



Is spelling related to written composition? A longitudinal study in French

Pascal Bressoux, Bernard Slusarczyk, Ludovic Ferrand, Michel Fayol

► To cite this version:

Pascal Bressoux, Bernard Slusarczyk, Ludovic Ferrand, Michel Fayol. Is spelling related to written composition? A longitudinal study in French. *Reading and Writing*, 2024, 37 (3), pp.615-639. 10.1007/s11145-023-10414-6 . hal-03947334

HAL Id: hal-03947334

<https://hal.science/hal-03947334>

Submitted on 18 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Is spelling related to written composition? A longitudinal study in French

Pascal Bressoux^{1*}, Bernard Slusarczyk¹, Ludovic Ferrand² and Michel Fayol²

1-Univ. Grenoble Alpes, LARAC, Grenoble, France

2-Université Clermont Auvergne, CNRS, LAPSCO, Clermont-Ferrand, France

Reading and Writing (2023). doi: <https://doi.org/10.1007/s11145-023-10414-6>

CRediT author statement:

Pascal Bressoux: Supervision, Methodology, Formal analysis, Writing of results; Bernard Slusarczyk: Investigation, Resources; Ludovic Ferrand: Writing - Reviewing and Editing; Michel Fayol: Supervision, Conceptualisation, Writing original draft.

***Corresponding author:**

Pascal Bressoux
Université Grenoble Alpes
LaRAC - Laboratoire de Recherche sur les Apprentissages en Contexte
UFR SHS - Bâtiment Michel Dubois
1251 Avenue Centrale, CS 40700, 38058 Grenoble Cedex 9
FRANCE
Email: pascal.bressoux@univ-grenoble-alpes.fr

Abstract

This research aims at exploring in an irregular orthographic system like French, if spelling is related to written composition. French spelling is particularly interesting because it includes phonographic irregularities (i.e., inconsistencies), lexical difficulties and numerous morphological silent marks (e.g., plural noun, adjective, and verb agreement). In a longitudinal study from the beginning of Grade 3 to the end of Grade 4, pupils (N = 173) were asked twice in every grade to compose narrative texts from strips. Text length, text completeness and three categories of spelling errors were coded and analyzed through multilevel growth curve models. Results show (1) a growing relationship between text production and orthographic performance, (2) that this relationship varied according to whether it related to text length or to text completeness, and (3) its strength and evolution depended on the type of errors (phonographic, lexical, or morphological).

Keywords: spelling; text composition; lexical difficulties in spelling; morphological errors in spelling

Highlights

- * We assessed four times narrative text composition and spelling longitudinally on 173 children from Grade 3 to Grade 4.
- * We aimed to find whether individual spelling performance is correlated with narrative text length and completeness.
- * French orthographic system is difficult to learn and use because it includes phonographic irregularities, lexical difficulties, and numerous morphological silent marks.
- * Spelling difficulties impact progressively the length and completeness of narrative productions in Grade 4 but not in Grade 3.
- * Lexical difficulties are related to text length.
- * Morphological difficulties are related to text completeness.

Introduction

The main problem that children face when producing written composition concerns the coordination of composing components, while calling on only a limited pool of cognitive resources in the highly limited capacity of working memory (Fayol, 1999; McCutchen, 2000; Olive, 2004). In the present research, we assume that cognitive capacity is a single resource shared across all concurrent processes (Fayol, 1999; Kellogg, 1987; McCutchen, 1996; Olive, 2004; Swanson & Berninger, 1996; for a review see Torrance & Galbraith, 2006). In the context of such a highly demanding task, dealing with orthographic difficulties may add extra cognitive load which might reduce the cognitive resources to be allocated to the coordination of all other composing components and, *in fine*, alter the length and/or the elaboration of the content of the text produced. Orthographic difficulties are specific to the different languages, some of which are transparent while others are opaque. In the latter case, the cognitive demand implied by the orthographic processing is higher and has a negative impact on composition performance as illustrated by several studies focused on the English orthographic system (i.e. low consistency) (Berninger et al., 1992; Graham et al., 1997; Juel, 1988). To the best of our knowledge, very few studies have investigated the relationship between spelling and composing in French (but see Broc et al., 2013; Girolami-Boulinier, 1984; Maggio et al., 2014). These studies focused mainly on concurrent relationships and provided evidence that the complexity of written composition is negatively associated with spelling performances. The current study aims to go beyond the finding of concurrent relationships between spelling performance and text production and investigate the relationship between spelling and written text composition in French in a longitudinal perspective, exploring the possibility that low spelling scores could have a negative impact on both immediate and future composition performances in third to fourth graders. We elaborate on this question using growth curve models which allow to explore how variations in the spelling errors might explain variations in text production, and we do so over a relatively long period of two years which maximizes our chances to capture important changes in both abilities. Also, this modelling strategy allows to model the shape of change and to distinguish within-individual variance from between-individual variance. We have

retained this period of elementary schooling because Third graders no longer write by referring only to phoneme-grapheme connections. They already have memorized a large number of orthographic regularities (Pacton et al., 2014) and a substantial orthographic lexicon (Lété et al., 2008). In addition, they have received instruction on morphosyntactic agreements (see hereafter) (Fayol, 2014). It is thus possible to follow the evolution of the three categories of errors (Catach et al., 1980; Daffern, 2017) and that of text production in parallel.

The French orthographic system is highly complex in both the rather low consistency of the phonemes-graphemes correspondences – according to Veronis (1988), only 52.7% of French written words could be predicted from their phonological forms - and the frequent occurrence of “silent” letters most of the time associated to morphology or morphosyntax (Dubois, 1965; Jaffré & Fayol, 2005; Lété et al., 2008). The high cost of dealing with these difficulties during writing could have a negative impact on performance in text production, in terms of length and / or quality.

The production of narratives

Written text composition is a highly complex and dynamic process encompassing several interacting subcomponents (Hayes & Flower, 1980; Hayes, 1996): Planning involves setting goals, retrieving ideas from memory, and organizing contents into a writing plan; translating consists in gradually constructing the text as a linear sequence of linguistic units which are ordered hierarchically by level (e.g., words, phrases, sentences, paragraphs) (Fayol, 2018); transcription constrains the flow of translation through writing and spelling (Berninger & Swanson, 1994; Hayes & Berninger, 2014); reviewing includes monitoring and editing the text produced so far.

Most, if not all of these (sub)components, make some demands on cognitive resources. The writers must thus adapt to the competing demands of the different subcomponents by prioritizing and processing the different tasks (Kim, 2019; Kim & Graham, 2022). As composing is a complex and hierarchically organized goal-directed activity, most resources are devoted to the higher-order global control of the production processes, taking into account the audience, the concepts and their organization and the way the linearization could be realized (Fayol, 1997; Herrmann & Grabowski, 1995; Olive, 2004). Take as an example the case of producing a narrative, the simplest type of text (Alamargot & Fayol,

2009; Chanquoy et al., 1990). The writers have to activate prior knowledge (i.e., the situation model) enabling them to elaborate the text content (i.e., the events and their unfolding over time as well as the characters involved and the places where the events occur) in order to ensure the coherence of the to be composed text through translation. The production then dynamically evolves over time due to the continuous updating of new information in relation to what has been written so far and what remains to be introduced as a function of the content and the potential audience. This controlled processing regarding planning and translating is slow and demands focused attention and conscious mental effort (McCutchen, 2011). Developmental studies reported that 6 to 10 year-old children are often able to answer questions about characters, motivations, goals and events (Trabasso et al., 1992) but find it difficult to mobilize these knowledges to produce coherent and well sequenced stories (Fitzgerald, 1984), with important and persistent individual differences (Peterson & McCabe, 1983). In order to prevent too difficult remembering of events, characters, places and risks of too short texts including too limited contents, we asked pupils to produce texts from series of images (kinds of comics). We thus provided pupils with a series of images aimed at enabling them to find the unfolding of events. In addition to activate, update and organize event knowledge, the writer also has to bring different lower-level processes or skills to bear, such as transcribing or grammatical encoding. At the lower levels, letters and words are activated enabling productions at the sentence level. Words are semantically integrated into coherent and meaningful syntactic representations. Furthermore, higher-order global coherence must be established among sentences and paragraphs to obtain an integrated representation at the text level (Kim, 2019).

Higher-order domain and discourse knowledge and lower-order transcription skills involved in text production compete for a limited pool of resources. Under certain task conditions, an inefficient (sub)component can disrupt performance because it draws resources away from focused activities. Such disruptions can affect either higher-order components or lower-level processes. For example, handwriting skills impact higher levels of processing (Berninger & Swanson, 1994; Bourdin & Fayol, 1994, 2002; Connelly et al., 2005; Limpo & Alves, 2013): low handwriting negatively affects fluency and quality of text production (Beers et al., 2017).

To prevent such disruptions, several studies have tried to reduce the costs of some components in order to make composition tasks more manageable. Some studies have been working on strategies, changing the relations between the components of the tasks and enabling self-regulated and adapted moves from parallel to serial processing (Graham, Harris & Mason, 2005; Limpo & Alves, 2013). Some other attempts have concentrated on improving transcription processes: increasing writing speed and quality had led children to produce longer and better quality texts (Alves, Limpo & Fidalgo, 2015; Christensen, 2005; Graham, Harris & Kink, 2000). Generally, these attempts have been successful. The load of a lower-level component (e.g., handwriting) through increasing the difficulty of managing the output modality was thus shown to affect a higher-level component, the length and quality of a text (Alves et al., 2018).

The French spelling system

Few research have been devoted to spelling, especially regarding French. As far as English is concerned, Juel (1988) reported that from Grade 1 to Grade 4 spelling skills were related to written composition and explained 29 % of the variance in quality of writing products in Grade 1 and 10 % of the variance in quality of writing products in Grade 4, when spelling skill had improved compared to Grade 1. Results of several other studies support the same conclusion in English-speaking children that spelling is related to written composition (Abbott et al., 2010; Alves et al., 2018; Berninger et al., 1992; Berninger et al., 2011; Graham et al., 1997) and can affect written text fluency and quality (Kent et al., 2014; Kim et al., 2011; Kim et al., 2015). These results are in line with conceptions assuming that resources likely to be allocated to the different components of the writing process are restricted by non-automated skills such as handwriting or spelling that require attention and control (Kellogg, 2001; McCutchen, 2000). As the English orthographic system is highly complex and difficult to learn and use, it most likely imposes heavy resource demands during production, even in adults. The situation is undoubtedly much simpler in the case of more regular spelling systems. The Spanish, Italian and Finnish orthographic systems are considered more transparent and thus easier to learn and use (Joshi & Aaron, 2006). For example, Arfé, Dockrell and De Bernardi (2016) reported that spelling skills contributed to accuracy and quality of written composition in Italian children but only in Grade 2 and Grade 3. As far as we know, few data are available regarding the French orthographic system and its potential

impact on text production characteristics in children (Girolami-Boulinier, 1984; Maggio et al., 2014).

The study of learning and using the French spelling system is particularly interesting because of its properties which make it one of the most complex systems to master. French is a Romance language in which words include a lexical component combined with a grammatical one. The French orthographic system for spelling is an alphabetic system of phoneme-grapheme correspondences (PGC) closely related to its oral language. It relies on two general principles: the phonographic principle and the semiographic principle (Fayol, 2014; Jaffré & Fayol, 2006). The phonographic principle establishes correspondences between graphic units (graphemes: G) and phonemes (P). Graphemes are more numerous than phonemes: one phoneme can thus be associated with several graphemes (Catach, 1980). For example, the French phoneme /E/ can be spelled differently (“è”, “ê”, “ai”, “est”, “et”, etc.) (Brissaud & Chevrot, 2011). In reading (from G to P), the G-P associations are relatively consistent (i.e. are almost always read the same way), and as a consequence, reading is rather easy to learn in French (Seymour et al., 2003). In spelling (from P to G), however, the P-G relations are much more complex (i.e. inconsistent: Bonin et al., 2008; Ziegler et al., 1996). A simple application of phoneme analysis is thus not sufficient to identify the corresponding orthographic units for spelling the majority of French words. Writers must also refer to lexical orthography (Martinet, Valdois & Fayol, 2004), statistical regularities, morphological knowledge (Pacton, Fayol & Perruchet, 2005), and morphosyntactic markers (Jaffré & Fayol, 2005; Weth, 2020). Spelling many French words requires mobilization of associations between sound and letter units larger than the PG correspondences and/or access to an orthographic lexicon of words with memorized and directly retrievable orthographic forms.

As in all alphabetic systems, it is possible to generate a phonologically plausible spelling of French words by applying phoneme to grapheme correspondences. These correspondences are explicitly taught in the first year of elementary schooling and allows pupils to decode many words. According to Share (1995, 2004), decoding (i.e. reading) becomes the “engine” of subsequent learning. It leads to the memorization of orthographic forms, those of words and also those of sublexical regularities according to their frequencies. For example, pupils learn very early that the phoneme /o/ is transcribed “eau” at the end of

words. However, even the most correct application of these correspondences allows to produce at most half of the French words (Ziegler et al., 1996). Some phonemes can be transcribed in several ways (e.g., /o/ -> o, au, eau) and some letters have no phonological counterpart (e.g., h in “théâtre” theater). Sources of information other than phoneme to grapheme correspondences must be used to spell most French words in a conventional way. This can be the spelling of specific words (e.g., the word “landau” is written with “an” and “au” and not with “en” and “eau”) that must be memorized (Fayol et al., 2013). This can be sequences of frequently associated letters (e.g., ard ; oire) so-called graphotactic patterns, for example that double consonants are never doubled at the beginning or ending of French words (Pacton et al., 2014). Lastly, this can be morphological or morphosyntactic informations: that “bavard” ends with a silent “d” because it belongs to a word family such as “bavarde, bavarder, bavardage” (Pacton et al., 2013); that plural nouns and adjectives end with a silent -s. The morphosyntactic marks and their use are explicitly taught from the second grade onwards and take time to become automatized (Fayol et al., 1999, 2006). The acquisition of words forms and of regularities takes times and remains difficult even in adults, leading to frequent and different error types.

Lexical production is a costly process; its cost depends on the difficulty of retrieving more or less frequent or familiar words from memory (Bonin et al., 1997). Frequent and consistent words are the easiest to retrieve and write down; however, rare and inconsistent words raise many issues: the processing difficulties related to spelling irregular words affect the time it takes to produce the handwritten responses, the latency (i.e. pause) and, sometimes the handwriting speed (Delattre et al., 2006; Foulon & Chanquoy, 2006; Maggio et al., 2012), in children as well as in adults. These changes in the management of the text production are expected to increase the cost of the orchestration of the involved components, and to impact the fluency and/or the quality of the product. For example, the slowing down of word production or the difficulty to choose a mark (e.g., between -s or -nt at the end of “timbre” [stamp]) can attract attention and affect retention of information already present in working memory (graphemic buffer) and even cause loss of thread of the narrative frame.

Regarding the semiographic principle, most words in French are composed of several morphemes (about 75% according to Rey-Debove, 1984). Morphological units increase the

degree of consistency characterizing the PG and GP mappings between spelling and sound. They thus help learners and users of French orthography to spell words, lexical morphology. Derivational morphological knowledge helps to spell correctly in a number of cases, for example, when words end with silent final letters, the most inconsistent part of words in French (Lété et al., 2008, e.g., “grand” [tall]). Referring to word families (“grande” fem [tall]; “grandeur” [greatness], etc.) leads to putting a final “-d” despite the lack of any phonological cue. Children can detect and apply morphological regularities through text including word families (Pacton et al., 2013) or through explicit teaching and training (Casalis et al., 2018). From 3rd grade pupils already have available a lot of morphological knowledge that will continue to increase, which is why we have chosen this school level to start our longitudinal study. However, referring to word families or searching for a derived word (e.g., “grande” fem [tall]) to correctly spell a related word (e.g., “grand” [tall]) captures some attention at the expense of retaining other information and updating ideas.

French has also a rich written morphosyntax (inflectional morphology), syntactic markers, serial graphemic marks that create coherence within phrases and between words or word groups on the clause level (Dubois, 1965; Weth 2020). However, the plural and feminine endings of nouns, adjectives, and verbs are rarely pronounced (Dubois, 1965). For example, “la poule” ([hen], singular noun) et “les poules” (plural [hens]) are pronounced the same way, as are the singular and the plural forms of the adjective “rousse” (“rousse” versus “rousses”; adj [red]) as well as the singular and plural forms of most verbs such as “picore” versus “picorent” ([is pecking] versus [are pecking]) (Largy & Fayol, 2001). All these specificities make the learning of French spelling both a complex and a long process. Only when readers and spellers of French have well established lexical representations can they refer to morphology and morphosyntax (Pacton et al., 2013; Quémart & Lambert, 2019).

As far as we know, there are mainly indirect indications that the management of written morphosyntax is cognitively costly (see Girolami-Boulinier, 1984; Maggio et al., 2018; Van Reybroeck & Hupet, 2009). Using a dual-task design involving the completion of a primary cognitive task (i.e., an agreement task) while simultaneously completing a secondary distractor task (in this case a working memory task), Fayol et al. (1994) and Largy et al. (1996) have induced agreement errors on verbs in highly educated adults having to write

down orally presented sentences. Due to the silent characteristics of written French morphology, adults inflected erroneously verbs such as “arrive” [arrive] when they had to both write a sentence like “Le chien des voisins arrive” [The dog of the neighbours arrives] and to temporarily store in working memory a series of five words: they wrote “arrivent” instead of “arrive”, making the verb agree with the closest (local) noun instead of making the correct subject-verb agreement. The same phenomenon occurred in children production. French children (e.g., Grade 1 to Grade 5) were able to use the relevant noun and verb inflections when explicitly asked to do so. However, they “omitted” to add these inflections when they had to recall (by writing down) the same linguistic patterns or to produce them when they composed a text (Fayol et al., 1996; 1999; Totereau et al., 1997).

Even the language production process of expert adults may become disrupted, although highly automated, when they have to deal with a secondary task, which attests the fragility and the cognitive cost of agreement processes in written French. Because they need to shift from higher-order components (e.g., retrieving or creating ideas or controlling for coherence) to lower-order skills (e.g., computing a distant subject verb agreement; looking back to the gender of a far subject, etc. see Alamargot et al., 2016) or because they have to temporarily focus their attention onto determining the ending of a word (e.g., “renard” – masc [fox] - takes a final “d” due to the feminine “renarde” – fem [fox]) skilled writers and more so novice writers can lose access to resources and knowledge to cope with writing demands. As a consequence, even good writers may fail to establish consistency or continuity between parts of their texts, leading to a decline in the quality of their textual production. The situation is much more complicated for poor spellers who have to deal very often with these concerns. We can thus expect that some information is missing in their texts resulting in weak coherency.

The present study

The study reported here explored the concurrent and longitudinal relationships between spelling and written text composition in a group of 173 pupils. These pupils have been followed from the very beginning of Grade 3 to the end of Grade 4 (20 months). These grades have been selected because associations between phonemes and graphemes are largely mastered (hence phonographic errors are rare), orthographic lexicon already contains a large number of words but the quality and precision of word forms remain rather

low (Lété et al., 1998). The difficulty, and the cognitive cost, are still more important for children who have a smaller and less precise orthographic lexicon (Bosse et al., 2020; Van Reybroeck & Hupet, 2009). In most cases, errors are phonologically plausible because they involve correct sound-spelling correspondences but not the standard correct spelling (e.g., “anfert” or “anfant” instead of “enfant” [child]). As a consequence, accessing and producing lexical items remains effortful for most children: the attention and time they have to devote to finding words in memory and deciding how to spell them could lead them to strategically restrict the length of their texts and/or to disrupt the flow of the production processes. Regarding morphology and agreements, written morphological derivation and inflection marks are explicitly taught from Grade 2, and become more frequent from Grade 3 onwards and therefore likely to influence the length and quality of written productions. For example, Pacton et al. (2005) observed that using the final “-eau” (/o/) to indicate the diminutive derivation (e.g., “un petit renard est un renardeau” [a small fox is a fox cub]) increases from Grade 3 onwards. The same trend was true regarding noun, adjective, and verb agreements inflection marks (Fayol et al., 1999, 2006). However, experimental studies provided evidence that even in educated adults the management of agreement remains highly costly: errors in inflections occur as soon as agreements are produced under attentional pressure, and more so in children (Fayol et al., 1999). The management of the agreement processes constrains the writers to keep in mind some cues about gender or number and to refer back to them when necessary: these comings and goings between what is already processed and what must be processed at a given time (see Alamargot et al., 2016 for illustrations) may have a negative impact on the fluidity of production processes, to the point of causing the loss of information in memory or the difficulty to ensure the links between ideas (i.e., the consistency of the text). The difficulties raised by lexical processing and by morphological processing differ (Morin et al., 2018). Some weak spellers produce mostly lexical errors, others mostly errors in agreeing nouns, adjectives and verbs. Hence the need to distinguish the types of errors committed and their relationship to the fluency and quality of the texts.

Pupils had to compose four narrative texts, two in Grade 3 (T1 and T2), and two in Grade 4 (T3 and T4), respectively at the beginning (T1 and T3) and end (T2 and T4) of the school years. All these texts were prompted by a series of 6 images describing different topics but following the same frames. These frames have been selected to make text

organizations as similar as possible to elaborate assessment grids (i.e., completeness) and thus facilitate comparisons. The length of every text has been evaluated by the number of words, and the three categories of spelling errors – phonographic, lexical, and morphological – have been collected. The ratio of each spelling error category has been computed based on the total number of words. The main hypothesis was that spelling performance would positively impact on text length and text completeness both concurrently and longitudinally. Two questions were addressed. First, are the scores for each type of error positively correlated with each other concurrently and longitudinally? Second, is spelling performance in T1 and T2 (Grade 3) and T3 and T4 (Grade 4) concurrently and longitudinally positively correlated to text length and text completeness? An empirical question remained: Are some categories of spelling errors more influential than some others?

Method

Participants

An initial sample of 263 pupils from 16 classrooms (from different schools in the French Département de la Haute Loire) was selected to be followed up from the beginning of Grade 3 (age 8-9) through to the end of Grade 4 (age 9-10). This is a convenient sample which captures socio-economic contexts that can be considered average. No schools were located in priority educational areas. Due to changes in teachers, moving out of the household and absences of pupils during the two school years, only 173 pupils (87 girls) from 15 classrooms participated in all the tests. The pupils whose productions were analyzed were all of normal age, without repeaters or pupils in advance.

Potential selection bias was investigated. All available variables at T1 were utilized as potential predictors of attrition in a logistic regression model. The results showed that none of the variables had a significant effect on the probability to leave the cohort, except for text quality: On this variable, pupils who left the cohort scored a little bit lower than their remaining counterparts. However, the odds ratio is weak ($OD = 1.28$). Furthermore, fit indices indicate poor adjustment of the model ($\text{Tau-}a = 0.12$) and a reduction of the deviance from an intercept-only model of only 5.6% (Menard, 2002). Therefore, selection bias was considered negligible.

Material and procedure

Four tests were used. The first one at Time 1, T1, at the beginning of Grade 3, was the production part of the standardized ECL-Collège test (Khomsy et al., 2005); the three others (Story 2 at Time 2, T2, end of Grade 3; Story 3 at Time 3, T3, beginning of Grade 4; Story 4 at Time 4, T4, end of Grade 4) were developed by the authors of this paper. To avoid pupils having to compose several times from the same prompts, these four tests followed the frame of the ECL test but evoked different topics. The similarity of the frames made the successive text productions easier to compare and score. In the four cases, a story in 6 images was presented on the same page. To facilitate the planning of the content of the text, the pupils had first to describe in writing each of the images, using some dedicated space under each image. Second, once completed this first phase, they had to write the story told by the strip on a specific place at the bottom of the same page. There were no time limits for either of the two writing tasks. Only the composed text was analyzed and rated. The same judge (the second author of the paper) scored all the texts regarding their length, completeness, and their spelling errors. The coding grid was very precise, and almost no room was left for interpretation. This is of course the case for text length (number of words). As for text completeness, the type and number of pieces of information (illustrated facts and inferences) were clearly determined: The rater only coded their absence/presence. As for spelling errors, their classification is almost unambiguous: The three categories of errors can be clearly distinguished.

Text length

For each text, the number of words was computed to determine text length. Text length was unambiguous since it was based on the number of words, which were identified by the blanks separating them.

Text completeness

Text completeness was scored using the ECL rating grid (Khomsy, Nanty, Parbeau-Guénou & Pasquet, 2005): Thirteen pieces of information were considered essential. Amongst them, 9 described illustrated facts (e.g., there are two characters; two phones; they use tools; etc.). These factual elements were unambiguously present in the images. In addition, 4

pieces of information constituted inferences (e.g., he is jealous; he is happy; etc.). The inferred elements correspond to information that cannot be obtained by a simple observation of the images. Inferred elements are processing of one or more pieces of information from one or several images. This processing leads to infer/generate additional information, not exposed directly in the images and useful for the flow of the story, ensuring its coherence. The same analysis grid including each time 13 items (9 + 4) was established for the other three stories. The correlations between factual element scores and inferential element scores were $r = 0.36$ ($p < 0.0001$) at T1, $r = 0.24$ ($p < 0.01$) at T2, $r = 0.63$ ($p < 0.0001$) at T3, and $r = 0.37$ ($p < 0.0001$) at T4. So, the two scores were summed up to form a global text completeness score at each time.

Spelling errors

All spelling errors were collected in each of the 173 texts and categorized following the classification from Catach et al. (1980): phonological (e.g., writing “bado” /bado/ instead of “bato” /bato/), lexical (writing “retart” instead of “retard”) and morphosyntactic word forms (writing “les timbre” or “les timbrent” instead of “les timbres” [stamps]) (see also Daffern, 2017; Richards et al., 2009). If a word contained several errors, each error was coded. Handwriting problems were very marginal: in the rare cases a word could not be read, no errors were coded. One spelling error variable has been calculated for each type of error: phonographic, lexical, or morphosyntactic. Each variable expresses the percentage of errors of a given type compared to the total number of words written in the text (number of errors of a given type divided by total number of words multiplied by 100) (see supplementary online material, Table S1).

Statistical analyses and modelling strategy

We used multilevel growth curve models using SAS 9.4 software. This allowed us to take into account the nesting of measures within pupils and the nesting of pupils within classes. Furthermore, this allowed us to model not only the final achievement but also the shape of the growth curve for the whole observed period, and to distinguish within-individual and between-individual variances.

We first specified a three-level unconditional growth curve model with measures at level 1, pupils at level 2 and classes at level 3 (where the class variable indicates pupils’

Fourth-grade class belongings). No class effect was found for text completeness, so classes were removed from subsequent analyses. However, a significant class effect was found for text length. Therefore, the models retained are a two-level growth curve model with measures at level 1 and pupils at level 2 for text completeness and a three-level growth curve model with measures at level 1, pupils at level 2 and classes at level 3 for text length.

Time was coded as the number of months elapsed since the beginning of the study. Therefore, the values for time are 0 at T1 (beginning of Grade 3), 8 at T2 (end of Grade 3), 12 at T3 (beginning of Grade 4) and 20 at T4 (end of Grade 4). We tested for non-linear effects using a quadratic variable of time. No quadratic effect was found for text completeness ($\Delta -2\log L = 2.2$ for 1 df; $p > .05$) and a significant quadratic effect was found for text length ($\Delta -2\log L = 6.8$ for 1 df; $p < .01$) meaning that change accelerates over time. However, the curvature is not very pronounced.

We tested for random effects of time. We specified the models so to test for random effects of the linear and quadratic variables of time and to test for the variance/covariance pattern that fitted best the data. As for text completeness, a significant random effect of (linear) time was found ($\Delta -2\log L = 21.2$ for 1 df; $p < .001$) at the pupil level (remember that there is no significant class effect). The covariance between the intercept and time slopes was not significant ($\Delta -2\log L = 1.4$ for 1 df; $p > .05$). As for text length, a significant random effect of (quadratic) time was found at the class level ($\Delta -2\log L = 21.9$ for 1 df; $p < .001$) and a significant effect of (linear) time was found at the pupil level ($\Delta -2\log L = 21.0$ for 1 df; $p < .001$). Specifying covariances between intercepts and slopes caused problems in the estimation process because the covariance matrix was not positive definite, indicating the random part was too complex for the data. Given that the covariance matrix was, in our case, a matter of statistical adjustment rather than a substantial issue, we decided to remove the covariances from the models.

After the random part of the model was determined, we specified the fixed part of the model using the same three-step strategy for each dependent variable. Model 1 was an unconditional growth curve model. Model 2 included all our independent variables (without interactions); apart from spelling errors, gender was also included since it might capture a part of the dependent variable variance. Model 3 was the final model including significant

interactions with time, in order to model differences in growth rate. Only significant interaction effects are presented.

Results

Descriptive statistics

Table 1 shows the means, standard deviations and correlations for the variables at each period time. Text length and text completeness increased over time while the number of errors tended to decrease. The text length standard deviations increased over time but this was expected since the means also increased over time. The coefficients of variation $((SD/Mean)*100)$ remained fairly stable over time, meaning that standard deviations and means increased proportionally.

The text completeness standard deviations increased over time, although not monotonically. The coefficients of variation tended to increase slightly over time.

The pattern of correlations between text length, text completeness and the different types of errors changed over time. As time passed, correlations increased. At T1, the correlations were low or not significant; they were generally significant and stronger at T4. The correlations between text length and text quality did not increase monotonically because the highest correlation appears at T2 ($r = .43$).

The total number and the proportions of orthographic errors decreased from T1 to T4. At T1, about one word out of three (21.3/60) was erroneously transcribed, about 50% of the errors were lexical (9.6) and 50% morphological (9.87). At T4, only one word out of 12 (14.6/121,8) was erroneous with 57.5% of them being morphological. The correlations between the different types of errors increased regularly over time. This increase suggests that orthographic acquisition became less type-dependent as time passed: Indeed, at T4, pupils who produced one type of error also tended to make other types of errors. The highest correlations were between lexical and morphological errors, reaching $r = .54$ at T4.

The correlations between text completeness and types of errors increased over time. This was true for any type of errors: From $r = -.06$ (ns) to $r = -.21$ for phonographic errors; From $r = -.04$ (ns) to $r = -.23$ for lexical errors; From $r = -.10$ (ns) to $r = -.31$ for morphological

errors. The same correlation pattern did not appear regarding text length: The increase over time was only valid with lexical errors (from $r = -.11$ (ns) to $r = -.20$).

Table 1. Descriptive Statistics for the Different Variables and Their Correlations

	Means	SD	2	3	4	5	6
Third grade T1							
1 Length (words)	60	22	.25**	-.06	-.11	-.01	-.08
2 Completeness (/13)	5.7	1.9	-	-.06	-.04	-.10	-.05
3 Phonographic errors	1.8	3.5	-	-	.22**	.23**	.55**
4 Lexical errors	9.6	6.4	-	-	-	.32**	.80***
5 Morphosyntactic errors	9.87	5.5	-	-	-	-	.74***
6 Total errors	21.3	11	-	-	-	-	-
Third grade T2							
1 Length (words)	65.47	28.5	.43***	-.03	-.03	.01	-.02
2 Completeness (/13)	6	1.7	-	-.16*	-.10	-.13	-.18*
3 Phonographic errors	1.3	2.9	-	-	.31**	.10	.50***
4. Lexical errors	6.9	5	-	-	-	.38***	.81***
5 Morphosyntactic errors	8.5	5.7	-	-	-	-	.79***
6 Total errors	16.7	9.98					

Table 1 (continued): Descriptive Statistics for the Different Variables and Their Correlations

	Means	SD	2	3	4	5	6
Fourth grade T3							
1 Length (words)	104	38	.32***	-.20**	-.19**	-.01	-.16*
2 Completeness (/13)	6.3	2.7	-	-.20**	-.17*	-.19**	-.24**
3 Phonographic errors	0.8	2.3	-	-	.27**	.26**	.53***
4 Lexical errors	7.3	5.5	-	-	-	.40**	.84***
5 Morphosyntactic errors	9	4.6	-	-	-	-	.78***
6 Total errors	17.1	9.5	-	-	-	-	-
Fourth grade T4							
1 Length (words)	121.8	43.5	.35***	-.08	-.20**	-.10	-.17*
2 Completeness (/13)	7	2.6	-	-.21**	-.23**	-.31***	-.32***
3 Phonographic errors	0.7	1.9	-	-	.32**	.32**	.58***
4. Lexical errors	5.4	4.5	-	-	-	.54***	.85***
5 Morphosyntactic errors	8.4	4.5	-	-	-	-	.86***
6 Total errors	14.6	8.9					

Estimates of the multilevel growth curve models

Text length

As for text length, Model 1 (Table 2) showed that the number of words increased significantly over time and this change accelerates over time although the curvature is not very pronounced: Each additional month corresponded to an average growth of about 2 words at the beginning of Grade 3 and of about 4 words by the end of Grade 4. Over the 20-month period, text length gained 62 words on average.

Model 1 also showed a significant level 2 time-slope variance which indicated that the growth rate varied between pupils. Over time, the between pupil variance increased, highlighting that heterogeneity between pupils was far higher at the end of Grade 4 than at the beginning of Grade 3. The significant level 3 (quadratic) time-slope variance indicated that the growth rate varied between classes and these differences are more pronounced at the end of the period under study.

Model 2 showed that girls wrote texts that were about 13 words longer than boys did. Regarding the relationships between text-length and orthographic errors, Model 2 also showed that on the one hand, phonological errors tended to be associated with text length although the relationship was only marginally significant ($p < .06$). On the other hand, lexical errors and morphological errors were not associated with text length.

Model 3 revealed a significant interaction between lexical errors and time: As time passed, a negative impact of the number of lexical errors appeared and increased. While its impact on text length was negligible (and not significant) at the beginning of Grade 3 (T1), it became substantial at the end of Grade 4 (see Figure 1). Note that the effect of phonological errors turned significant.

Table 2: Models explaining text length (standard errors within parentheses)

Parameters	Model 1	Model 2	Model 3
<i>Fixed effects</i>			
Intercept	59.783 (3.376)***	56.212 (4.661)***	51.726 (5.174)***
Time	2.094 (0.461)***	1.936 (0.471)***	2.634 (0.586)***
Quadratic time	0.052 (0.026)*	0.054 (0.026)*	0.042 (0.027)
Girl		13.548 (2.985)***	13.739 (2.990)***
Phonological errors		-0.912 (0.468) ^(t)	-0.970 (0.470)*
Lexical errors		-0.278 (0.248)	0.144 (0.329)
Morphosyntactic errors		0.088 (0.257)	0.125 (0.258)
Time*Lexical errors			-0.064 (0.032)*
<i>Random effects</i>			
Level 3 (classes)			
Intercept variance	87.248 (60.779)	84.5903 (54.761)	84.6729 (54.849)
Quadratic time slope variance	0.0029 (0.0015)	0.0028 (0.0014)	0.0027 (0.0014)
Level 2 (pupils)			
Intercept variance	125.840 (56.259)	71.878 (49.370)	81.3337 (49.926)
Time slope variance	1.5696 (0.419)	1.336 (0.384)	1.1755 (0.3731)
Level 1 Within pupil variance	731.01 (49.032)	746.02 (50.236)	748.180 (50.320)
-2 log L	6771.22	6746.73	6742.90

Fixed effects: ^(t) p<0.10; * p<0.05; ** p<0.01; ***p<0.001

Insert here

Fig. 1 Text length as a function of time and lexical errors

Text completeness

Model 1 (Table 3) provided evidence that text completeness significantly increased on average over time. The significant time-slope variance indicated that the growth rate in text completeness varied between pupils. Over time, the between pupil variance increased providing evidence of more heterogeneity between pupils at the end of Grade 4 than at the beginning of Grade 3.

Model 2 included all our independent variables. Note that it also included text length because, even though the number of errors was moderately related to the number of words in the text, we wanted to separate these two effects. Model 2 showed that, controlling for the number of errors and text length, text completeness remained approximately equal on average for girls and boys. The number of phonological errors and the number of lexical errors were not related to text completeness. However, the number of morphological errors was significantly and negatively related to text completeness: the more morphological errors, the poorer the text completeness. Text length was significantly and positively related to text completeness: the longer the text, the better the text completeness.

Model 3 revealed a significant interaction between morphological errors and time: the relationship between the number of morphological errors and text completeness increased over time. Although it was not significant at the beginning of Grade 3, as time passed, the number of morphological errors became more detrimental to text completeness as shown in Figure 2.

Table 3: Models explaining text completeness (standard errors within parentheses)

Parameters	Model 1	Model 2	Model 3
<i>Fixed effects</i>			
Intercept	5.624 (0.136)***	5.275 (0.265)***	4.846 (0.3014)***
Time	0.063 (0.011)***	-0.0062 (0.013)	0.053 (0.02351)*
Girl		-0.139 (0.199)	-0.166 (0.2012)
Phonological errors		0.007 (0.031)	0.0003 (0.031)
Lexical errors		-0.013 (0.016)	-0.014 (0.016)
Morphosyntactic errors		-0.054 (0.016)***	-0.002 (0.023)
Text length		0.019 (0.002)***	0.019 (0.002)***
Time*Morphosyntactic errors			-0.007 (0.002)**
<i>Random effects</i>			
Intercept variance	0.868 (0.251)	0.615 (0.209)	0.531 (0.321)
Time slope variance	0.0064 (0.0017)	0.0042 (0.0013)	0.005 (0.001)
Within pupil variance	3.189 (0.214)	3.011 (0.203)	2.690 (0.427)
-2 log L	2995.74	2920.25	2911.42

Fixed effects: * p<0.05; ** p<0.01; ***p<0.001

Insert here

Fig. 2 Text completeness as a function of time and morphological errors

Discussion

Faced with a task of producing a narrative text, French elementary pupils must mobilize and coordinate in working memory various low-level and high-level components, each having a cognitive cost (Fayol, 1999; McCutchen, 2000; Olive, 2014). Among these components, spelling raises many difficulties because the French orthographic system is as inconsistent as the English one in the Phoneme-Grapheme direction: many lexical items are difficult to spell. In addition, French morphosyntax has many silent marks for gender (“-e”) and number (“-s”, “-nt”) (Jaffré & Fayol, 2005). These characteristics contribute to a double disadvantage (Connelly & Dockrell, 2016): the low consistency between phonological and orthographic codes forces writers to slow down to find ways to spell words; to prevent themselves to produce spelling errors, writers tend to replace words they cannot spell by simpler words, which decreases lexical diversity. Such linguistic operations lead pupils to shift from time to time their attention from the dominant phonographic dimension to the orthographic dimension alone. They have to devote attention and cognitive resources to process spelling difficulties. As they have limited cognitive capacities, the resources devoted to spelling treatments reduce the resources available for other high-level treatments. Consequently, it is plausible to expect a negative relationship between the spelling performance of pupils and the characteristics of the texts they produce. The objective of this research was precisely to highlight this relationship through a short longitudinal study on a fairly large population of Third to Fourth graders. The design of the study offers a long period of observation (two school years) and provides multiple measurements of pupils’ writing skills. Furthermore, this allows to use multilevel growth curve models which partition pupils’ outcomes in between-context, between-individual and within-individual effects.

The results show a regular increase in the length of the texts between T1 and T4 as well as a more modest increase in the completeness of the content essentially between

Third and Fourth grades. These developments are consistent with what is reported in the literature on the development of written narrative (Alamargot & Fayol, 2009; Fayol et al., 2012). Despite the presence of the series of images, the elaboration by the pupils of the related content remains poor. At the same time, the proportions of errors tend to decrease slowly in all categories, more clearly for phonological errors and lexical errors, less strongly for morphosyntactic errors. Surprisingly, contrary to what was expected, the correlations between either text length or text content and the proportions of errors are very weak: close to zero in T1 and weakly significant with the content in T2. It is only in T3 and T4 that these correlations become systematically significant, especially with text content. These results raise the question of the relationship established by students during the production of texts between the management of content, fluency and spelling difficulties.

The results of the growth-curve models make clear that there is a growing concurrent and longitudinal relationship between text production and orthographic performance, from very weak in T1 to strong correlations in T4. In addition, this relationship also varies according to whether it relates to text length or to text completeness. Lastly, its strength and evolution depend on the type of errors: phonographic, lexical, or morphological.

Multilevel growth curve models reveal several trends. First, from T1 to T4, there was a large increase in both text length and related individual differences. Texts became longer on average but with important variability among pupils. Second, text completeness also improved with time, although more slightly, also with an increasing heterogeneity. In addition, text length and text completeness were positively correlated: The longer the texts the better the completeness. These results do not come as a surprise. They correspond to trends already reported in previous research (Girolami-Boulinier, 1984; Maggio et al., 2018). The number of errors decreased, and the pattern of errors changed: The proportion of lexical errors became rarer while that of morphological errors increased (Andreu & Steinmetz, 2016). In addition, lexical and morphological errors were increasingly correlated. These trends were in line with results from French national assessments and provided evidence for increasing individual differences, a few pupils cumulating the two categories of errors while others produced almost error-free texts (Andreu & Steinmetz, 2016; Morin et al. 2018).

Our first question was about the relationship between spelling errors and the length and completeness of texts. We hypothesized that pupils producing the greatest number of errors would also be those whose texts would be the shortest and of lower completeness. This result was clearly attested over time but not at the very beginning. To illustrate, all the negative correlations between the total number of errors and the characteristics of the texts evolved for text length from $-.08$ (ns) at T1 to $-.17$ ($p < .05$) at T4, and for text completeness from $-.05$ (ns) at T1 to $-.32$ ($p < .001$) at T4 (see Table 1).

Our second main question was about the specific relationships between the different categories of spelling errors and text production. Even if phonographic errors are still present in Grades 3 and 4 texts (as in Bahr et al., 2012), multilevel growth curve analysis showed that they had no impact on text completeness but had a negative effect on text length. Lexical errors only impacted the length of texts and this impact increased over time (Beers et al., 2017). As expected from the results reported in surveys on French spelling (Andreu & Steinmetz, 2016; Fayol et al., 1999; Luci & Millet, 1994), morphological errors had a negative effect only on the completeness of the texts (Figure 2), and this negative effect increased over time.

The results of this correlational study are in compliance with several previous research in English, Portuguese, Italian, and so on, and they confirm that spelling in French correlates negatively with text length and text completeness. Causal relationships cannot be directly deduced from results obtained from correlational data as reported here. However, causal inferences are supported by other studies providing evidence of the possible indirect negative effect of the load of lexical and morphosyntactic spelling in French, thus making it highly probable a negative impact of spelling difficulties onto written text production in children (Girolami-Boulinier, 1984; Maggio et al., 2018; Van Reybroeck & Hupet, 2009). This is exactly what we found. From the end of Grade 3, when handwriting was relatively automated and when orthographic knowledge could refer to memorized orthographic forms, individual differences in lexical spelling impacted text length and individual differences in morphosyntactic spelling affected text completeness. Poor lexical spellers produced shorter texts; poor morphological spellers composed less complete texts.

An important question concerns the absence of relations between either length or completeness of the contents of the texts and the proportions of errors in T1, and even T2. It is during this period that errors are most frequent and one would expect this to impact the length and completeness of the text, which is not the case here. However, this result could be interpreted in line with the predictions of the model of Berninger and Swanson (1994) and Berninger (1999): among the most novices (and therefore often the youngest), the processes involved in production – here the management of text content and the management of spelling treatments – are not yet integrated in working memory and are treated independently of each other. Such a result had already been reported in previous research on real-time production (Chanquoy et al., 1990). In contrast to adults who composed texts by distributing their attention and activities at different points in the production (i.e. pre-writing pause, between proposition pauses, and even modifying their flow of writing), 8-year-olds showed variations only during the pre-writing phase. In contrast, Tenth grade students exhibited almost the same modulations of writing as adults (see also Olive, 2004). Everything happens as if for the most novices, the processes of content management and those of management of translation (Fayol, 2016) and transcription (Alves et al., 2018) were only gradually integrated. Research analyzing production times and flows would be necessary to highlight this progressive integration (Maggio et al., 2012).

Progressing in writing necessitates the increasing fluency of linguistic processes, transcription skills and text generation skills involved in text production (Alamargot & Fayol, 2009; Berninger et al., 2002; Kim, 2019; McCutchen, 1996, 2000, 2011). Given that young writers have a limited amount of resources available to compose texts, the greater the amount of attention and effort paid to spelling, the fewer resources remain available for other lower-order processes or for higher-order processes. Depending on the intensity of the effort required, this could lead pupils to lose ideas or give up expressing them (e.g., composing less complete texts) or not being able to devote enough attention to establishing consistency between the events related. In addition, when the writers become aware of the difficulties, they can delay one of them and focus on the other(s), strategically allocating more or less time and effort to managing their own production (Levy & Ransdell, 1995). For

example, they can decide to substitute more frequent and simple words – but often less precise – or syntactic forms to rarer, inconsistent, and complex items.

Our results show that the inter-individual differences linked to lexical spelling difficulties (evidenced by a higher number of lexical errors) are associated with shorter texts. Several studies reported that dealing with lexical spelling difficulties takes time and requires effort to retrieve rare items in memory or to compose unfamiliar or new orthographic forms (i.e., through assembled spelling) and sometimes to decide between different spelling options (e.g., “enfant” versus “enfant” [child]) (Bonin et al., 1997; Bonin et al., 2018; Delattre et al., 2006). The reasons why lexical spelling difficulties are associated with the size of texts remains to be explored. Pupils could fail to find the precise spelling for specific words, and refrain from writing long texts. Later, they could have stored more lexical forms and become able to adaptively choose only the words they are able to spell correctly.

The relationship between morphosyntactic errors and text completeness is easier to explore. Several experimental studies using double tasks have provided evidence that agreement processes (e.g., subject-verb or adjective-noun) are difficult to learn and remain fragile even in adults. Fayol, Largy and Lemaire (1994) and Largy, Fayol and Lemaire (1996) have induced agreement errors on verbs in highly educated adults having to write down orally presented sentences. Due to the silent characteristics of written French morphology, adults inflected erroneously verbs such as “arrive” when they had to both temporarily store in Working Memory a series of five words while they were writing a sentence like “Le chien des voisins arrive” [The dog of the neighbours arrives]: they wrote “arrivent” (plural) instead of “arrive” (singular), making the verb agree with the closest (local) noun instead of making the correct subject-verb agreement. The same phenomenon occurred in children production: French children were able to use the relevant noun and verb inflections when explicitly asked to do so. However, they “forgot” to add these inflections when they had to recall (by writing down) the same linguistic patterns or to produce them when they composed a text (Fayol et al., 2006; Totereau et al., 1997). Disruptions could occur both in a bottom-up and in a top-down way: In the first case, the resources devoted to higher-level processes (e.g., dealing with coherency or introducing new ideas) prevent the correct realization of lower-level processes (e.g., agreement); in the second case, the cost of lower-

level processes (i.e., detecting agreement problems or revising agreement) hampers the management of higher-level processes. Indeed, the management of agreements necessitates attention shifts and eye-movements shifts from place to place (Alamargot et al., 2015). These changes in the focus of attention can make the pupils lose ideas, fail to generate ideas or fail to follow the thread or organization of their ideas. A decrease of text completeness could result from such attentional interferences.

Limitations of the present study and future directions

In terms of theoretical implications, results are coherent with the capacity theory of Just and Carpenter (1992). Due to the limitations on processing capacities, Third and Fourth graders can only manage a limited number of activities in parallel so as not to go beyond the available resources tied to the task (McCutchen, 2000; Maggio et al., 2018). They must keep in mind the goals of the production, the main contents related to the topic, and in parallel activate the words and their orthographic forms and process the syntactic structures and the agreement constraints they are linked to. In a number of situations, they cannot manage the multiple demands required by the writing task. Most probably, in a number of cases, they favor ideas and their organization at the expense of spelling. Hence the occurrence of many errors. In other circumstances, maybe related to the instructions and requirements of the task, their attention is focused on spelling to the detriment of the recovery and coherence of ideas. With the consequences of less completeness and less consistency of information. Additional research on the real-time process of production would be necessary to study these back-and-forth between management of high levels of production and processing of orthographic processes (see Alamargot et al., 2015; Maggio et al., 2012).

Although this study has explored the concurrent and longitudinal relationships between spelling and written text composition in an irregular orthographic system such as French and has shown (1) a growing relationship between text production and orthographic performance, (2) that this relationship also varied according to whether it related to text length or to text completeness, and (3) its strength and evolution depended on the type of errors (phonographic, lexical, or morphological), a number of questions remain open. As our study is correlational, causal relationships cannot be directly deduced from our results. An explicit experimental manipulation of cognitive load is thus needed to determine whether

limiting or increasing resources play a prominent role in the relationships between spelling and written text composition in French (e.g., Fayol, 1999; McCutchen, 2000; Olive, 2004). Also, since only one rater rated the samples, we could not calculate any reliability coefficient. Although a precise coding grid was established, which left few room for interpretation, one cannot assure that the coding was perfectly stable across samples.

Another limitation is that we focused our investigation on the possible effects of French spelling difficulties on the length and completeness of the texts produced. Our reasoning was that the cognitive cost (the load) of lexical processing and morphosyntactic agreements was likely to reduce the cognitive capacities that could be mobilized on the one hand, to find in memory or elaborate little-known spelling forms and on the other hand, to manage written marks without phonological traces. This mobilization of resources would be at the expense of the two dimensions of production: the quantity of information reported (the length) and the completeness thereof. However, data collection conditions prevented us from taking into account another source of transcription difficulty: the cost of handwriting. By studying the text productions of Third and Fourth grade students, we reduced the impact of this variable, as previous work attests (Bourdin & Fayol, 1994; Graham et al., 1997), but we were not able to completely eliminate it: Limpo et al. (2017) noted a transcription effect still in Seventh grade, and Bourdin and Fayol (2002) still in adults. It is therefore likely that the cognitive cost of handwriting also contributes, even if it is lower than for First and Second graders, to reducing the cognitive resources available to both manage the content of the texts and to process the spelling difficulties. Additional research will therefore be necessary to assess or control for the contribution of this variable to the trends that we have highlighted.

In terms of practical implications, the present findings are in line with previous investigations showing that children who have difficulties in text composition endure difficulties in text composition over time. Part of these difficulties seem to be associated to spelling problems. Until now, most studies have provided evidence that spelling errors occurred due to the heavy load of coordinating all the components involved in text composition. Our results suggest that things could also be going the other way: spelling difficulties could have a negative impact on the fluency and quality (i.e., here completeness)

of text production. Improving the production of texts in elementary schools would therefore require that spelling skills should be considered within text production interventions in both directions (Berninger et al., 2002).

If, as we have shown, spelling knowledge and its management by those who write pose a problem, it seems necessary to teach spelling, as had been proposed by Graham and Harris (1997). This is already what is done in France. However, this teaching must take two forms. On the one hand, as in all orthographic systems requiring the constitution in memory of an orthographic lexicon, certain lexical and sub-lexical orthographic forms must be taught (Pacton et al., 2013, 2014). Learning them should enable text authors to retrieve the most frequent forms directly from memory, and to hesitate less when producing inconsistent forms. On the other hand, morphosyntactic processing must also give rise to teaching – the silent marks of chords – and to training in their management. The data from the national assessments (Andreu & Steinmetz, 2016) show that this is the major difficulty in the production of texts in children. Our research highlights that the management of morphosyntactic processing has an impact on the processing of contents in memory: these are less easily recovered. It is therefore necessary to look for activities likely to facilitate the real-time management of these treatments (see Chanquoy et al., 1990, and Maggio et al., 2012, for insights). This should be the subject of further investigations. The latter should make it possible to understand precisely what the difficulties are and to determine the methods of intervention.

References

- Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology, 102*(2), 281–298. <https://doi:10.1037/a0019318>.
- Alamargot, D. & Fayol, M. (2009). Modelling the development of written composition. In R. Beard, D. Myhill, M. Nystrand, & J. Riley (Eds), *Handbook of Writing Development* (pp. 23-47). London: Sage.
- Alamargot, Flouret, L., D., Larocque, D., Caporossi, G., Pontart, V., Paduraru, C., Morrisset, P., & Fayol, M. (2015). Successful written subject verb agreement: an online analysis of the procedure used by students in Grades 3, 5 and 12. *Reading and Writing, 28*, 291-312. DOI: 10.1007/s11145-014-9525-0.
- Alves, R. A., Limpo, T., Fidalgo, R., Lenia, C., Alvares, P. L., & Castro, S. L. (2015). The impact of promoting transcription on early text production: Effects on bursts and pauses, levels of written language, and writing performance. *Journal of Educational Psychology, 108*(5), 665–679. <https://doi.org/10.1037/edu0000089>
- Alves, R. A., Limpo, T., Salas, N., & Joshi, R. M. (2018). Handwriting and spelling. *Best Practices in Writing Instruction, 3*, 211–239.
- Andreu, S., & Steinmetz, C. (2016). Les performances en orthographe des élèves en fin d'école primaire (1987, 2007, 2015). *Note d'Information de la DEPP, n°28*.
- Arfé, B., Dockrell, J., & Di Bernardi, B. (2016). The effect of language specific factors on early written composition: The role of spelling, oral language and text generation skills in a shallow orthography. *Reading and Writing, 29*, 501–527. <https://doi.org/10.1007/S11145-015-9617-5>
- Bahr, R.H., Silliman, E.R., Berninger, V.W., & Dow, M. (2012). Linguistic pattern analysis of misspellings of typically developing writers in grades 1-9. *Journal of Speech, Language, and Hearing Research, 55*, 1587-1599. [https://doi.org/10.1044/1092-4388\(2012/10-0335](https://doi.org/10.1044/1092-4388(2012/10-0335)
- Beers, S. F., Mickail, T., Abbott, R., & Berninger, V. (2017). Effects of transcription ability and transcription mode on translation: Evidence from written compositions, language bursts and pauses when students in grades 4 to 9, with and without persisting dyslexia or dysgraphia, compose by pen or by keyboard. *Journal of writing research, 9*(1), 1–25. <https://doi.org/10.17239/jowr-2017.09.01.01>
- Berninger, V. W., & Hart, T. M. (1992). A developmental neuropsychological perspective for reading and writing acquisition. *Educational Psychologist, 27*(4), 415–434. https://doi.org/10.1207/s15326985ep2704_2
- Berninger, V.W., & Swanson, H.L. (1994). Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writing. In E. Butterfield (Ed.), *Children's*

writing : Toward a process theory of the development of skilled writing (pp. 57-81).
Greenwich, CT: J.A.I. Press.

- Berninger, V. W., Vaughan, K., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., et al. (2002). Teaching spelling and composition alone and together: Implications for the simple view of writing. *Journal of Educational Psychology*, 94(2), 291.
- Bonin, P., Collay, S., & Fayol, M. (2008). La consistance orthographique en production écrite: une brève synthèse. *L'Année Psychologique*, 108, 517-546.
- Bonin, P., & Méot, A. (2019). Task differences and individual differences in skilled spelling. In C. Perret & T. Olive (Eds.), *Spelling and writing words: Theoretical and methodological advances* (pp. 131–150). Leiden: Brill.
- Bosse, M-L., Brissaud, C., & Le Levier, H. (2020). French pupils' lexical and grammatical spelling from sixth to ninth grade: A longitudinal study. *Language and Speech*, 64, 224-229.
<https://doi.org/10.1177/0023830920935558>
- Bourdin, B., & Fayol, M. (1994). Is written language production really more difficult than oral language production? *International Journal of Psychology*, 29, 591-620.
- Bourdin, B., & Fayol, M. (2002). Even in adults, written production is still more costly than oral production. *International Journal of Psychology*, 37, 219-227.
- Brissaud, C., & Chevrot, J.-P. (2011). The late acquisition of a major difficulty of French inflectional orthography: The homophonic /E/ verbal endings. *Writing Systems Research*, 3(2), 129–144. <https://doi.org/10.1093/wsr/wsr003>
- Broc, L., Bernicot, J., Olive, T., Favart, M., Reilly, J., Quémart, P., & Uzé, J. (2013). Lexical spelling in children and adolescent with specific lanbguage impairment: Variations with the writing situation. *Research in Developmental Disabilities*, 34, 3253-3266.
- Casalis, S., Pacton, S., Lefevre, & Fayol, M. (2018). Morphological training in spelling: Immediate and long-term effects of an interventional study in French third graders. *Learning and Instruction*, 53, 89–98. <https://doi.org/10.1016/j.learninstruc.2017.07.009>.
- Catach, N., Gruaz, C., & Duprez, D. (1980). *L'orthographe française : Traité théorique et pratique avec des travaux d'application et leurs corrigés*. Nathan.
- Chanquoy, L., Foulain, J.N., & Fayol, M. (1990). The temporal management of short text writing by children and adults. *European Bulletin of Cognitive Psychology*, 10, 513-540.
- Connelly, V., & Dockrell, J. (2016). Writing development and instruction for students with learning disabilities: Using diagnostic categories to study writing difficulties. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 349–363). The Guilford Press.
- Daffern, T., Mackensie, N.M., & Hemmings, B. (2017). Predictors of writing success: How important are spelling, grammar and punctuation? *Australian Journal of Education*, 61(1), 75-87. <https://doi.org/10.1177/0004944116685319>

- Delattre, M., Bonin, P., & Barry, C. (2006). Written spelling to dictation: Sound-to-spelling regularity affects both writing latencies and durations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(6), 1330–1340. <https://doi.org/10.1037/0278-7393.32.6.1330>
- Dubois, J. (1965). *Grammaire structurale du Français: Nom et pronom*. Paris: Larousse.
- Fayol, M. (1997). *Des idées au texte (From ideas to text)*. Paris : Presses universitaires de France.
- Fayol, M. (1999). From on-line management problems to strategies in written composition. In M. Torrance & G. Jeffery (Ed.), *The cognitive demands of writing* (pp. 13-23). Amsterdam: Amsterdam University Press.
- Fayol, M. (2014). Written language: Learning to read and to spell. In C. Fäcke (Ed.), *Manual of language acquisition* (pp. 162-176). Berlin: Walter de Gruyter.
- Fayol, M. (2018). Écrire (Writing). In O. Houdé et G. Borst (Eds.), *Le cerveau et les apprentissages* (pp. 69–87). Paris : Nathan.
- Fayol, M., Hupet, M. & Largy, P. (1999). The acquisition of subject-verb agreement in written French. From novices to expert errors. *Reading and Writing*, 11, 153-174. <https://doi.org/10.1023/A:1008038127807>
- Fayol, M., Alamargot, D., & Berninger, V. (2012). *Translation of thoughts to written text while composing: Advancing theory, knowledge, methods, and application*. New York: Psychology Press.
- Fayol, M., Largy, P. & Lemaire, P. (1994). Subject-verb agreement errors in French. *Quarterly Journal of Experimental Psychology*, 47A, 437-464. <https://doi.org/10.1080/14640749408401119>
- Fayol, M., Totereau, C. & Barrouillet, P. (2006). Disentangling the impact of semantic and formal factors in the acquisition of number inflections. Noun, adjective and verb agreement in written French. *Reading and Writing*, 19, 717-736. <https://doi.org/10.1007/s11145-005-1371-7>
- Fitzgerald, J. (1984). The relationship between reading ability and expectations for story structures. *Discourse Processes*, 7(1), 21–41. <https://doi.org/10.1080/01638538409544579>.
- Foulin, J.N., & Chanquoy, L. (2006). Online study of word spelling production in children. In L. Van Waes, M. Leijten, & C. Neuwirth (Eds.), *Writing and Digital Media* (pp. 145-152). Oxford, UK: Elsevier.
- Girolami-Boulinier, A. (1984). *Les niveaux actuels dans la pratique du langage oral et écrit*. Paris: Masson
- 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), 170–182. <https://doi.org/10.1037/0022-0663.89.1.170>.

- Graham, S., & Harris, K. R. (1997). It can be taught, but it does not develop naturally: myths and realities in writing instruction. *School Psychology Review*, 26, 414–424. <https://doi.org/10.1080/02796015.1997.12085875>
- Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology*, 30(2), 207–241. <https://doi.org/10.1016/j.cedpsych.2004.08.001>
- Hayes, J. (1996). A new framework for understanding cognition and affect in writing. In C. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 1–27). Mahwah: Lawrence Erlbaum.
- Hayes, J. R., & Berninger, V. W. (2014). Cognitive processes in writing: A framework. In B. Arfé, J. Dockrell, & V. Berninger (Eds.), *Writing development in children with hearing loss, dyslexia, or oral language problems: Implications for assessment and instruction* (pp. 3–15). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199827282.003.0001>
- Hayes, J. R., & Flower, L. (1980). Identifying the organization of writing processes. In L. W. Gregg, & E. R. Steinberg (Eds.), *Cognitive processes in writing: An interdisciplinary approach* (pp. 3–30). Hillsdale, NJ: Lawrence Erlbaum.
- Jaffré, J-P, & Fayol, M. (2005). Orthography and literacy in French. In R.M. Joshi and P.G. Aaron (Eds.), *Handbook of orthography and literacy*. Mahwah : L.E.A.
- Joshi, M., & Aaron, P. G. (2006). *Handbook of Orthography and Literacy*. New York: Routledge.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first to fourth grade. *Journal of Educational Psychology*, 80(4), 437-447. <https://doi.org/10.1037/0022-0663.80.4.437>
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149. DOI: 0033-295X/92/J3.00
- Kellogg, R. T. (2001). Competition for working memory among writing processes. *The American Journal of Psychology*, 114(2), 175–191. <https://doi.org/10.2307/1423513>
- Kent, S., Wanzek, J., Petscher, Y., Al Otaiba, S., & Kim, Y. S. (2014). Writing fluency and quality in kindergarten and first grade: the role of attention, reading, transcription, and oral language. *Reading and Writing*, 27(7), 1163–1188. <https://doi:10.1007/s11145-013-9480-1>.
- Khomsî, A., Nanty, I, Parbeau-Gu  no, A., & Pasquet, F. (2005). *ECL –   valuation des comp  tences linguistiques   crites au coll  ge*. Paris: Pearson ECPA.
- Kim, Y-S, G. (2019). Structural relations of language and cognitive skills, and topic knowledge to written composition: A test of the direct and indirect effects model of writing. *British Journal of Educational Psychology*, 90, 910-932
- 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. <https://doi:10.1016/j.lindif.2011.06.004>.

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), 79–95. <https://doi:10.1037/a0037210>.

Kim, Y-S., G., & Graham, S. (2022). Expanding the direct and indirect effects model of writing (DIEW): Reading-writing relations, and dynamic relations as a function of measurement/dimensions of written composition. *Journal of Educational Psychology*, <https://doi.org/10.1037/edu0000564>

Largy, P., & Fayol, M. (2001). Oral cues improve subject-verb agreement in written French. *International Journal of Psychology*, 36, 121-132. <https://doi.org/10.1080/00207590143000009>

Largy, P., Fayol, M., & Lemaire, P. (1996) The homophone effect in written French: The case of verb-noun inflection errors, *Language and Cognitive Processes*, 11(3), 217-256. DOI: 10.1080/016909696387178

Lété, B., Peereman, R., & Fayol, M. (1998). Consistency and word-frequency effects on spelling among first- to fifth-grade French children: A regression-based study. *Journal of Memory and Language*, 8, 952–977. <https://doi.org/10.1016/j.jml.2008.01.001>

Lété, B., Peereman, R., & Fayol, M. (2008). Phoneme-to-Grapheme Consistency and Word-Frequency Effects on Spelling Among First- to Fifth-Grade French Children: A Regression-Based Study. *Journal of Memory and Language*, 58, 952-977. doi:10.1016/j.jml.2008.01.001

Levy, C. M., & Ransdell, S. (1996). *The science of writing: Theories, Methods, Individual Differences and Applications*. Mahwah NJ: Lawrence Erlbaum Associates.

Limpo, T., & Alves, R.A. (2013). Modeling writing development: Contribution of transcription and self-regulation to Portuguese students' text generation quality. *Journal of Educational Psychology*, 105(2), 401-413.

Limpo, T., Alves, R. A., & Connelly, V. (2017). Examining the transcription-writing link: Effects of handwriting fluency and spelling accuracy on writing performance via planning and translating in middle grades. *Learning and Individual Differences*, 53, 26–36. <https://doi.org/10.1016/j.lindif.2016.11.004>

Lucci, V., & Millet, A. (1994). *L'orthographe de tous les jours*. Paris : Honoré Champion.

Maggio, S., Izaute, M., & Chenu, F. (2018). Spelling while writing texts. *L'Année Psychologique*, 118, 1-18. DOI 10.3917/anpsy1.181.0003

Maggio, S., Lété, B., Chenu, F., Jisa, H., & Fayol, M. (2012). Tracking the mind during writing: immediacy, delayed, and anticipatory effects on pauses and writing rate. *Reading and Writing*, 25, 2131–2151. <https://doi.org/10.1007/s11145-011-9348-1>

Martinet, C., Valdois, S., & Fayol, M. (2004). Lexical orthographic knowledge develops from the beginning of reading acquisition. *Cognition*, 91, B11-B22.

<https://doi.org/10.1016/j.cognition.2003.09.002>

McCabe, A., & Peterson, C. (1984). What makes a good story? *Journal of Psycholinguistic Research*, 13(6), 457–480. <https://doi.org/10.1007/BF01068179>

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review*, 8, 299–325. <https://doi.org/10.1007/BF01464076>

McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist*, 35(1), 13–23. https://doi.org/10.1207/S15326985EP3501_3

McCutchen, D. (2011). From Novice to Expert: Implications of Language Skills and Writing-Relevant Knowledge for Memory during the Development of Writing Skill. *Journal of Writing Research*, 3(1), 51-68. <http://dx.doi.org/10.17239/jowr-2011.03.01.3>

Menard, S. (2002). *Applied logistic regression analysis* (2nd ed.). Thousand Oaks: Sage.

Morin, M-F., Alamargot, D., Diallo, T.M.O, & Fayol, M. (2018). Individual differences in lexical and grammar spelling across primary school. *Learning and Individual Differences*, 62, 128-140. <https://doi.org/10.1016/j.lindif.2018.02.002>

Olive, T. (2004). Working memory in writing: Empirical evidences from the dual-task technique. *European Psychologist*, 9, 32–42.

Olive, T. (2014). Toward a parallel and cascading model of the writing system: A review of research on writing processes coordination. *Journal of Writing Research*, 6(2), 141-71. [doi: 10.17239/jowr-2014.06.02.4](https://doi.org/10.17239/jowr-2014.06.02.4)

Pacton, S., Borchardt, G., Treiman, R., L  t  , B., & Fayol, M. (2014). Learning to spell from reading: General knowledge about spelling patterns influences memory for specific words. *The Quarterly journal of Experimental Psychology*, 67 (6), 1019-1036. doi.org/10.1080/17470218.2013.846392

Pacton, S., Fayol, M., & Perruchet, P. (2005). Children’s implicit learning of graphotactic and morphological regularities in French. *Child Development*, 76, 324-339. https://doi.org/10.1111/j.1467-8624.2005.00848_a.x

Pacton, S., Foul  n, J.-N., Casalis, S., & Treiman, R. (2013a). Children benefit from morphological relatedness when they learn to spell new words. *Frontiers in Psychology*, 4, 696. <https://doi.org/10.3389/fpsyg.2013.00696>

Pacton, S., Sobaco, A., Fayol, M., & Treiman, R. (2013b). How does graphotactic knowledge influence children’s learning of new spelling? *Frontiers in Psychology*, 4, article 701. doi: 10.3389/fpsyg.2013.00701

Qu  mart, P., & Lambert, E. (2019). The influence of the morphological structure of words on the dynamics of handwriting in adults and fourth and sixth grade children. *Reading and Writing*, 32:175–195 <https://doi.org/10.1007/s11145-017-9762-0>

- Rey-Debove, J. (1984). Le domaine de la morphologie lexicale. *Cahiers de lexicologie*, 45, 3-19.
- Richards, T., Aylward, E., Raskind, W., Abbott, R., Field, K., Parsons, A., Richards, A., Nagy, W., Eckert, M., Leonard, C., & Berninger, V. (2006). Converging evidence for triple word form theory in children with dyslexia. *Developmental Neuropsychology*, 30, 547-589. https://doi.org/10.1207/s15326942dn3001_3
- Richards, T., Berninger, V., & Fayol, M. (2009). FMRI activation differences between 11- year-old good and poor spellers' access in working memory to temporary and long-term orthographic representations. *Journal of Neurolinguistics*, 22, 327-353. <https://doi.org/10.1016/j.jneuroling.2008.11.002>
- Seymour, P.H., Aro, M., & Erskine, J.M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143-174. <https://doi.org/10.1348/000712603321661859>
- Swanson, H.L., & Berninger, V.W. (1996). Individual Differences in Children's Working Memory and Writing Skills. *Journal of Experimental Child Psychology*, 63, 358-385. <https://doi.org/10.1006/jecp.1996.0054>
- Torrance, M., & Galbraith, D. (2016). The processing demands of writing. In C. A. MacArthur, S. Graham, J. Fitzgerald (Eds.), *Handbook of Writing Research, Second Edition* (pp. 67-80). New York: The Guilford Press.
- Totereau, C., Thévenin, M-T., & Fayol, M. (1997). The development of the number understanding of number morphology in written French. In C. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell* (97-114). Mahwah NJ: Lawrence Erlbaum Associates Publishers
- Trabasso, T., & Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes*, 15(3), 249–275. <https://doi.org/10.1080/01638539209544812>
- Van Reybroeck, M., & Hupet, M. (2009). Acquisition of number agreement: effects of processing demands. *Journal of Writing Research*, 1(2), 153-172.
- Veronis, J. (1986). Etude quantitative sur le système graphique et phono-graphique du français. *European Bulletin of Cognitive Psychology*, 6, 501-553. <http://www.jstor.org/stable/41557647>
- Weth, C. (2020). Distinguishing Syntactic Markers From Morphological Markers. A Cross-Linguistic Comparison. *Frontiers in Psychology*, 11, article 2082. doi: 10.3389/fpsyg.2020.02082
- Ziegler, J., Jacobs, A., & Stone, G.O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments & Computers*, 28, 504–515. <https://doi-org.insb.bib.cnrs.fr/10.3758/BF03200539>

Spelling and composing in French

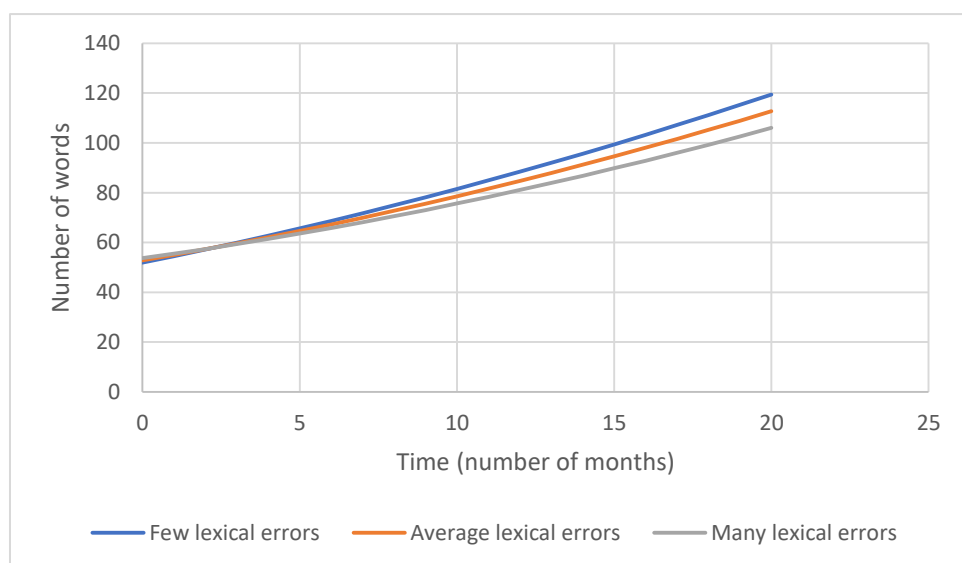


Figure 1. Text length as a function of time and lexical errors.

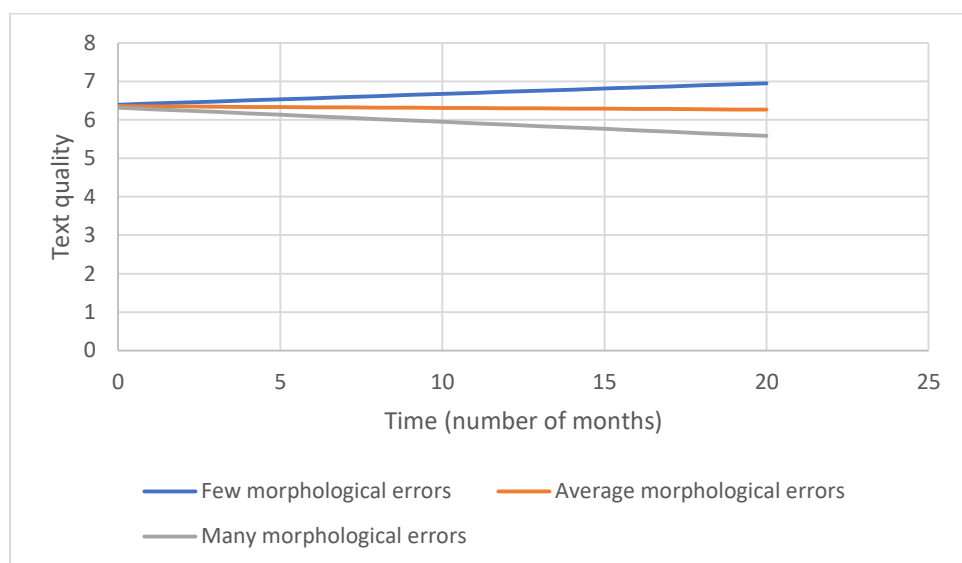


Figure 2. Text completeness as a function of time and morphological errors.

Supplementary material

Table S1. Descriptive Statistics for the Raw Variables Used to Compute Composite Variables Introduced in the Models

Time	Variable	N	Mean	SD	Minimum	Maximum
T1	Number of phonographic errors	173	1.03	1.95	0	15
	Number of lexical errors	173	5.60	4.01	0	19
	Number of morphosyntactic errors	173	5.91	3.78	0	18
	Factual completeness	173	4.90	1.32	2	9
	Inferential completeness	173	0.84	0.93	0	3
T2	Number of phonographic errors	173	0.82	1.70	0	11
	Number of lexical errors	173	4.47	3.63	0	16
	Number of morphosyntactic errors	173	5.62	4.36	0	20
	Factual completeness	173	4.95	1.18	2	8
	Inferential completeness	173	1.02	1.03	0	4
T3	Number of phonographic errors	173	0.67	1.59	0	9
	Number of lexical errors	173	7.15	5.58	0	40
	Number of morphosyntactic errors	173	9.46	6.39	0	39
	Factual completeness	173	5.25	1.83	1	10
	Inferential completeness	173	1.08	1.13	0	4
T4	Number of phonographic errors	173	0.80	1.98	0	15
	Number of lexical errors	173	6.23	4.99	0	29
	Number of morphosyntactic errors	173	10.02	6.09	1	28
	Factual completeness	173	5.84	1.97	1	9
	Inferential completeness	173	1.15	1.06	0	4