, 1999). The challenge of spelling, especially, is thought to affect linguistic expression in early written products, because children's attention span may be overloaded by their yet inefficient phoneme-grapheme coding skills (Berninger et al., 1992; Bourdin & Fayol, 1994). Accordingly, cross-linguistic studies of writing development in orthographies of differing levels of complexity provide useful, natural contexts for elucidating the specific effect of spelling constraints on the emergence and expression of the various linguistic dimensions of early written products.

Cross-linguistic studies of literacy are relatively rare, despite their importance for establishing whether developmental trends in reading and writing are similar across languages,

or whether there are language-specific differences that need to be taken into account both for theoretical and educational reasons. A major factor in cross-linguistic examinations of literacy development is the impact of the relative (in)consistency of phoneme-grapheme mappings, because it affects the *rate* of word-level—that is, reading and spelling—development (e.g., Caravolas, 2017; Caravolas & Bruck, 1993; Caravolas, Lervåg, Defior, Seidlová Málková, & Hulme, 2013; Landerl, Wimmer, & Frith, 1997; Seymour, Aro, & Erskine, 2003). Variations in orthographic consistency should have important effects in the text-writing domain, where development heavily depends on the automatization of low-level skills, such as spelling (e.g., Juel et al., 1986). Children who are learning to write need to allocate attention to spelling, potentially at the expense of sufficient cognitive resources for other aspects of writing (e.g., Berninger, 1999; Bourdin & Fayol, 1994; Juel et al., 1986; McCutchen, 2006). Indeed, there is some evidence that spelling poses fewer constraints on learning to write in more consistent orthographies (e.g., Babayiğit & Stainthorp, 2010; Hayes & Berninger, 2009; Lerkkanen, Rasku-Puttonen, Aunola, & Nurmi, 2004; Mäki, Voeten, Vauras, & Poskiparta, 2001), although none of these studies systematically compared a more inconsistent to a more consistent orthography.

Despite the well-attested effect of orthographic consistency on the rate of literacy development, there is evidence that the learning process may not be fundamentally different across languages, at least within alphabetic orthographies. For example, recent research has demonstrated that regardless of orthographic complexity, early literacy development is underpinned by the same subset of cognitive predictors for at least the first two years of formal literacy instruction (Caravolas et al., 2012, 2013). In addition, a systematic, cross-linguistic (Spanish–English) examination of early text productions showed that although Spanish children outperformed English children in spelling and spelling-related skills, such as capitalization and

word separation, there were no differences in other text features, such as vocabulary sophistication, syntactic complexity, or productivity (Salas, 2014). To the best of our knowledge, no published study has yet examined whether the dimensions of early writing vary cross-linguistically. Arguably, if spelling poses different constraints in consistent and in inconsistent orthographies, it may affect the nature and emergence of the various dimensions. For example, if Spanish children attain proficiency in spelling earlier than their English counterparts (for whom this skill continues to create a bottleneck in writing development), this may allow higher-level dimensions to emerge earlier in Spanish than in English. To investigate these issues, in this article we report a cross-linguistic, longitudinal comparison of the early (first- and second-grade) stages of writing development in English, a well-studied language with a complex, inconsistent orthography, and in Spanish, a less studied language with a much simpler orthography (Borgwaldt, Hellwig, & de Groot, 2004). Our main goal was to determine whether the emergence of writing dimensions in early text writing is characterized by strong language-specific trends or whether the same, language-general factors operate across languages.

<div align="center">**Background**</div>

Assessing Children's Written Productions

The assessment of writing has recently extended to a multidimensional perspective, where multiple, easy-to-measure, objective indicators of writing features and skills are obtained. Usually, these indicators are examined using factor-analytic techniques to establish the way in which they may be reflecting underlying dimensions of writing (e.g., Puranik et al., 2008; Wagner et al., 2011). For example, Y. S. Kim et al. (2011) assessed the writing products of kindergarten-level children and found evidence of the existence of a productivity dimension, whose indicators were the total number of words, the total number of sentences, and the total

number of ideas. Similarly, Wagner et al. (2011) identified a handwriting fluency dimension, indexed by children's (Grades 1 and 4 of elementary education) results on a timed alphabet writing task and a sentence copying task.

A multidimensional approach to the assessment of early writing is useful for at least two reasons. First, the field of early writing should benefit from a fine-grained, information-rich approach, at least until more research accrues. Such investigations should then inform the creation of standardized tests for assessing writing development with norms for children under the age of 8 years (or Grade 3), as these are currently not available (e.g., WIAT-III: Weschler, 2009; PROESC: Cuetos, Ramos, & Ruano, 2004). Second, multidimensional, multi-trait approaches allow for the examination of the underlying cognitive constructs or domains (e.g., Y. S. Kim et al., 2014; Puranik et al., 2008; Wagner et al., 2011). This approach is particularly relevant to the context of the present article, as it allows us to investigate whether the dimensions underlying a variety of text-based features are the same, and whether they emerge at the same points in development in two languages with contrasting orthographies.

Previous Studies on the Dimensionality of Writing

Only a handful of studies have assessed the dimensions of early writing in alphabetic orthographies. Most have been carried out with English-speaking participants and had a cross-sectional design (e.g., Puranik et al., 2008; Wagner et al., 2011). It should be noted, however, that while some of them focused solely on text-based features of writing (e.g., spelling accuracy, punctuation, and grammatical complexity, among others, as in Puranik et al., 2008), others focused on the component skills of writing, using measures that were not text-based, but obtained from independent tasks (e.g., spelling and handwriting, among others, as in Y. S. Kim et al., 2011). Still other studies have mixed the two types of measures (e.g., handwriting fluency,

measured by a standardized test, and a series of text-based measures, as in Wagner et al., 2011). In this article, we confined our analysis to text-based features of early (from mid-Grade 1) text products across two languages: Spanish and English.

The small body of literature on writing dimensions in English has converged on a number of factors. There is robust empirical evidence of a *productivity* dimension that, by measuring the number of words or ideas generated, captures children's skill to produce written output (e.g., Kim et al., 2011; Puranik et al., 2008; Wagner et al., 2011). Spelling or *transcription* (i.e., spelling + handwriting) constitutes another domain that plays a role, especially in the early stages of learning to write (presumably due to its aforementioned constraining effect on early writing development; Kim et al., 2014). In addition, support for the existence of a *quality* dimension has been reported, which usually taps the ability to organize and develop ideas coherently; notably, in these studies, quality has been measured either holistically or using qualitative scales. Thus, it is not entirely clear what specific text features might contribute to a quality dimension (Kim et al., 2014; Kim, et al., 2015; Wagner et al., 2011).

Less is known about the linguistic makeup of early text productions. While oral language skills, such as grammatical and vocabulary knowledge, are reported to affect the quality of children's early written products (Kim, Al Otaiba, Sidler, & Gruelich, 2013; Kim et al., 2014; Puranik et al., 2008), very rarely has the early development of specific text features been addressed. However, there is evidence that syntactic accuracy and complexity may emerge as early as Grade 1 (Wagner et al., 2011) and that aspects of vocabulary have been measured in Grade 2 children (Olinghouse & Leaird, 2009).

In sum, to date, studies on the dimensions of writing have been carried out primarily with English-speaking participants. The vast majority were cross-sectional, and the indicators of the

investigated factors took into account a mixture of text-based features (e.g., number of words, spelling accuracy, punctuation) and component skills measured independently of writing. This article thus expands on extant research by conducting a longitudinal (first- through second-grade) examination of the dimensionality of writing, while contrasting English-speaking to Spanish-speaking children's early writing development. Additionally, in each language, we exclusively analyzed measures of text-based features.

The Present Study

We investigated the development of several text-based features in the early writing products of English- and Spanish-speaking children on three occasions from Grade 1 to mid-Grade 2. In using repeated administrations of a simple writing task, we expected this developmental window to provide insights into the phase during which spelling ability should exert the greatest impact on emergent writing skills (Berninger & Swanson, 1994; Juel, 1988). Our main aim was thus to establish the similarities and differences in the role of spelling and in the emerging dimensionality of writing in the early stages of writing development of English- and Spanish-speaking children.

Consistent with the multidimensional approach, a series of features, assumed to index various dimensions of writing, was identified and extracted from the text of each child at each time point and was scored or quantified for analysis. Our choice of features was motivated by previous studies (e.g., Puranik et al., 2008; Wagner et al., 2011), by the nature of the writing task (see the "Method" section), and by the developmental window under study. Features included multiple indicators of text spelling and the spelling-related skills of word separation and capitalization, as well as a number of linguistic indicators of the hypothesized domains of vocabulary, syntactic complexity, and connectivity. In addition, we measured text productivity,

operationalized as the number of words per text, because it has been found to be a reliable proxy for the number of ideas and overall quality, and a robust measure of writing development (e.g., Berman & Nir, 2009; McMaster & Espin, 2007; Scott & Windsor, 2000).

Method

Participants

Altogether, 188 English-speaking children from the north of England and 190 Spanish-speaking children from Granada, Spain, who were drawn from a larger study on word-level literacy development across multiple languages, participated in this study. The English sample consisted of 98 boys and 90 girls, whose mean age at Time 1 was 60.27 months ($SD = 3.67$). The Spanish sample consisted of 104 boys and 86 girls, whose mean age at Time 1 was 66.72 months ($SD = 3.57$). The data reported exclusively in the present study, and derived from the first author's doctoral research, were collected when children were at mid-Grade 1 (Time 1, or T1), end of Grade 1 (Time 2, or T2), and mid-Grade 2 (Time 3, or T3). Missing data in the English sample were 7.98%, 9.04%, and 11.70% at T1, T2, and T3, respectively; in the Spanish sample, they were 12.10%, 15.26%, and 10.53% at T1, T2, and T3, respectively. The main causes of missing data were participant absences on the date of data collection or children having left the school. Children in the Spanish group were, on average, six months older than their English counterparts. This is because in England children enroll in kindergarten (reception year) classes the September after their 4th birthday, but in Spain they start the equivalent class the year they turn 5. Not surprisingly, a significant age difference was found, $t(376) = 17.11$, $p < .001$, $d = 1.76$. Controlling for age did not alter the pattern of results reported below; in addition, age did not correlate with virtually any of the target measures. For these reasons, age was not controlled for in subsequent analyses.

Teaching practices. Children in the English sample attended schools that followed the National Literacy Strategy, which adopted a phonics-based approach to literacy instruction. Children in the Spanish sample attended schools that followed the regional (i.e., Andalusian) framework for literacy instruction. Assessment of teacher practices was carried out through a series of questionnaires administered yearly to the teachers of all participating classrooms. We performed an analysis of the amount of time per week that the Grade 1 teachers devoted to a series of literacy activities (handwriting, spelling, and freewriting) in each language. Mann-Whitney U tests indicated that in both countries, children received a similar amount of instruction on handwriting ($U = 43.5$, $p = .449$), spelling ($U = 27.5$, $p = .247$), and freewriting ($U = 20.5$, $p = .074$).[1] For this reason, teaching practices were not controlled for in subsequent analyses.

Task and Procedure

Writing recent past events. Data were collected around mid-Year 1 (November–December, Time 1), end of Year 1 (May–June, Time 2), and again around mid-Year 2 (November–December, Time 3) at the children's schools. Children in both language groups were administered a text-writing task that required writing about recent past events. The administrator asked groups of five children about what they had done the day before after they had left school. Children wrote for five minutes and were encouraged to write as much as they could. They were praised for any efforts. No advice or more specific prompts were given. Once the five minutes elapsed, children were instructed to draw a picture illustrating their text. The administrator used that time to ask each child to read back the text she had produced. This procedure helped to elucidate each child's intended message, especially for children with illegible handwriting or

---

[1] The full descriptive tables are not reported here due to space limitations, but are available from the first author upon request.

very poor spelling skills (similar procedures were reported by Juel et al., 1986 and by Y. S. Kim et al., 2015). A number of text-based indicators was evaluated; they are described below, grouped according to the dimension they were hypothesized to index.

Writing conventions. This dimension was hypothesized to include aspects of children's accuracy in spelling and spelling-related skills.

Spelling accuracy. The percentage of words that were spelled correctly was used as an index of the child's text spelling skill. Numbers, symbols (e.g., *8, &, +*), and loan words (e.g., *Nintendo*, for the Spanish sample) were excluded from this count. Incorrect word separation was not penalized, so a child writing *Iwent* was given credit for two accurately spelled words, *I* and *went*. The scoring was binary: 0, 1.

Word separation. The percentage of words that showed conventional boundaries was used as an index of children's word separation skills, which are essential for applying some orthographic rules. When two words were written as one word (e.g., *Iwent*), both tokens were penalized, regardless of whether the words were spelled correctly. The scoring was binary: 0, 1.

Capitalization. Children's knowledge and use of lower- and uppercase letters was assessed calculating the percentage of words that showed conventional use of case. Words in which the uppercase version of a letter was used in an inappropriate context, for example, within a word as in *hoMe*, were counted as errors. The scoring was binary: 0, 1.

Productivity. This dimension was hypothesized to reflect children's text-generation skills at the word and letter levels. The two different transcription units (words, letters) were considered to ensure that language-specific traits, such as average word length, did not bias the results (Cutler, 2012; Cutler, Norris, & Sebastián-Gallés, 2004).

Number of words. Children's text-production skills were assessed by counting the total number of words written in the allotted time of five minutes. Incorrect spelling or unconventional word separation did not affect this count. For example, a child who wrote *Iplaid* for *I played* was given credit for writing two words.

Number of letters. Children's text-production skills were also assessed by counting the number of letters they wrote. Words for which no parallel could be established between the intended meaning and the actual written string were excluded from all counts. Letters were counted automatically in Excel.

Connectivity. This dimension was hypothesized to estimate children's skills to connect parts of the text they were producing, which is a key aspect of text organization and its resulting coherence.

Connectors. Conjunctions (e.g., *and*, *or*) and discourse markers (e.g., *then*, *first*) are fundamental to showing the links among parts of the text and thus contributing to text cohesion and coherence (Hickmann, 2003). Children's use of connectors was assessed by identifying all inter-clausal connectors in the texts and then calculating the percentage these represented of the total number of words in the text.

Punctuation. Punctuation marks signal the boundaries of different discourse segments, and there is evidence that children acquire them in a predictable order (Ferreiro & Pontecorvo, 1999). All types of punctuation marks—periods, exclamation points, commas, question marks, and so on—were counted manually, directly from the original texts. In order to control for the effect of overall text length, whereby longer texts might contain more punctuation marks, the total number of punctuation marks was divided by the number of words in the text and multiplied by 100.

Syntactic complexity. This dimension assessed children's grammatical skills, particularly the degree of density and embeddedness in syntactic constructions, which is a fundamental aspect of discourse elaboration and essential to the characterization of text composition skills across the life-span (Berman, 2008).

Words per clause. This measure of syntagmatic syntactic complexity was obtained by dividing the total number of words by the number of clauses produced by the child.

Subordination. This measure of the degree of syntactic nesting or embeddedness consisted of calculating the percentage of subordinate clauses out of the total number of clauses, thus constituting an index of the amount of embeddedness in children's texts.

Vocabulary. A set of measures was derived to estimate children's lexical choices and evaluate the lexical precision and elaboration in their written products.

Adjectives and adverbs. Optional lexical elements such as adjectives and adverbs add to the elaboration and precision of the message (Ravid & Levie, 2010), and were therefore tallied in all texts. Adverbs of obligatory expression, such as those in phrasal verbs (e.g., *take off*), were not counted. The score consisted of calculating the percentage of adjectives and adverbs of all words in the text.

Lexical density. The percentage of content words or open-class lexical tokens (e.g., nouns, verbs, adjectives, adverbs), as opposed to the proportion of closed-class tokens (e.g., prepositions, determiners, pronouns), was used as an indicator of lexical density, based on the assumption that a greater proportion of "semantically charged" tokens results in a richer and more informative text (Halliday & Hassan, 1985; Malvern, Richards, Chipere, & Durán, 2004). All words in the texts were labeled for grammatical class. Then, all content words were counted

and the resulting number was divided by the total number of words in the texts and multiplied by 100.

Lexical diversity. The type-token ratio[2] was calculated for each text as an indicator of lexical diversity in written composition (Malvern et al., 2004). Any modification to a word base, such as adding or subtracting inflectional endings or derivational morphemes, was interpreted as a different type.

Average content-word length. The average number of letters of all content words in children's texts was measured as a proxy for lexical sophistication and use of low-frequency tokens, based on the assumption that longer words are less frequent (Bybee, 2007).

Reliability

All texts were transcribed and coded (tagged) according to the units of interest described above and evaluated (scored) by the first author. Reliability between the first author and two external coders (native speakers of Spanish and English, respectively), blind to the objective of the study, was calculated on 20% of the sample. The coders were first trained on the coding criteria with an initial sample of 10 texts, after which they each coded 10% of the texts in each language group. Texts were selected in a quasi-random fashion to ensure an equivalent number of texts from each time point. Agreement was $r > .81$ or higher for all measures—spelling accuracy (.96 for English, .99 for Spanish), word separation, (.97 for English, .99 for Spanish), capitalization (.95 for English, .97 for Spanish), number of words (.94 for English, .99 for Spanish), punctuation (.81 for English, .94 for Spanish), and number and type of clauses (.90 for English, .96 for Spanish).

---

[2] The type-token ratio is a frequently used measure of lexical diversity. Each unique word is a type and each word is a token. Therefore, the ratio of types to tokens is a figure that oscillates between 0 and 1, where higher numbers indicate more lexical diversity; on the contrary, lower numbers reflect that the same tokens (words) have been used repeatedly.

For measures based on the tagging of texts for part of speech—that is, number of adjectives and adverbs, number of connectors, lexical density, and average length of content words—coding reliability was obtained by calculating the degree of agreement between the first author's manual tagging of each word and an automatic part-of-speech analyzer (CLAWS4 for English: Garside, 1987; MACO for Spanish: Civit, 2003) on a quasi-random selection of 10% of the texts in each language (20% of the total sample). Agreement between the two types of counts always exceeded 80%.

Preliminary Analyses

Because the study dealt with very early text productions, it was necessary to check for the existence of floor effects, especially in the earlier time points, as well as ceiling effects, as children made progress on various measures. An assessment of the psychometric properties of the measures was carried out in two steps: First, outlier scores were identified and corrected; second, the normality of the distribution for each variable was evaluated. Following Kline (2011, pp. 54–55), the outliers assessment consisted of looking for cases that were more than 3.0 standard deviations (*SD*s) above or below the mean for all relevant variables. Outliers amounted to only 1.06% of all scores (133 out of 13,107 data points). The average number of extreme scores per variable was 3.69 (*SD* = 2.27; range: 0–9), that is, less than 1%. Twenty-five children showed two or more extreme scores. We decided to keep these participants as we aimed to recruit an unselected sample, representative of the population found at schools. In order to avoid extreme scores distorting the distribution of values within a given variable and violating assumptions of ensuing statistical tests, scores exceeding 3 *SD* were Winsorized (following Osborne, 2013) to the value equivalent to exactly ±2.99 *SD*s. All the analyses reported below were carried out with the corrected scores, and these replicated analyses with the original scores.

The text writing measures were next checked for the normality of the distribution for each variable. None of the selected measures showed skewness or kurtosis values outside the recommended cutoff points (Kline, 2011). Skewness values ranged from –1.72 to 1.86, while kurtosis values ranged from –1.58 to 4.63. In sum, no major deviations from normality were observed for the final set of measures.

Results

The descriptive statistics for each measure are reported in Table 1, and these raised no concerns regarding floor and ceiling effects. Prior to exploring the dimensional structure, we examined the pattern of relationships between the multiple indicators using Pearson product-moment correlations for each language and time point (see Supplemental Tables 1, 2, and 3 for correlations at T1, T2, and T3, respectively).

[Insert Table 1 about here]

Writing conventions measures (spelling accuracy, word separation, and capitalization) were significantly associated throughout the study period in both languages. In English, the correlations between these three variables were consistently high, ranging from $r = .56$ to $r = .77$. In Spanish, they ranged from $r = .32$ to $r = .61$, with the exception of the correlation between capitalization and word separation at Time 1 (S-Table 1). Therefore, although correlations differed slightly in size, both language groups showed these three indicators to be tightly related in the early stages of learning to write.

Productivity measures, that is, number of words and number of letters, showed very high correlations throughout the study period ($r > .90$). The close relationship between different measures of productivity was to be expected and has been found before (e.g., Berninger et al., 1992; Puranik et al., 2008; Wagner et al., 2011). It has been suggested that number of words

constitutes the most representative of these measures and thus is the most widely used in studies

measuring productivity (Puranik et al., 2008). It is also evident from S-Tables 1, 2, and 3 that

number of words and number of letters were not only extremely highly correlated, but their

pattern of correlations with other measures was virtually identical.

The dimensions of *syntactic complexity*, *vocabulary*, and *connectivity* did not show such

clear patterns of relationships. With a few exceptions, correlations between the indicators of each

dimension were in the low range and usually nonsignificant. Syntactic complexity measures

were, as a rule, not significantly correlated with each other in either language, throughout the

study. The measures of the vocabulary dimension were mostly unrelated at Time 1 (S-Table 1),

with the exception of content-word length, which was negatively associated with the proportion

of adjectives and adverbs in English; in Spanish, lexical diversity was positively associated with

lexical density. This pattern was also observed at Time 3, but not at Time 2. At Times 2 and 3,

however, both language groups showed moderate correlations between the proportion of

adjectives and adverbs and lexical density (S-Tables 1 and 2). It is not surprising that these two

measures were related, given that the adjectives and adverbs measure is a more refined count of

lexical density (which includes the percentage of adjectives and adverbs, as well as of nouns and

lexical verbs). Of interest was the consistently strong negative correlation between lexical

diversity and productivity measures—number of words and number of letters. This may suggest

that lexical diversity is more sensitive to text length rather than tapping aspects of written

vocabulary richness (see Malvern et al., 2004).

Exploratory Factor Analyses

The hypothesized dimensions of writing conventions, productivity, syntactic complexity,

vocabulary, and connectivity were examined by conducting exploratory factor analyses with a

principal components factor extraction method. Analyses were run for each time point and language group separately in order to examine the longitudinal stability of the constructs. The Direct-Oblimin rotation was applied because it is the recommended approach when factors are likely to be correlated (Tabachnick & Fidell, 2007), as has been the case in previous studies of the dimensionality of writing (e.g., Y. S. Kim et al., 2015; Puranik et al., 2008). A Principal Components Analysis (PCA) approach was chosen also due to the absence of previous studies of a similar scope—that is, that dealt with very early text productions across two language groups. All text-based measures were entered and measures of sampling adequacy (MSA) were assessed for the model and for each individual variable. Variables with MSAs < .6 were dropped (Kaiser, 1974, as cited in J. O. Kim & Mueller, 1978), as well as those with communalities < .50 (Field, 2009), and the analysis was rerun with the remaining variables. In all cases, factors extracted were those with eigenvalues > 1.

Dimensionality of writing in English. Table 2 shows the standardized factor loadings for English at Times 1, 2, and 3. At Time 1, a four-factor solution was found, which explained 76% of the variance. At Times 2 and 3, in contrast, a three-factor solution was found, which explained 80.08% and 79.27% of total the variance, respectively. At all time points, spelling accuracy, word separation, and capitalization loaded strongly on a common, *writing conventions* factor, and this factor accounted for a consistently high proportion of variance (approximately 40%–49%). Also across time points, number of words and number of letters loaded on a common *productivity* factor, together with lexical diversity. It should be noted that although lexical diversity was initially considered an index of vocabulary-related aspects, it was always strongly—and negatively—associated with productivity measures, suggesting that longer texts were characterized by less diverse vocabulary. We return to this finding in the "Discussion"

section. The productivity construct itself consistently accounted for a significant proportion of variance, albeit lower than the contribution of writing conventions. At Times 2 and 3, two measures of vocabulary, percentage of adjectives and adverbs and lexical density, loaded on a common *vocabulary* factor. As predicted, punctuation and connectors loaded on a common *connectivity* factor, although this factor emerged only at Time 1; in addition, their loadings were of opposite signs, meaning that children chose either punctuation marks or inter-clausal connectors to link the different parts of the text. Percentage of subordinate clauses and percentage of adjectives and adverbs loaded on a common, unexpected factor at Time 1 only, perhaps because they were both tapping the relative degree of informativeness and precision that these measures contribute to. Finally, there was no evidence of a syntactic complexity factor at any time point.

[Insert Tables 2 and 3 about here]

In summary, two factors were consistently obtained across time points in English: a writing conventions factor, composed of spelling accuracy, word separation, and capitalization, and a productivity factor, composed of number of letters, number of words, and lexical diversity (or type-token ratio). In addition, a vocabulary factor emerged at Times 2 and 3, composed of lexical density and percentage of adjectives and adverbs. The Time 1 factor of connectivity did not endure to the later time points. Moreover, there were small to moderate, negative correlations between the writing conventions and productivity factors of $r = -.286$, $r = -.353$, and $r = -.210$ at Times 1, 2, and 3, respectively. This suggests that even though writing conventions measures controlled for text length (i.e., they were calculated as a proportion of all words in the text), the more children wrote, the less accurate they were across all writing conventions measures. The writing conventions factor was also moderately and positively correlated with the vocabulary

factor ($r = .372$) at Times 2 and 3, while the correlation between the productivity factor and the vocabulary factor was $r = -.154$ at Time 2 and $r = -.137$ at Time 3, suggesting that as children wrote more, their vocabulary tended to be less sophisticated.

Dimensionality of writing in Spanish. Table 3 shows the standardized factor loadings for Spanish at Times 1, 2, and 3. At Times 1 and 2, a two-factor solution was found, which explained 81.74% and 76.69% of the variance, respectively. Spelling accuracy and word separation, but not capitalization, loaded heavily on the first, *writing conventions*, factor. Only at Time 3 did capitalization additionally load on this factor. As in English, the writing conventions factor accounted for the largest proportion of variance at each time point, but unlike in English, its contribution systematically decreased over time. Similarly to English, number of words, letters, and lexical diversity loaded on another factor, *productivity*, across time points, and the strength of this factor appeared stable over time. At Time 3, a three-factor solution was found, accounting for 77.85% of the variance. The third factor was akin to the vocabulary factor found for English, which was also composed of percentage of adjectives and adverbs and lexical density. Syntactic complexity measures and connectivity measures did not load on any factor. As in English, the writing conventions factor was moderately related to the productivity factor at T1 and T2, but not at T3 (correlations were $r = .297$, .220, and .043 at Times 1, 2, and 3, respectively). Note that for Spanish, the correlations between these two factors were positive; thus, the more children wrote, the more accurate they were in spelling, word separation, and capitalization. Finally, the vocabulary factor showed a correlation of $r = -.108$ with the writing conventions factor and of $r = .052$ with the productivity factor. In sum, also in Spanish the writing conventions and the productivity dimensions were salient constructs throughout the study period. A vocabulary factor was also identified, composed of adjectives and adverbs and lexical

density, but only in the last time point. In contrast to English, there emerged fewer factors, while the relationship of writing conventions to other dimensions decreased with time. In both languages, the examination of the correlations (S-Tables 1, 2, and, 3) across time points revealed some tendencies for other dimensions, but these did not surpass the loading thresholds of the exploratory factor analyses. This suggests that some aspects of text composition may emerge later in development, and during their emergence they lack stability.

Discussion

This article set out to establish the nature and complexity of early text-writing development among children learning to write in English and those learning to write in Spanish. For this purpose, children's text productions were systematically coded and evaluated on a large number of text-based measures at three time points in Grades 1 and 2. In previous studies, we found that Spanish children showed an advantage in single-word spelling accuracy and other spelling-related skills, such as word separation and capitalization, but that beyond these word-level features, children learning to write in both languages showed very similar trends (Caravolas et al., 2012; Salas, 2014). In the current study, the English and Spanish data from each time point were subjected to exploratory factor analyses in order to reveal the factor structure of text-generation skills, via an assessment of the various text-based features making up these early text productions. Additionally, we aimed to determine the stability of these emerging constructs over the first year-and-a-half of formal literacy instruction.

Results showed that to a large extent, the same dimensions operated in both languages. Two dimensions stood out as well established constructs already by mid-Grade 1 across the two languages: writing conventions and productivity. Other studies have also found support for a productivity construct and a writing conventions or "accuracy" construct (Puranik et al., 2008;

Wagner et al., 2011). The identification of these robust, early emerging constructs is thus in line with previous studies and extends our current understanding of the dimensionality of writing to more consistent orthographies (e.g., Y. S. Kim et al., 2014; Puranik et al., 2008; Puranik & Al Otaiba, 2012; Wagner et al., 2011).

Over and above the two foundational constructs of writing conventions and productivity, other text-based indicators tended to group together as we had hypothesized, but they did so at later time points. A vocabulary factor emerged in English and Spanish by the end of Grade 1 and by the middle of Grade 2, respectively. Percentage of adjectives and adverbs and lexical density were the variables loading on this factor across languages. This is not surprising, because the percentage of adjectives and adverbs is a subset of the percentage of content words, measured by the lexical density measure. However, the fact that they were not consistently related in English until the end of Grade 1, and in Spanish until mid-Grade 2, suggests that some constructs of text-level performance may not consolidate until later stages of writing development. Taken together, the finding that writing conventions and productivity factors emerge and stabilize fairly early on, while other, more rhetorical aspects of text construction emerge later and are less stable constructs at these early stages can be considered to reflect children's knowledge-telling processes of text composition (Bereiter & Scardamalia, 1987).

Of interest is the finding that lexical diversity, calculated as the type-token ratio, was not associated with any vocabulary measure, but loaded heavily on the productivity factor across time points and language groups. Other studies observed a similar behavior of the corrected type-token ratio measure (Olinghouse & Leaird, 2009). This result further substantiates previous observations about the nature of lexical diversity as a measure extremely sensitive to text length, but that poorly captures vocabulary-related aspects of text construction (Malvern et al., 2004).

In contrast to previous studies assessing the dimensionality of early writing, we did not find support for a "complexity" factor. Syntactic complexity measures were characterized by their instability, both in terms of their lack of associations with one another throughout the study period, as well as in their pattern of relationships with other measures, which was erratic. In addition, they were often excluded from factor analyses due to their lack of a contribution to the models. Other studies measuring written syntax have reported inconsistent patterns in relation to measures such as these. For example, Beers and Nagy (2010) reported negative correlations between their measures of words per clause and of clauses per t-unit. The differences between the present results and those of Puranik et al. (2008), who obtained support for a complexity factor, may be related to the fact that participants in their study were older (starting at Grade 3 and up to Grade 6), and therefore their written syntactic complexity skills could have been more consolidated. Certainly, syntactic complexity in written composition has a protracted development (Berman & Nir-Sagiv, 2007; Berman & Verhoeven, 2002; Ravid & Berman, 2010). The differences between the present results and those reported by Wagner et al. (2011), who found that a measure of t-unit length (in words) and a measure of clauses per t-unit were individual indicators of a latent complexity factor in first and fourth graders, may be related to (a) the slightly different types of syntactic measures; (b) the use of a different prompt (explaining which animal they would like as a pet for their class) and hence a different genre, which may have affected syntactic complexity (Beers & Nagy, 2009, 2010); and (c) children having twice as much time to produce the texts. Finally, there was little support for a connectivity dimension. All in all, the lack of stability of the vocabulary and syntax factors could also be partially related to the fact that children were evaluated on a type of writing task that, while chosen to facilitate children's opportunities to generate content and ideas, did not relate to a concrete or well-known

discourse genre that children could have drawn on. Future research should strive to obtain multiple text samples from each child and at each time point in order to better estimate their text-writing skills.

While the commonalities across languages far exceeded the discrepancies, some language-specific trends were identified. First, the writing conventions dimension was fairly stable in both languages, but it explained progressively less variance in Spanish. This contrast might reflect the more prominent and protracted role of spelling and spelling-related skills in English, given that it is a much more inconsistent orthography than Spanish (Caravolas et al., 2012; Salas, 2014). As Spanish children are closer to automatizing these skills, the construct itself may increasingly fail to capture individual differences. Second, but related to the above, the writing conventions dimension showed stronger relationships with other dimensions, such as vocabulary, in English than in Spanish. This could also be reflecting the more constraining role of spelling in English, as compared to the simpler Spanish orthography. In sum, writing conventions seems to be a common dimension operating across languages, although its weight will vary as a function of children's progress on these skills (Juel et al., 1986), and therefore it will be subject to cross-linguistic differences. Productivity, on the other hand, appeared to have the potential for a robust, cross-linguistic dimension of writing development less affected by orthographic consistency (Salas, 2014), as it was fairly stable in both language groups and showed very similar patterns of relationships with other domains. Previous cross-linguistic studies with older children and adults had already pointed out that text length is an excellent developmental marker, both for writing and oral text production (Berman & Nir, 2009; Berman & Verhoeven, 2002). As such, it may be a privileged indicator of text quality and of language development across languages and across the life-span. This hypothesis, however, would require

the validation of these exploratory findings through the use of confirmatory factor analysis techniques. A final note pertaining cross-linguistic differences is the inherent difficulty of conducting this type of research design, which involves participants across countries. Our decision that participants were matched in terms of the amount of formal literacy instruction was at the expense of a significant difference in chronological age. The confound between the variable of interest, language/orthographic consistency, and age is hard to fully disentangle with natural samples being recruited in different countries, although in our study age seemed to have little effect on the patterns of results obtained. Future research should endeavor to compare writing development across populations that start literacy instruction at the same chronological age—for example, comparing Spanish children to English-speaking children from education systems where literacy instruction begins later.

Overall, our results suggest that across languages, some constructs of text-generation performance may only be emerging during the first and second years in school, hence accounting for their instability over time. These findings are consistent with the view that some writing components develop before others (e.g., Berninger & Swanson, 1994; McCutchen, 2006) and should be taken into consideration for early writing assessment. This means that even very early written composition can be evaluated on at least these two dimensions (writing conventions and productivity). Arguably, testing at these foundational stages of writing development should enable early detection of writing difficulties and could possibly improve the effectiveness of the remediation strategies that are put into action. The striking cross-linguistic similarity in terms of writing dimensions contributes to a growing body of evidence proposing that despite differences in orthographic complexity, teaching practices, and cultural contrasts, there exists a foundational literacy development stage that is characterized more by commonalities than by differences, at

least in alphabetic orthographies. Moreover, where between-language differences are observed, these may usually be ascribed to the moderating effects of phono-graphemic consistency. This line of research has found robust evidence showing that although the rate of reading and spelling development is sensitive to variations in orthographic consistency, these skills draw on the same cognitive precursors (e.g., Caravolas et al., 2012, 2013). This study builds on that notion by proposing that writing dimensions may also be language-general and that, with the exception of writing conventions that depend on phoneme-grapheme mapping consistency, they develop at a similar pace.

References

Alamargot, D., & Chanquoy, L. (2001). *Through the models of writing: With commentaries by Ronald T. Kellogg & John R. Hayes* (Vol. 9). London, UK: Kluwer Academic.

Babayiğit, S., & Stainthorp, R. (2010). Component processes of early reading, spelling, and narrative writing skills in Turkish: A longitudinal study. *Reading and Writing*, *23*(5), 539–568.

Beers, S. F., & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Reading and Writing*, *22*(2), 185–200.

Beers, S. F., & Nagy, W. E. (2010). Writing development in four genres from Grades three to seven: Syntactic complexity and genre differentiation. *Reading and Writing*, *24*(2), 183–202.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum.

Berman, R. (2008). The psycholinguistics of developing text construction. *Journal of Child Language*, *35*(4), 735–771.

Berman, R., & Nir, B. (2009). Cognitive and linguistic factors in evaluating text quality: Global versus local? In V. Evans & S. Pourcel (Eds.), *New directions in cognitive linguistics* (pp. 421–440). Amsterdam, Netherlands: John Benjamins.

Berman, R. A., & Nir-Sagiv, B. (2007). Comparing narrative and expository text construction across adolescence: A developmental paradox. *Discourse Processes*, *43*(2), 79–120.

Berman, R. A., & Verhoeven, L. (2002). Cross-linguistic perspectives on the development of text-production abilities: Speech and writing. *Written Language & Literacy*, *5*(1), 1–43.

Berninger, V. W. (1999). Coordinating transcription and text generation in working memory during composing: Automatic and constructive processes. *Learning Disability Quarterly*, *22*(2), 99–112.

Berninger, V. W., & Swanson, H. L. (1994). Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writing. *Advances in Cognition and Educational Practice*, *2*, 57–81.

Berninger, V., Yates, C., Cartwright, A., Rutberg, J., Remy, E., & Abbott, R. (1992). Lower-level developmental skills in beginning writing. *Reading and Writing*, *4*(3), 257–280.

Borgwaldt, S. R., Hellwig, F. M., & de Groot, A. (2004). Word-initial entropy in five languages: Letter to sound, and sound to letter. *Written Language & Literacy*, *7*(2), 165–184.

Bourdin, B., & Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. *International Journal of Psychology*, *29*(5), 591–620.

Bybee, J. (2007). Review of Goldberg's constructions at work: The nature of generalization in language. *Journal of Child Language*, 3, 14–16.

Caravolas, M. (2017). Learning to read Czech and Slovak. In L. Verhoeven & C. Perfetti (Eds.), *Learning to read across languages and writing systems* (pp. 371–392). Cambridge, UK: Cambridge University Press.

Caravolas, M., & Bruck, M. (1993). The effect of oral and written language input on children′s phonological awareness: A cross-linguistic study. *Journal of Experimental Child Psychology*, 55(1), 1–30.

Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science*, *24*(8), 1398–1407.

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., ... Seidlová-Málková, G. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science*, *23*(6), 678–686.

Civit, M. (2003). *Guía para la anotación sintáctica de Cast3LB: un corpus del espanol con anotación sintáctica, semántica y pragmática. X-TRACT-II WP-03-06 y 3LB-WP-02-01*. Barcelona, Spain: CliC–UB.

Cuetos, F., Ramos, J. L., & Ruano, E. (2004). *PROESC. Batería de Evaluación de los Procesos de la Escritura*. Barcelona, Spain: TEA.

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Cambridge, MA: MIT Press.

Cutler, A., Norris, D., & Sebastián-Gallés, N. (2004). Phonemic repertoire and similarity within the vocabulary. In *Proceedings of the 8th International Conference on Spoken Language Processing* (Vol. 1, pp. 65–68). Seoul, Korea: Sunjin Printing.

Ferreiro, E., & Pontecorvo, C. (1999). Managing the written text: The beginning of punctuation in children's writing. *Learning and Instruction*, *9(*6), 543–564.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London, UK: Sage.

Garside, R. (1987). The CLAWS word-tagging system. In Garside R., F. Leech, and G. Sampson, (Eds.). *The Computational Analysis of English: A Corpus-Based Approach*. Harlow:Longman.

Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology*, *89*(1), 170.

Halliday, M. A. K., & Hassan, R. (1985). *Language: Context and text: Aspects of language in a social-semiotic perspective.* Oxford: Oxford University Press.

Hayes, J. R., & Berninger, V. W. (2009). Relationships between idea generation and transcription. In C. Bazerman, R. Krut, K. Lunsford, S. McLeod, S. Null, P. Rogers, & A. Stansell (Eds.), *Traditions of writing research* (pp. 166–180)*.* London, UK: Routledge.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30)*.* Hillsdale, NJ: Lawrence Erlbaum.

Hickmann, M. (2003). Coherence, cohesion, and context in narrative development: Some comparative perspectives. In S. Strömqvist & L. Verhoeven (Eds.), *Relating events in narrative: Typological and contextual perspectives* (pp. 81–306). Mahwah, NJ: Lawrence Erlbaum.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, *80*(4), 437.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal

    study of children in first and second grade. *Journal of Educational Psychology*, 78(4), 243.

Kim, J. O., & Mueller, C. W. (Eds.). (1978). *Introduction to factor analysis: What it is and how

    to do it*. London, UK: Sage.

Kim, Y. S., Al Otaiba, S., Folsom, J. S., Greulich, L., & Puranik, C. (2014). Evaluating the

    dimensionality of first-grade written composition. *Journal of Speech, Language, and

    Hearing Research*, *57*(1), 199–211.

Kim, Y. S., Al Otaiba, S., Puranik, C., Folsom, J. S., Greulich, L., & Wagner, R. K. (2011).

    Componential skills of beginning writing: An exploratory study. *Learning and Individual

    Differences*, *21*(5), 517–525.

Kim, Y. S., Al Otaiba, S., Sidler, J. F., & Gruelich, L. (2013). Language, literacy, attentional

    behaviors, and instructional quality predictors of written composition for first graders. *Early

    Childhood Research Quarterly*, *28*(3), 461–469.

Kim, Y. S., Al Otaiba, S., Wanzek, J., & Gatlin, B. (2015). Toward an understanding of

    dimensions, predictors, and the gender gap in written composition. *Journal of Educational

    Psychology*, *107*(1), 79.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. London, UK:

    Guilford Press.

Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on

    dyslexia: A German-English comparison. *Cognition*, *63*(3), 315–334.

Lerkkanen, M. K., Rasku-Puttonen, H., Aunola, K., & Nurmi, J. E. (2004). The

    developmental dynamics of literacy skills during the first grade. *Educational

    Psychology*, *24*(6), 793–810.

Mäki, H. S., Voeten, M. J., Vauras, M. M., & Poskiparta, E. H. (2001). Predicting writing

    skill development with word recognition and preschool readiness skills. *Reading and*

    Writing, *14*(7–8), 643–672.

Malvern, D., Richards, B. J., Chipere, N., & Durán, P. (2004). *Lexical diversity and language*

    development. New York, NY: Palgrave Macmillan.

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition.

    Educational Psychology Review, *8*(3), 299–325.

McCutchen, D. (2006). Cognitive factors in the development of children's writing. *Handbook of*

    Writing Research, 115–130.

McMaster, K., & Espin, C. (2007). Technical features of curriculum-based measurement in

    writing a literature review. *The Journal of Special Education*, *41*(2), 68–84.

Olinghouse, N. G., & Leaird, J. T. (2009). The relationship between measures of

    vocabulary and narrative writing quality in second- and fourth-grade students. *Reading and*

    Writing, *22*(5), 545–565.

Osborne, J. W. (2013). *Best Practices in Data Cleaning: A complete Guide to Everything You*

    Need to do Before and After Collecting Your Data. Thousand Oaks, CA: Sage.

Puranik, C. S., & Al Otaiba, S. (2012). Examining the contribution of handwriting and spelling

    to written expression in kindergarten children. *Reading and Writing*, *25*(7), 1523– 1546.

Puranik, C. S., Lombardino, L. J., & Altmann, L. J. (2008). Assessing the microstructure of

    written language using a retelling paradigm. *American Journal of Speech-Language*

    Pathology, 17(2), 107.

Ravid, D., & Berman, R. A. (2010). Developing noun phrase complexity at school age: A text-

    embedded cross-linguistic analysis. *First Language*, *30*(1), 3–26.

Ravid, D., & Levie, R. (2010). Hebrew adjectives in later language text production. *First Language*, *30*(1), 27–55.

Salas, N. (2014). *Early development of text writing in two contrasting orthographies: English and Spanish* (Unpublished doctoral dissertation). Bangor University, Bangor, Wales.

Scott, C. M., & Windsor, J. (2000). General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *Journal of Speech, Language and Hearing Research*, *43*(2), 324.

Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143–174.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics.* Boston, MA: Allyn & Bacon/Pearson Education.

Wagner, R. K., Puranik, C. S., Foorman, B., Foster, E., Wilson, L. G., Tschinkel, E., & Kantor, P. T. (2011). Modeling the development of written language. *Reading and Writing*, *24*(2), 203–220.

Wechsler, D. (2009). *Wechsler Individual Achievement Test* (3rd ed.). San Antonio, TX: Pearson.

Table 1

Means and standard deviations of text-based measures for each language group at three time points

	English			Spanish		
	Time 1	*Time 2*	*Time 3*	*Time 1*	*Time 2*	*Time 3*
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Spelling	47.34 (23.77)	64.08 (21.15)	69.48 (19.37)	55.30 (26.66)	81.16 (12.28)	85.25 (9.58)
W. separation	48.41 (36.58)	68.58 (33.37)	79.02 (21.17)	54.88 (34.22)	87.79 (13.79)	92.13 (10.19)
Capitalization	65.43 (26.31)	75.28 (21.57)	77.19 (18.15)	64.71 (32.65)	91.70 (8.36)	94.10 (5.53)
No. of words	15.36 (8.52)	25.08 (12.57)	31.90 (15.89)	14.36 (8.41)	24.04 (8.55)	34.47 (13.29)
No. of letters	56.60 (31.21)	92.15 (46.15)	117.79 (57.82)	53.64 (29.83)	88.79 (30.35)	129.07 (50.46)
Punctuation	5.41 (8.68)	6.03 (6.53)	6.53 (6.28)	1.29 (4.49)	4.07 (5.43)	5.45 (4.99)
Connectors	4.07 (6.27)	5.73 (5.98)	8.13 (5.72)	6.42 (7.32)	9.34 (6.63)	10.55 (5.89)
Words/Clause	5.08 (1.30)	5.83 (1.90)	5.54 (1.20)	4.97 (2.52)	5.79 (2.03)	5.66 (1.57)
Subordination	5.77 (11.42)	5.78 (10.70)	12.43 (14.97)	6.59 (14.05)	9.04 (14.31)	12.58 (14.78)
Adj+Adv	2.26 (4.15)	4.73 (5.71)	6.18 (5.26)	1.94 (3.58)	3.69 (4.84)	4.83 (4.34)
Lexical density	46.48 (16.78)	43.96 (11.05)	46.02 (7.91)	48.78 (17.29)	44.17 (8.15)	46.02 (7.91)

Table 2

Standardized regression coefficient factor loadings for English at Times 1, 2, and 3[†]

	Time 1				Time 2			Time 3		
	Writing Conventions	*Productivity*	*Connectivity*	*Subord. + Adj+Adv*	*Writing Conventions*	*Productivity*	*Vocabulary*	*Writing Conventions*	*Productivity*	*Vocabulary*
Variance explained by factor	39.52%	14.15%	12.08%	10.25%	48.89%	17.21%	13.98%	44.72%	21.25%	13.31%
Spelling	.821	--	--	--	.897	--	--	.873	--	--
Separation	.853	--	--	--	.932	--	--	.913	--	--
Capitalization	.755	--	--	--	.792	--	--	.727	--	--
No. of words	--	-.731	--	--	--	-.763	--	--	-.843	--
No. of letters	--	-.722	--	--	--	-.786	--	--	-.852	--
Punctuation	--	--	-.800	--	--	--	--	--	--	--
Connectors	.665	--	.569	--	--	--	--	--	--	--
Subordination	--	--	--	.830	--	--	--	--	--	--
Lexical density	---	--	--	--	--	--	.767	--	--	.790
Adj+Adv	--	--	--	.771	--	--	.908	--	--	.875
Lexical diversity	--	.941	--	--	--	.909	--	--	.877	--

[†]Loadings < .50 are not displayed

Table 3

Standardized regression coefficient factor loadings for Spanish at Times 1, 2, and 3[†]

	Time 1		Time 2		Time 3		
	Writing Conventions	*Productivity*	*Writing Conventions*	*Productivity*	*Writing Conventions*	*Productivity*	*Vocabulary*
variance explained by factor	60.08%	21.66%	51.57%	25.12%	31.80%	27.03%	19.02%
Spelling	.689	--	.801	--	.869	--	--
Separation	.931	--	.887	--	.825	--	--
Capitalization	--	--	--	--	.765	--	--
No. of words	--	.912	--	.947	--	.982	--
No. of letters	--	.858	--	.905	--	.964	--
Lexical density	---	--	--	--	--	--	.800
Adj+Adv	--	--	--	--	--	--	.881
Lexical Diversity	--	-.908	--	-.782	--	-.719	--

[†] Loadings < .50 are not displayed

S-Table 1

Correlations between text-based measures at Time 1

	1	2	3	4	5	6	7	8	9	10	11	12	13
1.Spelling	--	.70*	.77*	.60*	.53*	.16	.35*	-.05	.12	.13	.25*	-.14	-.27*
2. W. separation	.43*	--	.59*	.46*	.42*	.30*	.37*	.01	.05	.16	.13	-.09	-.14
3. Capitalization	.32*	.21	--	.54*	.51*	.14	.29*	-.03	.06	.20	.20	-.07	-.16
4. No. of words	.57*	.29*	.33*	--	.96*	.09	.29*	.06	.22*	.12	.34*	-.53*	-.19
5. No. of letters	.54*	.33*	.34*	.96*	--	.09	.32*	.08	.19	.13	.31*	-.49*	-.02
6. Punctuation	.08	.18	-.02	.13	.15	--	.08	.00	-.01	.12	.07	.06	.06
7. Connectors	.35*	.27*	.22	.39*	.40*	.11	--	-.15	.14	.03	-.01	-.11	-.04
8. Words/clause	.23*	-.01	.08	.39*	.35*	.01	.07	--	-.10	-.10	.12	.05	.01
9. Subordination	-.05	-.05	-.02	.02	-.01	-.06	-.04	.07	--	.02	.31*	-.02	-.22*
10. Lex. density	-.03	.28*	.14	-.20	-.10	-.02	-.15	-.34*	-.24*	--	.21*	.16	-.06
11. Adj+Adv	.21	.25*	.11	.23*	.30*	.07	.21	.24*	.01	.05	--	-.01	-.32*
12. Lex. div.	-.31*	-.09	-.10	-.70*	-.61*	-.07	-.23*	-.33*	-.01	.33*	-.01	--	.03
13. Word length	-.13	.04	.06	.03	.21	-.01	.04	.12	-.11	.04	.06	.09	--

Note. Correlations for the English and Spanish samples appear above and below the diagonal, respectively.
*p < .01.

S-Table 2

Correlations between text-based measures at Time 2

	1	2	3	4	5	6	7	8	9	10	11	12	13
1.Spelling	--	.68*	.71*	.54*	.52*	.25*	.27*	-.09	.22	.18	.20	-.24*	-.15
2. W. separation	.43*	--	.63*	.41*	.40*	.23*	.19	-.02	.20	.24*	.25*	-.08	-.06
3. Capitalization	.35*	.32*	--	.53*	.52*	.21	.27*	-.06	.19	.28*	.37*	-.25*	-.12
4. No. of words	.29*	.13	.39*	--	.98*	.06	.29*	-.01	.27*	.35*	.33*	-.53*	-.08
5. No. of letters	.28*	.18	.40*	.94*	--	.07	.29*	.01	.27*	.36	.35*	-.48*	.05
6. Punctuation	.09	.14	-.09	-.12	-.03	--	-.16	-.07	.01	.22	.29	.05	-.09
7. Connectors	.01	-.01	-.02	.14	.08	-.18	--	.03	.07	-.01	.01	-.21	.13
8. Words/clause	.06	-.11	-.19	-.05	-.11	-.03	-.32	--	-.01	-.09	-.13	.08	.11
9. Subordination	.10	-.01	-.18	.30*	.26*	-.18	-.05	-.21	--	.03	.07	.05	-.01
10. Lex. density	-.22*	-.08	.12	.15	.15	-.11	-.05	-.15	.18	--	.47*	-.03	-.02
11.Adj+Adv	-.03	.09	.29*	-.06	.03	-.02	-.14	.20	.09	.47*	--	.08	-.21
12. Lex. div.	-.18	-.07	-.19	-.60*	-.54*	.16	-.15	.01	-.03	.21	-.03	--	.10
13. Word length	-.18	.13	-.01	-.17	.03	.16	-.13	.10	-.19	-.08	-.17	.11	--

Note. Correlations for the English and Spanish samples appear above and below the diagonal, respectively.
*p < .01.

S-Table 3

Correlations between text-based measures at Time 3

	1	2	3	4	5	6	7	8	9	10	11	12	13
1.Spelling	--	.65*	.65*	.42*	.41*	.15	.25*	.11	.09	.29	.34*	-.06	-.18
2. W. separation	.61*	--	.56*	.36*	.36*	.18	.21	.11	.18	.21	.32*	-.01	-.05
3. Capitalization	.53*	.43*	--	.38*	.37*	.11	.21	.07	.15	.35*	.44*	-.03	-.23*
4. No. of words	.03	.09	.13	--	.98*	-.26*	.29*	.25	.31*	.47*	.21	-.56*	-.23*
5. No. of letters	.05	.11	.13	.97*	--	.24*	.28*	.25*	.30*	.46*	.22*	-.54*	-.10
6. Punctuation	.04	-.02	-.08	-.21	-.18	--	-.20	-.06	-.03	-.06	.21	.29*	.24*
7. Connectors	.09	.10	.12	.10	.12	-.28*	--	-.01	.00	.06	.06	-.23*	-.04
8. Words/clause	-.09	-.21	.07	.04	.01	-.19	-.30*	--	-.03	.23*	-.17	.00	.02
9. Subordination	.08	.03	.22	.22	.21	-.11	-.04	-.24*	--	.22*	.06	-.08	-.13
10. Lex. density	-.06	-.06	-.25*	.01	.04	-.04	-.10	-.10	.01	--	.44*	.12	-.11
11. Adj+Adv	.28*	.35*	.10	-.10	-.03	.09	.01	-.42*	.10	.46*	--	-.07	-.24*
12. Lex. div.	.14	-.03	-.04	-.63*	-.56*	.11	-.09	-.10	-.02	.34*	.22	--	.09
13. Word length	-.05	-.05	.17	-.10	-.00	-.07	-.10	.19	-.09	-.14	-.19	.10	--

Note. Correlations for the English and Spanish samples appear above and below the diagonal, respectively.
*$p < .01$.