Bangor University

DOCTOR OF PHILOSOPHY

Early development of text writing in two contrasting orthographies: English and Spanish

Salas, Nayme

Award date:
2013

Awarding institution:
Bangor University

Department of Psychology
Bangor University

Early Development of Text Writing in Two Contrasting Orthographies: English and Spanish

Naymé Salas

Supervisor
Dr. Markéta Caravolas

Submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
Bangor University

May 2013

Summary

Current thinking about writing considers spelling skills to be a crucial factor in its development. However, most research evidence to date comes from studies in English, which has a highly inconsistent orthography. Since the rate of spelling development is affected by degrees of orthographic consistency, cross-linguistic investigations are thus needed to determine whether English-based assumptions may be extended to more consistent orthographies.

This work reports the results of a longitudinal investigation of early writing development from a cross-linguistic viewpoint. The writing development of children learning to write in English, a very inconsistent orthography, and children learning to write in Spanish, a very consistent orthography, was contrasted. Specifically, the studies included in this thesis explored the extent to which orthographic complexity moderated gains in a number of microstructural writing features, as well as whether it shaped the relationships between different levels of text construction. Finally, preschool-Year 2 predictors of writing outcomes were explored.

The results showed that orthographic complexity moderates gains in word-level features of text writing, but, beyond the word level, both language groups showed remarkably similar performances. Spelling was found to contribute to development in the amount of text produced (text length) over a relatively short period of time. Moreover, the underlying factors driving the development of writing were common until the middle or end of Year 1. In short, orthographic consistency seems to moderate word-level writing, but has a reduced effect in text-level performance. The discussion considers the implications for models of writing development and the divide between word- and text-level writing performance.

Preface

Chapters 2 through to 6 present data collected within the framework of the *Marie Curie* FP7-People Initial Training Network grant, entitled *Enhancing Literacy Development in European Languages* (henceforth, ELDEL; reference: FP7-PEOPLE-2007-1-1-ITN 215961, PI: Markéta Caravolas. For a description of the project see http://www.eldel.eu/workpackage1). Section 3.1 offers a more detailed account of the characteristics of the ELDEL project.

Work-package 1, *Establishing the Foundations of Literacy Development in European Languages,* was an integral part of ELDEL and the project to which the PhD candidate was appointed. Work-package 1 collected linguistic, cognitive, and other literacy-related performance data of the same sample of children who informed this thesis. The text writing task and the data obtained from it are an original, unique contribution of Naymé Salas, under the supervision of Dr. Markéta Caravolas. Therefore, all research questions and designs were conceived by the PhD candidate, under the supervision of Dr. Markéta Caravolas. Upon completion of the present body of work, no member of the ELDEL project had approached the development of text writing using the data reported here. The table below provides a list of all the different tasks and instruments that were the source of information used in the thesis, as well as it specifies the precise purpose with which they were used. Note that of all members of the ELDEL project, only the PhD candidate, Naymé Salas, pursued the study of text writing.

Chapter	Data/Task used	Description
2	Teacher questionnaires	Questionnaires for teachers of all participating classrooms were collected yearly. Part of the data reported in this chapter was extracted from these questionnaires to compare the amount of instruction in each language group, with special emphasis on writing instruction.
3	Parents' questionnaires	Parents' of all participants were invited to fill in a questionnaire. Only information concerning parents' education level was used as a proxy for children's SES in this chapter, as well as in ensuing chapters as a control variable.
	Text writing	This task was conceived of by the PhD candidate as the source of information for the main outcome measures of her thesis. It was collected simultaneously in the UK and in Spain. By the time this thesis was submitted no other member of the ELDEL project was working on this data or in the development of writing.
4	Text writing	Same as above
5	Word writing	This task was collected as a main outcome of the ELDEL project. Here it was used as a predictor measure for the main outcome of the thesis: text writing skills.
	Text writing	Same as above
6	Block design Vocabulary Word writing Letter writing Word reading Rapid Associative Naming Word span Letter knowledge Phoneme isolation Phoneme blending Morphological awareness Syntactic awareness Sentence repetition	This series of tasks were administered and collected for the ELDEL project to obtain estimates of cognitive skills associated with literacy development. In this chapter they were used to explore the cognitive predictors of writing skills.
	Text writing	Same as above

Para mamá

Acknowledgements

There have been many times when I was not entirely sure I would ever be writing these lines. I feel immensely grateful to many people who have been instrumental in this weird but fascinating journey that has been my PhD.

I wish to thank my supervisor, Dr. Markéta Caravolas, for granting me the amazing opportunity of being part of a truly unique international project. The experience was certainly an enriching one, in both the exciting as well as in the more challenging moments. I also want to thank the members of ELDEL. In particular, I am extremely grateful to Edu Onochie and Prof. Sylvia Defior for collecting the Spanish texts; to Sam Crewe and Andrea Reynolds, for their help in collecting the UK data; and to Anna Samara and Betty Mousikou, whose cheerful support was the greatest company during my time in Bangor. I am also indebted to Emma Campbell and Paola Rodríguez who coded dozens of texts and showed the best possible disposition without asking for anything in return.

A big part of the write up of this thesis was done at the University of Barcelona, in collaboration with the GRERLI group. I am very grateful for how welcome they all made me feel after 3 years of absence and for always making it look like I am at home when I go to work. I especially wish to thank Joan Perera, for being one of the most generous people I have ever met, and Anna Llauradó, also for generosity and constant support.

I cannot thank enough the help, love, and encouragement of my good friends Elisa and Melina, whom I adore and admire more and more every day. I feel very grateful to Mila and Rodrigo, for their love and support during the Bangor years. I also want to thank those who are always in my mind (and heart) despite how far we may be: my friends Martín and Rafa.

My life would be entirely different, and possibly quite miserable, without Lain. There are no words to express how grateful and lucky I feel for having you in my life.

Having been away from home for a good number of years, I feel most indebted to the love and support of my family: my brother, Gabi, and my sister, Salomé, and my parents, who have always trusted in me regardless of what I decided to do and wherever in the world I decided to go.

Most of all, I feel immensely grateful to Liliana Tolchinsky for being the best teacher (in the broadest sense) one could hope for. She has always been in my mind while I was writing almost every word of this work and has been a constant source of inspiration. She once opened a door that, even in the most difficult times, I have never considered closing.

Contents

List of figures

List of tables

Chapter 1

General introduction and review of the literature

1.1 Introduction to the Study

This is a thesis about language development. As such, it deals with a defining trait of human nature, which is the faculty to make use of a limited set of discrete, symbolic units, which are combined—following set rules—to form a potentially infinite number of expressions. Noam Chomsky (1957/2002) refers almost indistinctively to the spoken and written modalities when providing his very definition of *language*,

> [...] I will consider a *language* to be a set (infinite or finite) of sentences, each finite in length and constructed out of a finite set of elements. All natural languages **in their spoken or written** form are languages in this sense, since each natural language has a finite number of phonemes (or letters in its alphabet) and each sentence is representable as a finite sequence of these phonemes (or letters), though there are infinitely many sentences. (p. 13, my emphasis in bold)

The studies within this thesis focus specifically on one aspect of language development, which is the process of learning to write. In this sense, I will consider writing development within the framework of *later language development* (LLD). LLD refers to the linguistic growth that takes place once children have acquired the core grammar of their mother tongue(s), a process that is assumed to occur at around age 5 (Berman, 1997, 2004; Karmiloff-Smith, 1986). Gains in linguistic competence and performance after the age of 5 are thought to involve deepening knowledge about already acquired structures, particularly in accordance with specific communicative purposes and situations. Therefore, it does not entail acquiring new grammatical structures; rather, it promotes the reconfiguration of extant linguistic knowledge (Berman & Slobin, 1994; Karmiloff-Smith, 1992). Literacy thus represents a central aspect of LLD, provided that several of such later developments are rooted in and fostered by literacy practices (Olson, 1996; Tolchinsky, 2004). In sum, from the wider lens of LLD, writing is an essential element in the long-term process of becoming a proficient language user.

Not only is writing (and, more broadly speaking, literacy) an integral part of language development; also, oral and written language interact throughout the life-span. Given the phylogenetic and ontogenetic precedence of speech over writing, there is little controversy in claiming that the emergence and development of writing cannot be dissociated from oral language in any clear-cut way. As a consequence, the characteristics of spoken language will necessarily impact on the individual's conceptualization of how writing represents language. In addition, it has also been claimed that, in literate societies, written language shapes and re-shapes people's linguistic and metalinguistic representations (Ferreiro, 2009; Olson, 1994; Vernon & Ferreiro, 1999), to a great extent determining the course of language development. For example, the working definition of many basic linguistic concepts (e.g., word, sentence, etc.) derives from written language (Ferreiro, 2000; 2009).

There is general consensus in that mastery of written language is essential in today's literate societies, where it is present in virtually any kind of social interaction. As a consequence, proficient reading and writing skills are crucial to the full integration of an individual in everyday social practices, as they constitute the basis for personal, professional, and intellectual fulfilment. However, writing is a complex process that requires the integration of linguistic, cognitive, and procedural knowledge. Consequently, its development is extremely time-consuming (Alamargot & Chanquoy, 2001).

Understanding writing development is interesting from, at least, three main perspectives. First, from a linguistic perspective, writing is crucial for native speakers to become proficient language users (Berman & Ravid, 1999). Also, the written product exposes children's knowledge of all levels of linguistic analysis: phonological, morphosyntactic, lexical, pragmatic, and discoursive (Alamargot & Chanquoy, 2001, p. 3). Second, from the viewpoint of cognitive psychology, the complexity of the writing process and the variety of cognitive processes involved may serve to better comprehend how text composition is resolved at different developmental stages, as well as to determine in which ways immature writers' solutions resemble those of adult, skilled writers. Third, from a socio-educational perspective, becoming a skilled reader-writer is the primary goal of basic education and paves the way to most professional and intellectual achievements. Thus, a detailed description of the early steps in learning to write, as well as its linguistic and cognitive underpinnings, seems to be crucial to produce effective teaching practices, early identification of at-risk factors, and successful remediation strategies.

Although cross-linguistic comparisons have contributed greatly to our current understanding of the language faculty, whether to emphasize commonalities or language-specific traits, this type of approach is still far from constituting an ordinary practice in relation to literacy studies. Although some important large-scale studies have been made in the sub-area of *word* reading and spelling research (e.g., Caravolas, et al., 2012; Caravolas, Hulme, & Snowling, 2001; Caravolas, Volín, & Hulme, 2005; Georgiou, Parrila, & Papadopoulos, 2008; Seymour, Aro, & Erksine, 2003; Vaessen et al., 2010; Ziegler et al., 2010), cross-linguistic studies of *writing* development are scarce. A notable exception is the work that resulted from a cross-linguistic collaboration led by Prof. Ruth Berman, *Developing Literacy in Different Languages and Contexts* (1998-2002), which investigated the effects of variations in discourse genre (expository-narrative), production modality (spoken-written), and schooling experience (grade-school, junior-high, high-school, and college students). The project comprised data from seven different countries and covered an almost 10-year time span (from 9-years-old until adulthood). To the best of my knowledge, no large-scale study has been carried out looking at the earlier phases of learning to construct written text in the first years of primary education, from a cross-linguistic perspective. This is surprising, especially considering that (as it shall become apparent in subsequent sections) there are important cross-linguistic differences in the rate at which children master a fundamental aspect of writing: spelling.

The current body of work is one of the outcomes of one of such rare, large-scale, cross-linguistic endeavour (*Enhancing Literacy Development in European Languages,* ELDEL, FP7-PEOPLE-2007-1-1-ITN 215961, PI: Markéta Caravolas), aiming to clarify the shared and unique characteristics in the process of writing development in alphabetic orthographies. Within this framework, the studies contained in this thesis were designed to address the specific need in the field for longitudinal, cross-linguistic data on early writing development. With that goal, the early (mid-Year 1 to mid-Year 2) text composition skills of a group of monolingual English-speaking children were systematically compared to a similar group of Spanish-speaking peers. In addition, precursors of writing ability were obtained from both groups during Reception Year. In what follows, I review the main theoretical approaches and models of writing, both for skilled and novice writers; then, I make a critical analysis of the role of transcription factors in early writing development, which constitute one of the main sources of motivation for a cross-linguistic approach to writing development.

1.2 Models of Writing

1.2.1 Models of Skilled Writing

Models of adult or skilled writing have constituted the point of departure of virtually all developmental accounts of writing. Above all, they are concerned with the higher-level processes that result in efficient written communication. Consequently, they are mostly based on what expert or highly skilled writers do. Hayes and Flower's (Hayes & Flower, 1980) is the most relevant and influential of all models proposed to date (but see Nystrand, 2006, for a short history of writing research, showing notable influences that led to Hayes and Flower's proposal). In spite of having been conceived to explain the nature of the writing processes as they take place in "competent writers" (Hayes & Flower, 1980, p. 29), the model has had a remarkable impact on writing development research. Since it has constituted the reference point for much of the subsequent modelling of writing both experienced and novice, a review of the original 1980 model, as well as the most important revisions to it—in 1996 (Hayes, 1996), 2001 (Chenoweth & Hayes, 2001), and, more recently, in 2012 (Hayes, 2012)—are summarised below.

1.2.1.1 Hayes and Flower (1980). Hayes and Flower's (1980) seminal model of writing attempted to account for the cognitive processes that take place during adult, skilled writing. Their approach was a strong reaction to previous models of writing, which conceived the writing process in a linear fashion, usually incorporating stages both before and after the actual text production (e.g., Britton, Burgess, Martin, McLeod, & Rosen, 1975; Graves, 1983; McCormick-Calkins, 1986). Although the latter approach drew attention to the importance of processes beyond the text—such as planning or editing— and was very influential for the pedagogy of writing, it was inadequate to explain the interplay of cognitive processes within the writer (Flower & Hayes, 1981). In Hayes and Flower's (1980) model, writing was conceptualized as a complex activity that involved several cognitive processes and types of knowledge (of language, orthographic conventions, pragmatic and audience-related aspects, topic content, etc.). In other words, it was regarded as a problem-solving space that transcended the pedagogical and teaching realm, thus becoming a core interest of cognitive psychology (Alamargot & Chanquoy, 2001).

In accordance with the relevance assigned to the cognitive mechanisms involved in writing, the main source of information for the model was not constituted by texts, but by think-aloud protocols collected from a group of adult writers. The model (Figure 1.1) identified three main domains, two of which were subject–internal: the writing processes and the writer's long-term memory; and a third domain, external to the subject: the task environment. The latter involved the topic of the text, together with all the constraints and demands of the communicative situation, such as the ones imposed by the writing prompt and the characteristics of the target audience. Long-term memory stored the writer's knowledge about the topic, and also ready-made writing plans, such as the writing plan for an argumentative essay as opposed to the writing plan for a story narrative.

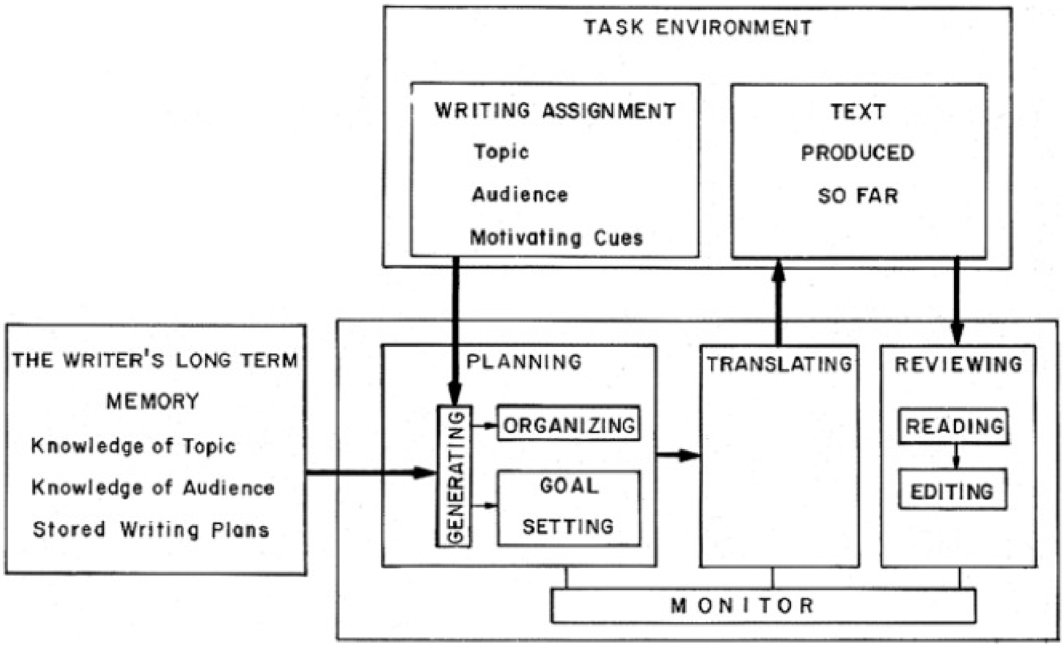


Figure 1.1. Hayes and Flower (1980) model of writing.

The heart of the model, which paved the way for the bulk of subsequent research on the topic, was constituted of three writing processes: *planning, translating*, and *reviewing*, together with the *monitor*, a cognitive device overseeing the transitions among the other three. Two main features stand out in the conception of the model: firstly, that the writing processes are hierarchical in nature, with each one including one or more component processes—with the crucial exception of the translation process—; secondly, that there is no order or linearity in the way processes take place. On the contrary, it is assumed that writers may transfer from one process to the other many times during composing. The lack of linearity in the occurrence of writing processes or of their sub-components, together with their hierarchical nature, yields a recursive model of writing (Flower & Hayes, 1981; Hayes & Flower, 1980, 1986).

The process of planning involved three sub-processes: (1) the generation of content, a sub-process labelled *generating*; (2) the structure and organization of such ideas, or *organizing*; and (3) the writer's goals regarding the tone and other audience- and rhetoric-related aspects of the text, referred to as the *goal setting* sub-process. Translating, on the other hand, was the process by which ideas generated during planning are encoded into written language. At this instance, all the constraints of the written language system have to be considered and attended to by the writer, a fact that Flower and Hayes (1981) acknowledged as potentially having a detrimental impact on the other writing processes, especially for developing or struggling writers, because of the likely overload to short-term memory resources (p. 373). As it shall become evident later on, this component process of writing and its cascading effect on the writing process and product will be the centre of attention of developmental accounts of writing.

The process of reviewing was constituted by two sub-components: *reading* and *editing*, which could be either planned or spontaneous. The writer may choose to revise any aspect or level of writing, thus determining which of the other processes or sub-processes needs to be evoked. For example, if during revision a writer decides to correct a spelling mistake or the legibility of a certain string of letters (assuming the text is handwritten), he will need to go back to the translating process. Similarly, if during one such episode of revision—conscious or otherwise—the writer considers that something in the text does not suit the communicative goals set during the planning process, most likely he will try to generate new ideas and eventually translate them into written language to find the wording that corresponds best to his communicative objectives. Finally, the *monitor* was the cognitive device that regulated the transition from one process or sub-process to another by determining which one was necessary at each point during the composing task. In other words, it accounted for personal styles of writing (Hayes, 2012).

In sum, Hayes & Flower's (1980) model should be seen as a turning point in the studies of writing because it signals a landmark in the history of writing research within cognitive psychology. The main concern was the nature of the relationship between an individual's mental processes during the construction of a written text and the final characteristics and quality of a text. The different units, processes and sub-processes, together with the recursivity inherent to the model, reflected the complexity of a most challenging cluster of cognitive skills.

1.2.1.2 Hayes (1996). Hayes' revision of the 1980's model preserved the essence of the original one, but put a stronger emphasis on affective and motivational aspects, specified the role of the writer's memory resources—especially working memory, which had been practically overlooked in the 1980 model—, and showed an increasing concern for reflecting the impact of the different writing media (e.g., handwriting vs. typing).

The author's interest in the role of motivation and affect was partly related to findings suggesting that students who believed that writing effectively was a matter of aptitude ("a gift", p. 9) showed more anxiety towards writing and lower images of themselves as writers. In a complex task where the writer not only needs to be able to cope with multiple processes but also with multiple goals, his motivation and attitude towards writing was seen as vital. The role of motivational factors in writing and, particularly, in writing development will be further elaborated by Hayes in later proposals (Hayes, 2011, 2012).

The number of domains in the model was reduced from three to two: an individual-internal domain, consisting of the writer's long-term and working memory capacities, and the writing processes; and an individual-external domain, the task environment. The components of the task environment were slightly more specified than in the original model. It was defined as including a social component (e.g., audience characteristics, other relevant texts, the writing prompt, etc.) and a physical component, to account for both the text produced so far and the writing medium. Subsuming the writer's (working and long-term) memory within the individual's domain achieved a two-fold goal: on the one hand, it provided a better description of the role of different types of memory resources during text composing, in the light of research evidence that had been produced since the 1980 paper (e.g., Kellogg, 1996; McCutchen, 1986); on the other hand, it served to make explicit the tight relationship between every step of the writing process and the writer's memory resources. The role of long-term memory had already been established in relation to topic knowledge, storage of writing schema, and so on (Hayes & Flower, 1980). In Hayes' (1996) model, with Baddeley's (1986) model of working memory as reference, it was further assumed that all writing processes, "have access to working memory and carry out all non-automated activities in working memory." (p. 8).

The more general cognitive processes of *reflection, text interpretation,* and *text production* were proposed instead of the writing processes of planning, translating, and reviewing, which underwent major reformulations. Planning was reconceptualised as a component, among others, of the more general "superprocess" of *reflection.* Reflection involved problem-solving strategies, of which planning was a subtype. The author reasoned that, when writers do not have an available task schema to complete a writing task, they often need to resort to general problem-solving and decision-making abilities (p. 21; for a more detailed account of planning processes, see Hayes & Nash, 1996). The role of planning within the model called for it to be downgraded from its prominent position in the 1980 model, especially after evidence that time spent in the entire writing task—and not solely in planning—tended to be associated with higher text quality. Research on the cognitive processes related to planning or reflection established that it was planning strategies, rather than time dedicated to it, that may be used as a marker for writing quality (e.g., Hayes & Nash, 1996). In addition, skilled writers seemed to engage in *conceptual planning*, which involves a concern for getting through to a certain audience, or expressing a specific tone or attitude in their messages. This type of planning was found to be rare, with most writers understanding planning solely in terms of *content generation*; that is, the ideas or propositions relevant to the writing task.

Reviewing also underwent a substantial reformulation and was replaced by reading. Hayes, Flower, Schriver, and Carey (1987) had focused on reading as an activity central to writing, partly due to evidence suggesting, for example, that text quality ratings appear to be related with performance on reading tests (Spivey, 1984; as cited in Hayes, 1996). Hayes (1996) brought that perspective into his reformulation of the model of writing. Reading was thus regarded as a subtype of text interpretation processes, where text interpretation stood for "a function that creates internal representations from linguistic and graphic inputs" (p. 13). In this way, different reading modalities contributed to writing in different ways. For example, reading-to-revise was crucial to detect textual problems (local/superficial or affecting the general organization of the text). In contrast, reading-to-evaluate allows one to form "a representation of the text's meaning" (p. 14), attending to the message that the writer intends to convey, rather than to text issues.

Translating was replaced by *text production,* and a more detailed account of the nature of this process was provided. Hayes (1996) based this part of the model chiefly on the findings of a previous study by Kaufer, Hayes, and Flower (1986), who noted that writers appeared to produce text not in whole sentences but, rather, in sentence parts. Sentence parts were identified either by a pause of two or more seconds or by “a grammatical discontinuity indicating that the current language represents a revision of earlier language.” (p. 23). The amount of text produced in each of these writing bursts was moreover sensitive to the experience of the writer, with more competent writers producing longer bursts (Kaufer et al., 1986). These sentence parts were considered to be formed in accordance with writing plans and their content stored in working memory. Such content would take a surface form in the articulatory buffer and may be articulated “vocally or subvocally” (p. 23), to be finally written down. The author recognized that this system made clause boundaries particularly challenging in terms of working-memory resources—presumably because native speakers’ automatized morphosyntactic abilities make within-clause parsing and production easier than inter-clause production. As a consequence, the model would predict that sentence parts should normally coincide with clause boundaries.

Task schemas were not a new addition, but they were given more prominence in Hayes’ (1996) model. In contrast to the monitor, which was an all-purpose cognitive device for regulating when each component had to be evoked, process-specific task schemas were proposed instead. There may be task schemas specific not only to a particular genre or communicative situation, but also to a specific to a particular process, like revision. In this way, task schemas formed the basis of many aspects of the resulting text and, ultimately, in their communicative success or failure. An inadequate task schema may wrongly guide the execution of a text, for example, when a writer selects casual vocabulary to construct a formal text. It was proposed that another way in which a task schema may impact on text composition is by their absence. When a writer lacks a task schema s/he would need to devise one during text composition, leading most likely to a cognitive overload that would negatively affect the quality of the text. It was argued that writers may simply neglect some aspects of text construction—in the sense that a writer does not necessarily undergo all writing processes, not in the sense that the writing product would remain unaffected—, as in the case of revision or planning. Indeed, Hayes & Flower (1980) were aware that the writing processes they had identified corresponded to the “competent writer”, but that they may be absent in several other writers and/or writing circumstances. In the case of novice writers, several studies have indicated that children rarely, if at all, plan or revise their texts (McCutchen, 2006; Perfetti & McCutchen, 1987). This may well be a result of their lacking task schemas for these processes; alternatively, the cognitive cost assigned to planning or revising (or any other specific task schema) may be impossible to take on in addition to the demands that transcription alone entails. Crucially, there is evidence that task schemas are sensitive to instruction (e.g., Wallace & Hayes, 1991).

1.2.1.3 Chenoweth and Hayes (2001). The third major revision to the Hayes and Flower (1980) model of writing focused primarily on expanding previous research on the translation process. In addition, the structure changed into a representation based on levels of *processes*, *resources*, and *control* (Figure 1.2). The resource level was constituted of the writer's long-term and working memory, as well as by general cognitive processes. The control level was constituted of the task schemas, which determined the nature and organization of the interactions among components of the process level. In this sense, Chenoweth and Hayes gave task schemas even more prominence than in Hayes (1996), since they directed all the writing process. The process level was divided into an external and internal component. The external component, in spite of only specifying physical influences on the writing process (text-written-so-far, dictionaries, and task materials), was described as also including the task environment in virtually the same way as in previous versions, particularly Hayes' (1996). Therefore, although motivational and affective aspects—which were assigned considerable weight in Hayes (1996)—would be captured by this external component, motivational aspects were clearly downplayed in this new version.

The basis for the most important changes in Chenoweth and Hayes (2001) model, with respect to its previous versions, can be found in some of Kaufer et al.'s (1986) findings about the characteristics of sentence production, which chiefly affected the original translating process. In particular, they directed their attention to the relationship between fluency of text production and writing quality, noting that, on average, longer bursts of proposed language were characteristic of more competent

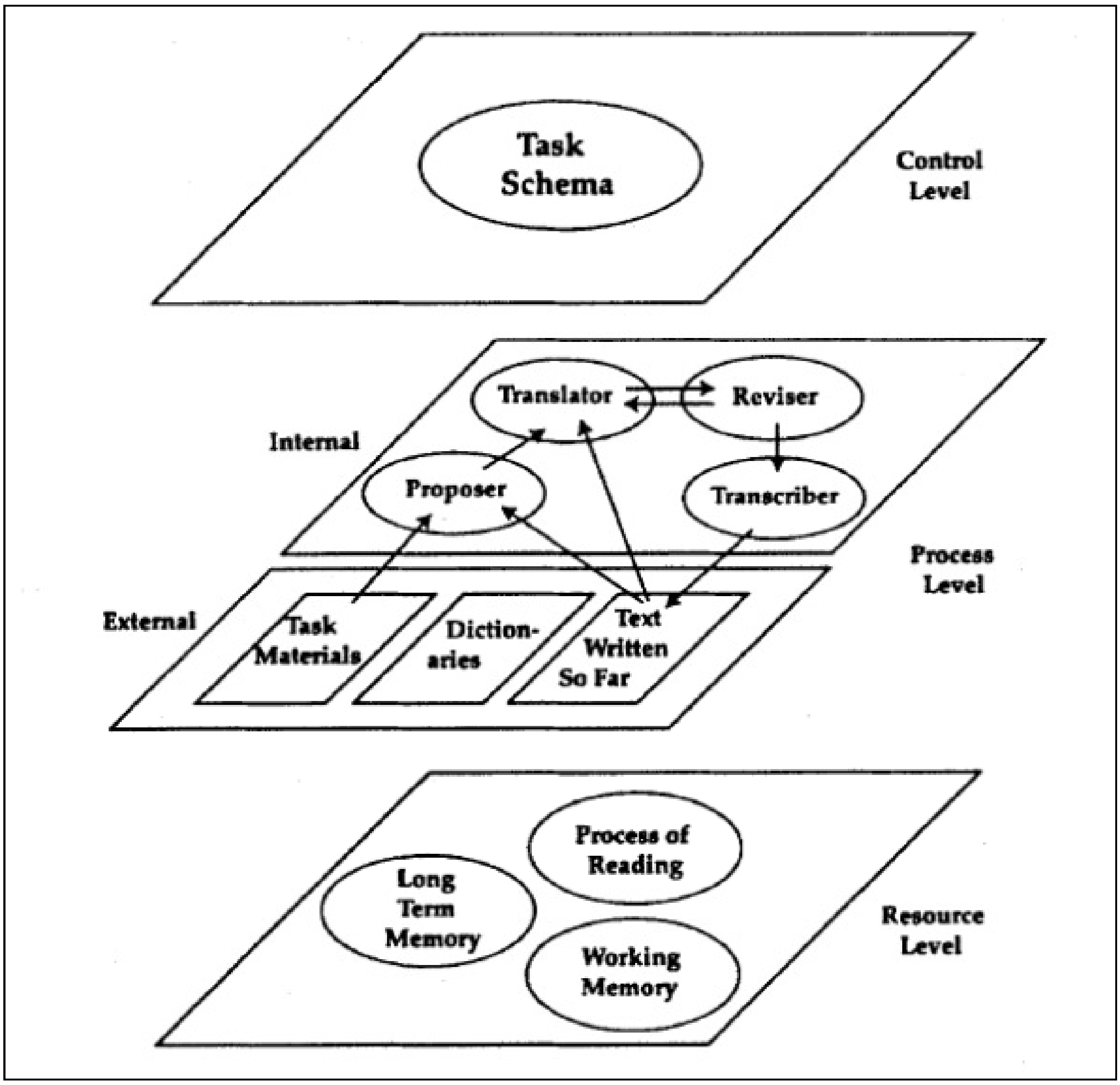


Figure 1.2. Chenoweth and Hayes (2001) model of writing.

writers. Although they viewed writing fluency—i.e., "the rate of text production" (p. 81)—as distinct from writing proficiency—i.e., "accuracy, grammatical and lexical complexity, and appropriateness" (p. 81)—, data such as those obtained by Kaufer et al. (1986) suggested that these two aspects were related. Consequently, the focal point turned to production factors, as they began to be regarded as crucial within the writing process.

Chenoweth and Hayes' (2001) model assumed that pre-linguistic ideas are generated by the *proposer*, which the *translator* converts into linguistic propositions. The role of the *reviser* is to supervise the output of both the proposer and the translator, evaluating it according to varying criteria (e.g., of grammatical well-formedness, meaning, goals, etc.). Finally, the *transcriber* would be responsible for turning linguistic content into written language. In this way, the gap between (pre-linguistic) idea generation and the actual written language output was filled by introducing the translator-transcriber distinction. The authors maintained that these processes do not occur in a linear manner, but that they are in constant interaction, thus allowing for mutual influences.

The model was put to the test by analyzing the think-aloud protocols produced by adult speakers of English, who were learning a second language (L2), in two conditions: during the composition of a text in English and during the composition of a text in the L2 (French or German). The working hypothesis was that writers' reduced linguistic skills in the L2 would result in an overload of cognitive resources during the translation process either because (a) the translator would be less efficient in the L2 to convert pre-linguistic ideas into grammatically acceptable linguistic propositions; or (b) they would need to revise more, given that they cannot trust their automatic second language generation abilities to the same extent as they do during L1 composition. The protocols were transcribed and segmented to identify the rate of text production, measured in number of words per minute; the average burst length, in number of words; and the proportion of proposed words that was finally transcribed in the text. Their results showed a very clear effect of linguistic experience on text production fluency. When writers were composing in their L2, they produced fewer words per minute, shorter bursts, and a lower percentage of their proposed language ended up in their written texts. In addition, writers who had had three semesters of exposure to the L2 showed a poorer performance than those who had been learning the L2 for five semesters. These findings were interpreted as evidence in support of the model, showing the interactions between the writer's skills (i.e., linguistic experience in the text's target language) and the different writing processes (i.e., translation) and between writing processes (i.e., between the reviser and the translator). This flexibility attributed to writing processes to adjust to the writer's knowledge (here, interlanguage levels), may be particularly relevant from a developmental viewpoint, especially when linguistic experience as well as other emerging skills are considered.

1.2.1.4 Hayes (2012). *Modelling and remodelling writing* (Hayes, 2012) may be best described as an elaboration of Chenoweth and Hayes' (2001) version of the model of writing. The 3-level structure was kept (Figure 1.3). The control level was left virtually unchanged, with the only addition of *attention* as another general-purpose cognitive process to be employed by the rest of writing processes. The process level also kept the internal and external distinction, but were relabelled as *writing processes* and *task environment*, respectively, making the domain labels more similar to the original Hayes and Flower model (1980). More detail was provided in relation to the task environment components: the role of collaborators and critics, of the technology for transcription, as well as previously specified components, like the text-written-so-far and task materials.

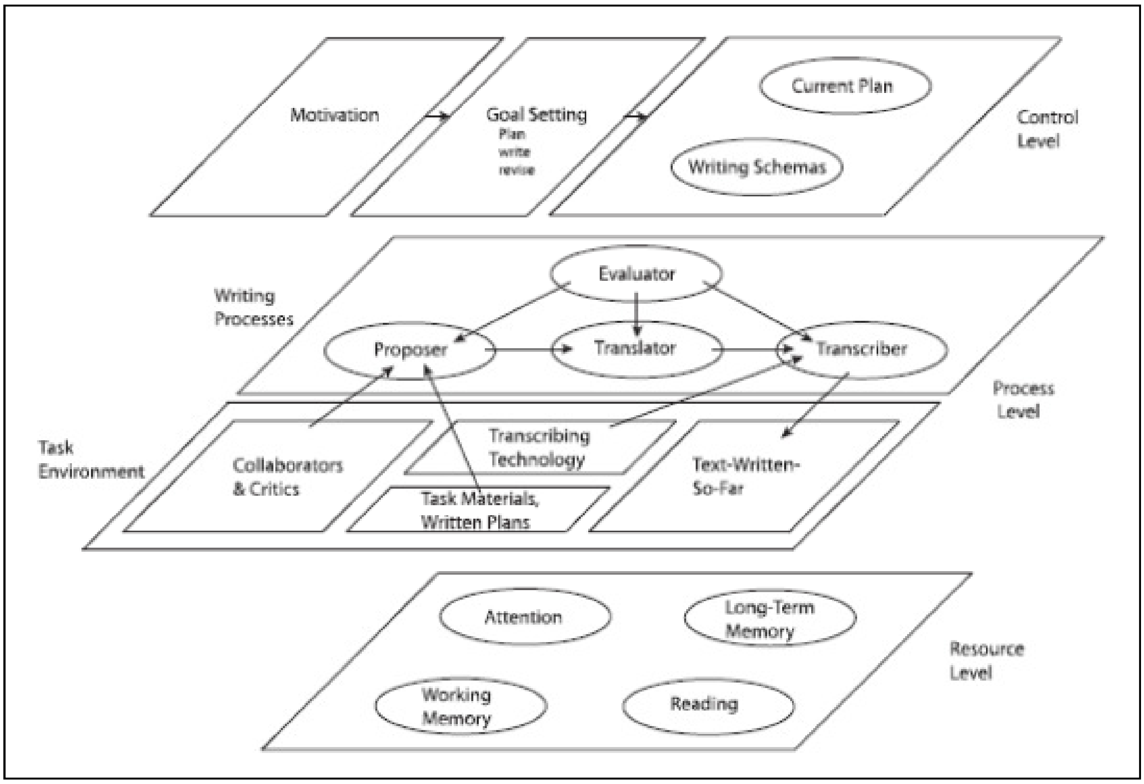


Figure 1.3. Hayes (2012) model of writing.

Two features of the model stand out: the role of motivation and the reconceptualization of revision and planning. Changes to the control level were marked by the return to motivation as a one of its key components. Motivation is hypothesized to influence the writing process in a number of ways, particularly determining the amount of writing people do, as well as having an impact on text quality. However, the model only accounts directly for its influence on goal setting, but does not specify whether or how motivation would have an effect on other writing processes or subprocesses (p. 5). Planning and revising are no longer regarded as writing processes by Hayes (2012), but as special sub-types of writing tasks. Given that "[...] the purpose of dividing writing into subprocesses was to try to understand writing as the interaction among subprocesses, each of which does part of the writing job but not the whole job" (p.7), the author reasons that creating a writing plan (whether it becomes transcribed or not) is a form of writing in the same way that "revising a text is [...] a specialized writing activity." (p. 8). For the sake of parsimony, planning and revising should be considered as two forms of writing, to the extent that each involves the activation of writing processes, subprocesses, and resources.

To conclude, the initial 1980 model underwent considerable revision, modifications, and expansions in almost every component cognitive process. The tendency was thus to increase the model's versatility, so that it could accommodate a broader array of writer profiles. In this sense, affective factors and the role of memory resources constituted a major driving force of such changes (McCutchen, 2006). The initial oversight of the distinctions within the translation component—i.e., between text generation and transcription—have been argued to stem from the fact that the 1980 paper aimed at modelling *expert* writing processes (Alamargot & Chanquoy, 2001; McCutchen, 2006), where translation is supposed to be fairly automatic and not pose a challenge to that kind of writer. Notably, however, the omission did not lessen the impact of the model on developmental studies of writing.

1.2.2 Models and Theories of Writing Development

Research on writing development has tended to focus on either the writing process or the writing product (Berninger, 2009). Writing development from a product perspective typically described the path from toddlers' first realization of the permanence of marks obtained by using a crayon, to the differentiation of drawing from writing, to the increasing awareness of the relationship between writing symbols and oral language, and, finally, to the proficient use of writing as a means of linguistic expression (Gibson & Levin, 1975; Graham & Winetraub, 1996; Traweek &

Berninger, 1997). Ferreiro and Teberosky's (1979) pioneering study demonstrated that children's mere exposure to print—as is the case for children growing up in literate communities—, allows them to develop hypotheses about the nature, purpose, and function of written language. In this sense, it has been argued that, by the time children begin to be formally taught to read and write, they are already in possession of a wealth of knowledge about writing. Moreover, there is evidence that they show early metalinguistic and cognitive predispositions in line with the specific characteristics of the writing system of their community (Tolchinsky, Levin, Aram, & McBride-Chang, 2012). Children have been shown to exhibit different performance levels in preschool measures of phonological awareness, letter knowledge, and visuo-spatial perception, as a function of whether they have been raised within a community that uses an alphabetic (Spanish), abjad (Hebrew), or morphosyllabic (Chinese) writing system (Tolchinsky, et al., 2012).

Within the framework originated by Ferreiro and collaborators, it is claimed that children's knowledge about writing, as well as their written productions, follow a predictable order of acquisition. This line of reasoning about written language development has been termed the *linearity hypothesis.* Proponents of the linearity hypothesis of writing development (e.g., Ferreiro & Teberosky, 1979; Gibson & Levin, 1975; Tolchinsky, 2003) argue that children incorporate knowledge about writing in a stage-like manner, with universal features of writing (like symbolic representation of language, linearity, or discreteness of symbols) developmentally preceding and setting the bases for the learning of language-specific traits (e.g., directionality, letter forms, etc.).

The linearity hypothesis has not gone unchallenged. An alternative account, the *unified hypothesis*, claims that preschool children learn writing features in no particular order (e.g., Treiman, Cohen, Mulqueeny, Kessler, & Schechtman, 2007). Instead, children have been argued to pay attention to "salient features" of writing, such as the letters in their own name (Treiman, et al., 2007; but see Puranik & Lonigan, 2011, for a discussion of methodological issues with some of these studies). Recently, however, the linearity hypothesis has received strong empirical support in a study with over 300 English-speaking children, with ages ranging from 3 to 5 years old (Puranik & Lonigan, 2011). Applying a scaling procedure to a group of text features (including universal, as well as English-specific writing features), which had been ordered according to a hypothesized degree of sophistication and specificity, they found robust indicators of a sequential path in the learning of such features. All 3-year-olds in the sample had acquired most of the universal writing features, while more refined features could only be found in the writing samples of 4- and 5-year-olds. Importantly, the finding held for five different writing tasks (writing of letters, CVC words, child's name, sentences, and short text). Different writing tasks had been included, in view of arguments about the influence of task type in determining the maturity of writing forms (i.e., degree to which spellings represent the phonological structure of words) used by preschoolers (Sulzby, Barnhart, & Hieshima, 1989). Indeed, Puranik and Lonigan (2011) found that the simpler the writing task (e.g., writing letters or writing their own name), the more advanced the writing forms children used (e.g., invented spelling vs. scribble). Finally, although improvement over time was attested for all writing tasks, significant differences were only found for word- and subword tasks (writing letters, child's name, and CVC words); performance on sentence- and, especially, text-level tasks did not improve with time. In sum, Puranik and Lonigan's (2011) findings indicate that children show knowledge about writing from as early as age 3, and they seem to move towards an appropriation of the symbolic function of writing well before being aware of the exact nature of the relationships between oral and written language (e.g., knowledge of phonographic correspondences). Importantly, in the age ranges included in Puranik and Lonigan's (2011) study, children showed a remarkable amount of word and sub-word knowledge, but their sentence- and text-level performance seemed to lag far behind, suggesting that these levels of writing performance may constitute a protracted development.

Approaches that have focused on the cognitive processes occurring during text composition originated in attempts to describe adult or expert writing (e.g., Hayes & Flower, 1980). The shift from product to process approaches has been attested ever since in psycholinguistic studies (Berninger et al., 1996). It should be noted, however, that more recent explorations of writing products do not confine themselves to describing the written outcomes, but are regarded as a window to access the cognitive skills involved in text composition (Tolchinsky, 2006).

In what follows, I review the most important models of the writing process that were devised to describe and explain the characteristics of beginning and developing writers. These include the two text production strategies proposed by Bereiter and Scardamalia (1987): *knowledge-telling* and *knowledge transforming*; the *simple view* of literacy (Juel, Griffith, & Gough, 1986); and those models more explicitly based on the original Hayes and Flower (1980) model of writing, but modified in order to accommodate evidence from beginner and developing writers. Often, the various models proposed do not constitute alternatives, but may complement each other, in the sense that some of them have a larger scope (e.g., Hayes & Flower, 1980), while others may focus on a specific aspect or phenomenon of the writing process (Alamargot & Chanquoy, 2001).

1.2.2.1 Bereiter and Scardamalia's model (1987). This seminal work on writing development offered a dual framework for approaching the writing process. One alternative saw the writing process as a natural, easy task, drawing on effortlessly acquired linguistic knowledge and everyday social interaction. The other alternative involved a conception of writing as a highly complex skill, one that goes well beyond naturally endowed faculties and ordinary social interaction, requiring conscious effort and specific training. These two approaches to the writing process were labelled by *knowledge telling* and *knowledge transforming* models of writing, respectively.

Within the knowledge-telling model, writers may extrapolate their spoken sentence- and text-production skills to construct a written text. Given that children have already mastered the core grammar of their L1 by the time they begin to be formally taught to write (Berman, 1997, 2004; Karmiloff-Smith, 1986), this view of writing should develop relatively easily. There are a number of caveats, however, derived from the very nature of writing. One of them is, of course, the written code. Children need to learn the written language code, that is, the orthography, to be able to write. Another issue is the fact that writing, as a mode of production, is defined by the absence of an immediate addressee or interlocutor, in sharp contrast with conversation, where the support of the interlocutor helps with staying on topic, maintaining coherence, and may even intervene to require clarifications as needed (p. 7). This issue, however, may be resolved by resorting to the cues provided by the text produced, the writing assignment, and by task schemata (e.g., narrative text, argumentative essay, etc.). The graphic representation of the knowledge-telling model is shown in Figure 1.4. Once an assignment has been proposed (e.g., by a teacher or the writer himself), genre markers and topic markers need to be identified to begin what will mostly be a search for relevant content. As is evident from the representation of the model, the resulting text is highly dependent on the child's knowledge about the topic in question, as well as on discourse knowledge; that is, task schemata. Bereiter and Scardamalia claimed that variation in discourse knowledge is thus related to whether the writer has a more or less detailed representation of a given discourse schema, such as that of an argumentative essay:

> Some immature writers may have an opinion-essay schema that contains only two elements—*statement of belief* and *reason* [...]. Others may have more complex schemas that provide for multiple reasons, anticipation of counterarguments, and so on. (p. 9, emphasis in the original).

It follows from this account that beginner writers, who may have no stored writing schemata for many text types, may simply produce one statement (sentence) after another, until they feel they cannot produce more content that is relevant to the topic. Indeed, the authors report a description provided by a 12-year-old student about what the writing process entails:

> I have a whole bunch of ideas and write down until my supply of ideas is exhausted. Then I might try to think of more ideas up to the point when you can't get more ideas that are worth putting down on paper and then I would end it. (p. 9)

In this sense, they argued that the knowledge-telling framework supplies immature writers with an efficient strategy to produce written text, provided the prompt is clear enough, without the need for interlocutor support. Within the knowledge telling strategy no regard is paid to global coherence, communicative effectiveness, or rhetorical aspects of text construction.

In contrast, the knowledge-transforming model of writing (Figure 1.5) is one where what the writer knows, his/her ideas, may be modified or elaborated over the course and as a result of the writing process. A typical knowledge-transforming text would produce the impression that the text was constructed with a defined purpose, and would show great regard for the specific needs of its distant reader (p. 29). The knowledge-transforming model incorporates the knowledge-telling model as well, but only as a problem-solving process that may be called upon by two main "problem spaces": the *content* and the *rhetorical* domains. The content problem space comprises knowledge and beliefs about a topic, while the rhetorical problem space is concerned with the writing goals. It is assumed that these two domains are in constant interaction and may influence each other. For example, in trying to resolve a text coherence problem (i.e., rhetorical problem space), the writer may be led to

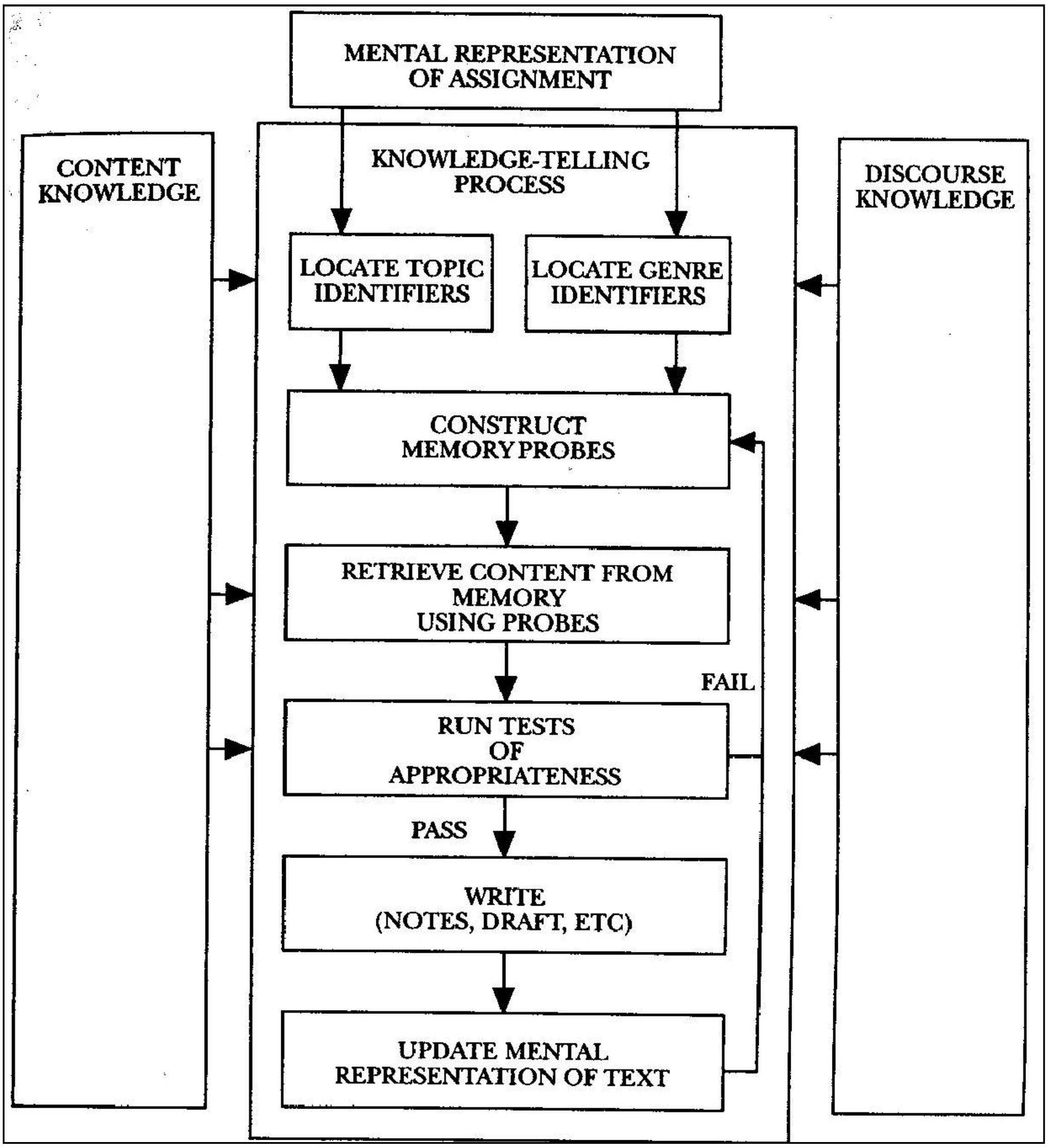


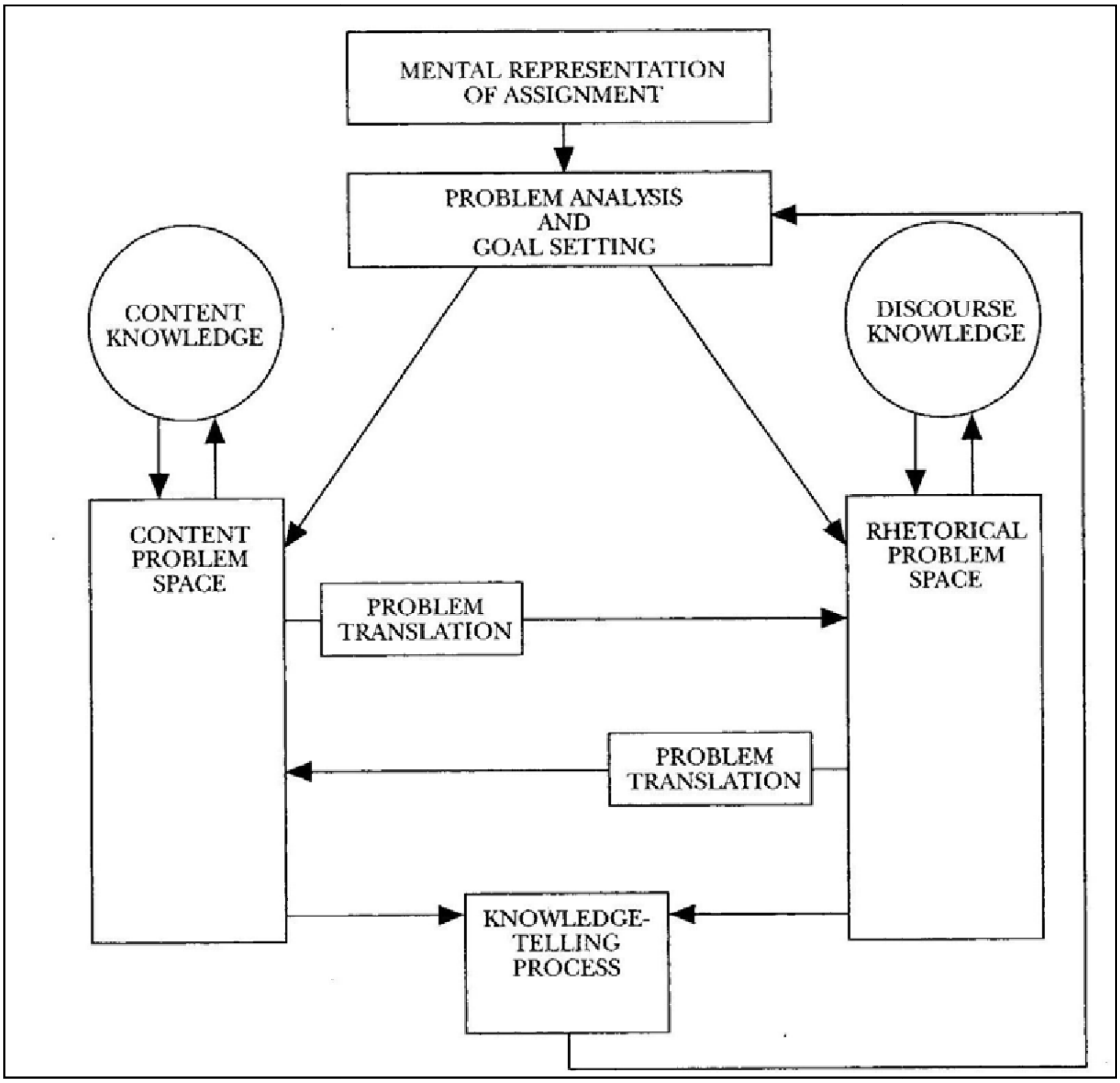
Figure 1.4. Bereiter and Scardamalia's (1987) knowledge-telling model

question whether his own understanding of the topic makes sense, and may choose to either reformulate or reconsider his/her ideas.

It is because of their entirely different nature, that two models—and not one—are needed to reflect alternative strategies in the composition of a written text. The authors emphasized that the knowledge-transforming model, though more characteristic of adult, skilled writers, should not be considered as an improved version of the knowledge-telling model. Actually, they pointed out how the knowledge-telling model may sometimes constitute a more efficient or sensible choice in particular text construction circumstances; for example, when the writer has memorized verbatim what s/he wants to say, in casual writing (e.g., a letter to a friend), or in a first raw draft of a text.

Given that the two models were not simply meant to describe the cognitive processes underlying written composition, but to provide *explanatory* parameters of observable behaviour, its usefulness was demonstrated by testing some of the predictions based on the models. For example, the authors reasoned that the time before actually beginning to write would be shorter for writers adopting a knowledge-telling strategy, than for those adopting a knowledge-transforming strategy, given that the latter would need more time to consider rhetorical, audience-related, and global text construction aspects, in addition to content-retrieval processes.

Bereiter and Scardamalia's dual model of text construction has had a strong impact in developmental accounts of writing. Two related studies are noteworthy since they further elaborated the outcomes of the knowledge-telling model, identifying various strategies of text construction by beginner and developing writers. The first study was the result of a doctoral dissertation by Fuller (main findings reported in Berninger et al., 1996) who looked at primary school (grades 1 through to 9) children's elicited narrative and expository texts. Her aim was to identify child-developed rules of thumb or writing schemes that children used when composing texts. Fuller thus identified several writing rules that children seemed to follow during text composition. One example of such early writing plans was labelled the "chain" structure. The child begins by producing a sentence, usually in response to the topic or prompt. This first sentence, in turn, provides the prompt for the topic of the next sentence. For example, in response to the prompt given by the experimenter of completing an open-ended sentence starting *I like_____ because…*, the first sentence in the "chain" structure may read *I like coloring because it's not boring*. The following sentence keeps with the "coloring" theme, thus producing, *I like to color cats.* Then, "cats" activates a new theme, giving place to a sentence like *I have a black cat at home*. Another child-developed writing scheme termed "the wheel" involves producing a series of propositions all around the same topic, but without any concern for marking how each sentence relates to the other. The example provided by the authors is one in



. *Figure 1.5.* Bereiter and Scardamalia's (1987) knowledge-transforming model.

which the writer produced nine different sentences whose only shared property is a friend of hers, named Ashley: *I like Ashley. She is nice. Ashley is my friend. I like people and Ashley is one. I like Ashley cus shis nis* (Berninger et al., 1996, p. 206). Fuller interpreted these types of texts as having no regard for creating a global topic.

Fuller's dissertation findings motivated a reformulation by Hayes (2011, 2012), who placed these early child-developed writing plans within Bereiter and Scardamalia's (1987) knowledge-telling model, regarding them as sub-types of knowledge-telling strategies. Hayes reanalyzed the corpus used in Fuller's thesis and reckoned that three knowledge-telling strategies could account for at least 90% of the texts written by first- to ninth-graders. These strategies were labelled *flexible focus, fixed focus,* and *topic elaboration*. Children who compose written text using the flexible focus strategy produce sentences which vary in their topic. Fuller's "chain" structure would thus fall under this category. Children who use the fixed-topic strategy produce texts of which the sentences all share the same topic, as in Fuller's "wheel" structure. Finally, children who use the topic elaboration strategy produce sentences of varying topics, but some of them are an elaboration of one sub-topic or idea. After elaborating on one sub-topic, the writer typically comes back to the original topic and so on. Interestingly, each one of the three strategies seems to follow their own particular developmental trajectory. While the fixed-topic strategy was more typical of texts produced by first- to fourth- graders, topic elaboration texts increased as a function of grade level, from 13% in first grade to 63% in ninth grade (Hayes, 2011, p. 88). Flexible-focus texts did not vary with grade level and never accounted for more 10% of the corpus.

In sum, Bereiter and Scardamalia's knowledge-telling and knowledge-transforming models provided a useful framework to understand the process of text composition, accommodating both skilled and immature writers' profiles. It should be stressed that the emphasis of both the knowledge-telling and the knowledge-transforming strategies is on higher-level aspects of text construction. Although transcription skills were ignored by the authors (Bereiter & Scardamalia, 1987, Ch. 4; Scardamalia, Bereiter, & Goleman, 1982), the two models do not factor in the development of transcription components. This particular characteristic of Bereiter &

Scardamalia's proposal makes it stand out from the rest of developmental models of writing processes.

1.2.2.2 Juel, Griffith, & Gough (1986). The simple view. The *simple view* of reading acquisition (e.g., Hoover & Gough, 1990) suggested that reading ability results from the interaction of decoding, on the one hand, and listening comprehension, on the other. Either component was moreover assumed to be insufficient to account for reading on its own; nevertheless, it was ascertained that a certain amount of decoding was necessary before reading comprehension could take place. Underlying these claims was the notion that reading comprehension entails higher–level processes, which require a degree of automaticity in their component lower-level processes—i.e., decoding. In other words, before readers can deal with a cognitive task as challenging as reading comprehension, more "basic" decoding abilities need to absorb the smallest possible amount of cognitive resources.

Juel, Griffith, and Gough (1986) applied the *simple view* to the study of writing development. They maintained that writing ability involves two main processes: spelling and *ideation* (organization and generation of ideas). In this way, spelling was taken to mean to writing what decoding meant to reading comprehension in the reading studies (Gough & Tunmer, 1985). A point was made that, in spite of the simplicity of the model, there was no assumption that the component processes are simple in themselves, although there was no elaboration or assessment of the processes underlying ideation. As for spelling skills, Juel et al. (1986) assumed that both decoding and spelling rely on the same sources: knowledge of sound-spelling correspondences, termed the *orthographic cipher,* together with lexical knowledge and exposure to print. The cipher would thus account for all regular spellings, while lexical knowledge and print exposure would account for irregular spellings, as well as to aid in deciding among plausible spellings in regular words, something particularly relevant for inconsistent orthographies like English. Finally, phonological awareness was hypothesized to be a precursor of spelling and of decoding skills. Similarly to the reading studies, the model thus predicted that lower–level skills (spelling) determine to a large extent writing ability during the early years of writing development, while higher-level skills (ideation) become progressively more important, once spelling has become automatized.

The authors found supporting evidence for the model, in which decoding and reading comprehension were also included. Their findings also suggested, in line with their predictions, that decoding and spelling do rely on the same set of subskills—namely, knowledge of the cipher, lexical knowledge and print exposure, and phonological awareness—, though reading comprehension and writing were not significantly correlated in their longitudinal study. The authors concluded that the lack of a strong relationship between the higher-level literacy domains was the result of the higher-level components of reading and writing—i.e., listening comprehension and ideation, respectively—not being "isomorphic".

The simple view has the merit of having produced a most parsimonious model of reading and writing development that offered clear, testable predictions. However, it may be accused of oversimplifying vital aspects of text construction, given that the "ideation plus spelling" definition fails to account for writing-specific rhetorical style(s), text structure, and other higher-level text composition features. It completely ignores the task environment and its effect in the writing process and product, as well as other higher-level processes such as text planning and editing. It could be argued that the model only intends to account for emerging and developmental writing—where most of the higher-level processes and text features are usually absent. In such a case, however, it would be poor in explanatory value, since it would be very hard to elucidate the what, when, and how, of the transition between this simple model and a model better fit to accommodate more mature writing. The simple view has moreover been claimed to embody the *reductionist* approach to literacy (Tolchinsky, 1996), given that it is based on the general assumption that "the ability to read and write can be decomposed into a number of simpler components", where "knowledge of texts is explained in terms of knowledge of words, knowledge of words in terms of knowledge of sounds." (Tolchinsky, 1996; p. 103).

1.2.2.3 Berninger, Mizokawa, and Bragg (1991). Developmental constraints model. During the mid- and late 1980s, a group of investigators led by Prof. Virginia W. Berninger at Washington University, in Seattle, started conducting a series of research projects on the development of literacy, with a particular focus on writing (for a review of the full research program and its major findings see Berninger, 1999). Central tenets to this line of research included an interest in the process, rather than the product, and in the clinical uses of writing research. Essential to those aims was a detailed investigation of the developmental characteristics of each writing process. Early research findings (e.g., Berninger, Mizokawa, & Bragg, 1989) led them to believe that Hayes & Flower's (1980) model was not fully adequate to reflect the development of text writing ability.

Berninger, Mizokawa, and Bragg (1991) proposed a multiple constraints theory of writing development to provide both a framework for studying emerging and developing text composition skills, and to serve as reference for clinical purposes (e.g., identification and remediation). Based on Lurian neurodevelopmental theory (Luria, 1973, cited in Berninger et al., 1991), which postulates that sensory/motor skills develop before sensory-motor integration skills, which in turn develop prior to higher-level language and other cognitive processes, they developed a model in which constraints of different sources—neuropsychological/neurodevelopmental, linguistic, and cognitive—operate at different stages in writing development.

Low-level *neuropsychological constraints* were argued to have the largest impact at the earliest stages of writing development (i.e., grades 1-3), given their importance in determining handwriting skill and in the automatization of spelling. These skills were (1) the retrieval of alphabet letters from visual memory; (2) the "neurological soft-signs elicited by finger tasks", estimated using a series of tests of finger movements; (3) visual-motor integration, typically measured by tasks requiring the child to imitate geometric shapes of increasing difficulty using pencil and paper; and (4) orthographic and phonological coding, comprising word-, letter-, and letter cluster-spelling tasks). A second stage (grades 4-7) would be characterized by the pre-eminence of *linguistic constraints*, which would become relevant only after the lower-level skills have been relatively automatized. These constraints would operate at different "levels of language": word, sentence, and paragraph. Lastly, the *cognitive constraints* of planning, translating, and reviewing are likely to constrain the writing process in grade 7 and above, once low-level processes are automatized and the language levels of word, sentence, and paragraph have been mastered.

The multiple-constraints theory is interesting from a cross-linguistic viewpoint because some of the constraints may present more cross-linguistic variation than others. For example, differences in spelling (or orthographic coding) have been largely attested in several studies (see Section 1.4 below). However, it is not clear whether one should expect cross-linguistic variation at the paragraph level and, even less, in the characteristics of the writing process as a whole. To the best of my knowledge, there is not one paper arguing against Hayes and Flower's (1980) model of writing (or against any model of writing, for that matter) on the grounds of its inadequacy to fit cross-linguistic data.

1.2.2.4 Berninger and Swanson (1994): Modifying Hayes & Flower model. The authors proposed a number of modifications to Hayes & Flower's (1980) model of skilled writing in order to accommodate the nature of beginning and developing writing. Based on a series of cross-sectional studies with primary, intermediate, and junior-high students (Berninger et al., 1992; 1994; Whitaker, et al., 1994), in which they analyzed the writing products as well as several literacy-related cognitive skills, their contribution may be summarized in four main points: (1) the specification of components within the translation process—the only writing process not including sub-processes in Hayes and Flower's (1980) model—; (2) the distinction between different types of planning and reviewing, distinguishing online, local from off-line (or post-translation), global types; (3) the order of acquisition of the different writing processes and sub-processes; and (4) the respecification of the role of memory resources at different developmental stages.

A fundamental contribution of Berninger's laboratory to theorising about writing development was their call for a detailed breakdown of the processes within the translation component, in view of the cognitive effort that they entailed for novice writers (e.g, Berninger et al., 1991; 1992). Unlike the other two writing processes in Hayes & Flower's (1980) model, planning and reviewing, which included sub-components, translating had been left without further specification. As a point of departure, Berninger and Swanson (1994; see also Berninger et al., 1992; 1994) distinguished between *text generation* and *idea generation.* The writer engages in idea generation when s/he proposes pre-linguistic content, a process that does not belong within the translation process, but within planning. Translation was thus subdivided into *text generation* and *transcription*. Text generation is concerned with transforming those pre-linguistic ideas into language, according to the lexicon and grammar of the language. Transcription is the process by which those ideas are encoded into written language. The distinction was deemed relevant because of evidence indicating that both sub-processes may be dissociated in developing writers (Berninger, et al., 1992) and in some writing disorders (Graham, 1990). Moreover, the transcription sub-component was argued to be supported by spelling and handwriting, which constitute the lower-level components of writing (Berninger et al., 1994).

A final revision within the translation component was related to the role of oral language skills and written language skills at different levels—i.e., word, sentence, text—, which were found to affect text production development (Abbott & Berninger, 1993; Berninger, et al., 1994). In this sense, it was suggested that writing skills are not at the same stage developmentally in all linguistic units, but that different levels of proficiency may be found for the word, sentence, and discourse levels. Writing and, more generally, literacy skills, are also more strongly related within the same language level (e.g., word writing and word reading) than across levels (e.g., word writing and text writing).

A second point in Berninger & Swanson's (1994) revision of the Hayes & Flower (1980) model involved the emergence and development of the planning and reviewing processes. Based on empirical studies, they found evidence for distinct types of planning and revising at different points in the development of writing. Although Hayes and Flower did refer to various types of planning and revision, no explicit differentiations were incorporated to the original model, making it difficult to apply to developing writing. Berninger and Swanson observed that online planning tends to be concerned with content generation, while online reviewing chiefly attends to superficial aspects like correcting spelling, handwriting legibility, and some word choices. In contrast, global, advanced planning and post-translation reviewing are more concerned with the overall coherence and communicative effectiveness of the text, attention to audience, adequacy to rhetorical goals, and so on. In a later study (Berninger et al., 1996), claimed that temporal and spatial dimensions should also be considered in the development of planning and revising. They argued that whether planning and revising occur before, during, or after translation depends on the maturity of the writer. In addition, writers seem to follow a clear developmental pattern: planning simultaneous to translation precedes advanced planning, while online revision emerges later than guided, after-translation revision (Berninger, Whitaker, Feng, Swanson, & Abbott, 1996). Similarly, revision of parts of the text seems to develop earlier than whole-text revisions (Berninger, et al., 1996).

The third important finding reported in Berninger and Swanson (1994) was concerned with the developmental order in the emergence and development of writing processes. Although young writers may engage in planning, translating, and reviewing, these processes are not all present from the very outset of development. Translating seems to operate primarily in the first stages of writing development, with transcription preceding text generation. Moreover, the authors claimed that both transcription and text generation emerge at different rates as a function of the language units in question: word level seems to develop before sentence level, which in turn precedes text level. Planning and revising, on the other hand, would emerge later on, with local/online types emerging prior to global/off-line types.

The fourth key contribution reported in Berninger and Swanson (1994), highlighted the essential role of short-term memory in writing development, over and above the functions of long term memory, as initially recognized by Hayes and Flower (1980). As mentioned above, short term memory was included in the revised versions of the model (Hayes, 1996; 2012; Chenoweth and Hayes, 2001), as well as by other models of skilled writing (e.g., Kellogg, 2001; McCutchen, 1986). In their review, Berninger and Swanson did not find working memory to be a particularly relevant factor at the very early stages of writing development (grades 1-3). At these stages it would seem that mastering transcription is perhaps the top priority for children, so that enough cognitive resources may be freed up to be devoted to other aspects of text generation. However, as writing tasks increase in complexity, working memory and the particular way in which processes and subprocesses are orchestrated during text composition become more important. For these reasons, they concluded that working memory must be included in any developmental model of writing, especially when dealing with intermediate and junior-high writers.

Finally, Berninger and Swanson revisited the adequacy of Berninger et al.'s (1991) developmental constraints theory of writing. They concluded that, in light of the findings reviewed, the neurodevelopmental, linguistic, and high-level cognitive constraints, seem to be operative throughout the development of writing, rather than constituting discrete stages. In addition, they pointed out that, although the influence of transcription and text generation in text production fluency and quality tended to decrease over time—in favour of higher-level cognitive processes, whose weight tended to augment—, transcription skills maintained their influence over compositional fluency and quality even up to adolescence. Further support for this standpoint was found by other investigations, suggesting that, at least under certain circumstances or in particular populations, the mechanics of writing seem to exert a long-lasting influence over the writing process (e.g., Bourdin & Fayol, 1994, 2002; Graham, 1990; Jones & Christensen, 1999).

1.2.2.5 Berninger and Amtmann (2003) and Berninger and Winn (2006): the writing triangles. The work carried out by Berninger and collaborators insisted on the ongoing influence that low-level transcription skills have on text generation, in interaction with executive function factors. For these reasons, Berninger and Amtmann (2003), proposed the second simple view of writing (Figure 1.6). The triangular representation clearly reflects the conceptualization of transcription skills, as well as cognitive, emotional, and other writing-specific regulatory factors, as the foundations upon which text generation rests. Also apparent from the model's representation, is the idea that text generation comprises word, sentence, and text levels of performance, allowing for differences and similarities within and across these levels.

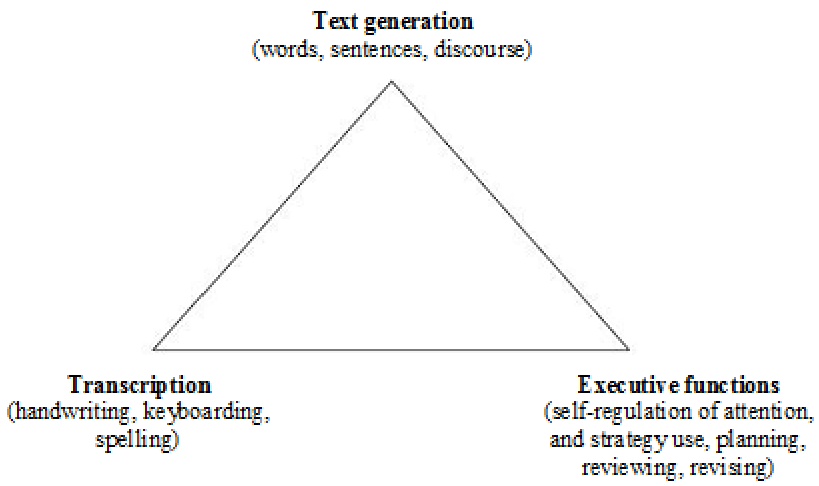


Figure 1.6. Berninger & Amtmann (2003) second simple view of writing.

A revision of Berninger and Amtmann's (2003) simple view of writing found its motivation in the need to accommodate the increasing complexity and more multidimensional nature of more advanced writing. The vertices of the triangle remained virtually identical to Berninger and Amtmann's (2003) model, maintaining transcription and executive functions at the base of text generation (of words, sentences, and text). Berninger and Winn's (2006) model further assumed that interactions among these processes take place in a working memory environment (Figure 1.7; Berninger, 2008; Berninger & Winn, 2006). The role of working memory as particularly relevant from the intermediate grades and beyond was not a new idea, since it had been already stated by Berninger and Swanson (1994). Therefore, this last specification of the model aimed to provide a more succinct, parsimonious model of writing processes by beginner writers and the characteristics of the transition into more advanced forms of writing, for which working memory is assumed to be at the core (Berninger, García, & Abbott, 2009).

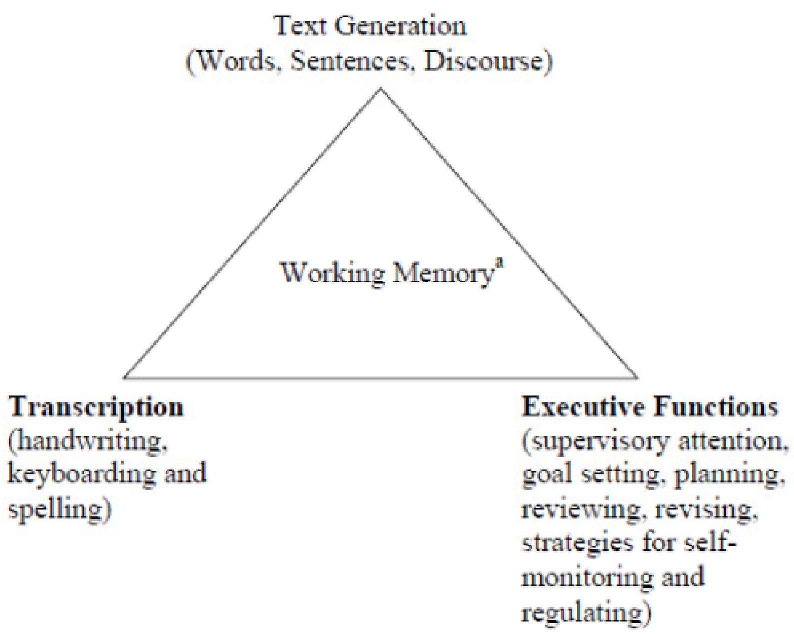


Figure 1.7. Berninger & Winn (2006) not-so-simple view of writing.

To sum up, transcription skills had been majorly overlooked in early models of skilled writing (particularly Hayes & Flower's, 1980), whose chief interest was to elucidate higher-level aspects of written text construction. In contrast, low-level writing factors quickly became the centre of attention of virtually all accounts of writing development. While some early studies considered that the effect of non-automatized/immature transcription skills in beginning writing did not compromise text quality (e.g., Bereiter & Scardamalia, 1987; Scardamalia, Bereiter, & Goleman, 1982), it is now generally accepted that low-level transcription skills are a pre-requisite for children's attention to be devoted to other aspects of the writing process and, thus, for the emergence and development of higher-level writing processes (Berninger & Swanson, 1994). Research by Berninger and collaborators (e.g., Abbott et al., 2010; Berninger et al., 1992; 1994; Berninger & Amntmann, 2003) moreover pointed to a sustained influence of spelling and handwriting throughout the primary grades, and even into adolescence.

1.3. The Role of Transcription Skills in Writing Development

Writing is often described as a complex activity that requires the orchestration of a myriad of cognitive processes and subprocesses (Chanquoy & Alamargot, 2001; Hayes & Flower, 1980). Current theories of writing and writing development converge on the understanding that the different processes are in constant competition for the limited pool of available cognitive resources (e.g., Berninger & Swanson, 1996; Bourdin & Fayol, 1994, 2002; Kellogg, 2001; McCutchen, 1986). Alternative explanations of cognitive resource management during composing (e.g., Torrance & Galbraith, 2006) suggested that writing development may not necessarily be a matter of process competition, but rather that processing of some writing tasks (e.g., spelling) may interfere with processing of other writing tasks (e.g., text generation). In either case, it is clear that some sort of hierarchy is established among writing processes, in which favouring (or focusing on) some particular aspect of text construction, negatively affects performance levels of other aspects. In this sense, lower-level transcription skills—i.e., spelling and handwriting—appear to be good candidates for skills that may be automatized to liberate cognitive resources for higher-level writing processes. Indeed, developmental accounts of writing view handwriting and spelling as comprising children's central concerns in the early stages of learning to write.

The impact of lower-level writing components on the writing process and product may be explained in several, non-exclusive ways. As stated before, from a cognitive resources perspective, mastery of handwriting and spelling skills frees up resources, so that they may be used by other higher-level writing processes, such as planning or content retrieval. From a motivational/affective perspective, children who are successful at spelling words and who do not struggle with handwriting, may show a better disposition and feel more confident about producing written text, than struggling writers (Hayes, 2011, 2012; Pajares & Valiante, 2006). From a strictly writing process perspective, when only minimal attention to low-level aspects of writing is required, then transcription should not interfere with other processes involved in writing (Torrance & Galbraith, 2006), making handwriting and spelling skills a top priority in writing development.

Not surprisingly, spelling and handwriting have been shown to be tightly related to one another during early writing development and to have a direct or indirect impact on text writing fluency (Berninger et al., 1992; 2002; Graham et al., 1997) and quality (Abbott, Berninger, & Fayol, 2010). A serious gap within early writing development research, however, is the scarcity of cross-linguistic comparisons. The vast majority of investigations to date have been carried out with children learning English orthography, which is usually quoted as a prime example of complex, opaque orthography, at least as far as the alphabetic orthographies are concerned (e.g., Share, 2008). Although there are some studies on early writing development involving more consistent orthographies (e.g., for Spanish: Borzone & Signorini, 1998; Borzone de Manrique & Diuk, 2003; and Sánchez Abchi, Borzone, & Diuk, 2007; for Finnish: Lerkkanen, Rasku-Puttonen, Auriola, & Nurmi, 2004; and Mäki, Voeten, Vauras, & Poskiparta, 2001; for Turkish: Babayiğit & Stainthorp, 2010), no systematic cross-linguistic comparisons have been made for children at the very early stages of learning to write. Given the central role bestowed on transcription factors in early writing development, it seems imperative to address this issue from a cross-linguistic viewpoint to test whether English-based conclusions are generalizable to other orthographies. This is a crucial point, considering the theoretical and educational implications.

Studies on the abilities that underlie both handwriting and spelling seem to indicate a tight relationship between the two skills (e.g., Graham et al., 1997). In one of the few studies looking at both transcription skills simultaneously, Berninger et al., (1992) conducted a series of multiple regressions to identify common and unique predictors of handwriting, spelling, and composition in a sample of children attending 1st, 2nd, and 3rd grade,. Predictor measures included a large battery of tasks tapping low-level neurodevelopmental, cognitive, and linguistic skills. A letter retrieval measure (*alphabet task*, Berninger et al., 1992), and orthographic coding measures explained unique variance in both spelling and handwriting skills in children from grade 1 to grade 3 (Berninger et al., 1992). However, orthographic coding of letter *clusters* (identifying whether a sequence of letters had been shown previously or not) was a significant predictor of handwriting, while orthographic coding of *whole words* contributed to explaining variance in spelling. Predictors unique to handwriting involved low-level neuro-motor function skills (a timed finger succession task), while unique predictors of spelling were verbal IQ, visual-motor integration (a non orthographic task where children copy geometric designs), phoneme segmentation, and nonword reading. In short, both handwriting and spelling appear to be driven by underlying skills that are subject to ample inter-individual variation. Given that the uniqueness of handwriting seems to be related to neuro-motor, low-level developmental skills, it should not constitute an important source of cross-linguistic variation. Unique facets of spelling skill, on the other hand, appear to be strongly rooted in the particular interface between the characteristics of oral language and how it is encoded into written language, at least at the phonological level of description. For this reason, it is essential to explore spelling/writing from a cross-linguistic perspective.

1.4 Cross-linguistic Approaches to Spelling

Differences in literacy development across languages have been claimed to arise, to a great extent, as a consequence of the consistency of the phoneme-grapheme mappings. The *Orthographic Depth Hypothesis* (ODH, Frost, Katz, & Bentin, 1987; Katz and Frost, 1992) was proposed to account for such differences. According to the ODH, alphabetic writing systems may be placed at some point in an orthographic transparency continuum according to how consistently the orthographic code represents the phonology of the language. A completely consistent system would be one in which, for every phoneme in oral language, there is only one grapheme to represent it and, vice versa. Conversely, orthographies are considered to be inconsistent when several possible graphemes may represent a single phoneme, and vice versa. Within this view, when phoneme-grapheme mappings are not very consistent, then users of that orthography may need to develop additional, non-phonological strategies for word decoding and encoding, such as lexically based strategy (though see Harm & Seidenberg, 2004 and Pollo, Treiman, & Kessler, 2008, for alternative views). The ODH may be thought of, then, as a theory about how efficient phoneme-grapheme mappings are for learning to read and spell. Clearly, the need to develop two instead of one strategy for cracking the alphabetic code would carry an associated cost for learners of deep orthographies. Accordingly, the ODH predicts that the more consistent the mappings, the faster it is to learn that orthography. Indeed, several studies have supported the view that spelling takes more time to be mastered in deep as opposed to transparent orthographies (e.g., Caravolas & Bruck, 1993; Carrillo, Alegria, & Marin, 2012). It should be noted, however, that although orthographic consistency affects *rate* of word-level literacy development, the cognitive and linguistic skills underlying basic spelling and decoding abilities appear to be quite comparable across alphabetic orthographies (e.g., Caravolas, et al., 2012; Landerl & Wimmer, 2008; Vaessen et al., 2010; Ziegler et al., 2010), at least over the foundational period of literacy learning (Caravolas et al., 2012). Finally, orthographic complexity in a more general sense, not only varies as a function of orthographic depth, but also according to specific linguistic parameters, such as syllabic complexity or prosodic patterns (Seymour, et al., 2003; Share, 2008; Treiman, 1993). These other language-based factors, however, appear to become particularly relevant in later stages of spelling development.

In suggesting that radically different strategies (phonological and non-phonological) of reading and spelling resulted from the consistency of phono-graphic mappings, the ODH failed to accommodate the strong evidence for the existence of a phonological component driving literacy development in an overwhelming number of alphabetic orthographies (e.g., Caravolas et al., 2012; Vaessen, et al., 2010; Ziegler et al., 2010). An improvement in that respect was Ziegler and Goswami's (2006) Grain-Size Theory (GST), which also aimed to explain cross-linguistic differences in rates of literacy development, but within a fully phonological framework. According to the GST, children base their learning of spelling and reading on the unit that presents the highest degree of phonology-to-print consistency. For shallow orthographies it is then sufficient to learn the phoneme-to-letter mappings, while more inconsistent orthographies, like English or French, would also need to attend to larger units, thus involving several letters at once. This theory was partly based on the realization that English orthography is much more consistent at units larger than the phoneme and at the level of morphophonology, than at the pure phonemic level. As with the ODH, it follows from this theory that users of deep orthographies would need considerably more time to learn to read and spell than those of more transparent orthographies, since the former would need to learn a large pool of units and their correspondences, in addition to learning phoneme-to-grapheme mappings.

The GST has been challenged on a number of fronts. Firstly, it has been criticized for not successfully integrating the different levels of linguistic analysis and performance. In this way, the GST may have overlooked the implications of proposing an account based on units larger than the phoneme without an adequate consideration of the potential overlaps with other linguistic units with proven psychological reality, such as morphemes (Durgunoğlu, 2006) and, more generally, by neglecting the influence of oral language input in shaping reading and spelling development (Caravolas, 2006). Secondly, it has been pointed out that the GST is only concerned with gains in reading accuracy, while failing to account for reading fluency, which is a critical marker of literacy development once accuracy levels reach ceiling—something likely to occur after a few months of instruction in most transparent orthographies (Wimmer, 2006).

Both the ODH and the GST apply to writing at the word level, that is, to spelling (or word reading). Although related (e.g., Abbott, et al., 2010; Graham et al., 1997), spelling and text composition constitute distinct skills (Abbott & Berninger, 1993). However, considering the central role of spelling skills within current models of writing development (e.g., Juel et al., 1986; Berninger & Amtmann, 2003; Berninger & Swanson, 1994; Berninger & Winn, 2006), the high sensitivity of spelling to variations in orthographic complexity justify the need for cross-linguistic (i.e., cross-orthographic) studies to determine to what extent such variations affect both the writing process and product.

1.5 Spelling Development

Many studies and scholars have been occupied with the nature and development of spelling skills, making it spelling the most thoroughly studied aspect of writing development. Spelling involves writing at the word level, that is, it applies to the within-word domain (Ravid &Tolchinsky, 2002). Although spelling is a low level, more "basic" skill within the larger scope of text composition, it is not exempt from complexities. It has been argued that spelling is a more difficult skill to acquire than its counterpart, word reading (e.g., Bosman & Orden, 1997). This is because spelling a word requires a full phonological and orthographic representation of the word, whereas only a partial representation may result in successful word identification.

The first research efforts to elucidate what spelling entailed were carried out looking at the intricate spelling system of English, in which it was assumed that irregularities were the norm and spelling consisted of rote memorization printed words (Templeton & Morris, 2001). Recognition that English orthography includes regular patterns that are, nevertheless, not always phonological in nature (e.g. Chomsky & Halle, 1968; Read, 1971; Venezky, 1970), led to regard spelling development as a complex process in which different sources of knowledge need to be taken into account. As a consequence, the development of spelling ability has been looked at in terms of the point at which different types of knowledge—and their corresponding strategies—become available to the child, assuming that the development of spelling skill follows a predictable order. Several influential models endorsed this stage-like nature of spelling development in alphabetical writing systems (e.g., Frith, 1985; Gentry, 1982; Henderson & Beers, 1980). In spite of several differences among them, they share a number of characteristics. Firstly, stage models identify an initial or baseline stage at which children seem to be aware of the differences between writing and other types of symbolic representation (e.g., drawing), but are still ignorant of the link between the oral language phonology and the alphabetic code; that is, they assume some kind of pre-phonological state.

Secondly, each phase or stage is discreet in nature, so that mastery of stage 1 is a prerequisite to move on to stage 2, and so on. Finally, and stemming from the last point, transitions from one to the following developmental stage are qualitative in nature, since they imply learning an altogether new strategy or resorting to a different source of knowledge to face spelling tasks.

A paradigmatic stage-like model of spelling development is Frith's (1985) seminal work, which proposed a model of typical reading and spelling development as the basis for comparison with the literacy development of dyslexic populations. Her model identified three main phases in literacy development: *logographic, alphabetic*, and *orthographic*, which occur at different rates in reading as opposed to writing; that is, reading and writing are seen as different, though related, processes, the relative weight of which varies throughout development. The logographic stage is characterized by children's ability to recognize whole words, without yet a full awareness of the phonological structure of words or of the fact that such phonological units are related to letters. Frith points out that the logographic stage already presupposes a wealth of knowledge on the part of the child. She argues that at least a vague notion of what constitutes a word or a sentence, together with an understanding the symbolic nature of written language, is required. This stage is thus characterized by logographic skill being operative, defined as the "instant recognition of familiar words", where "contextual or pragmatic cues" play an important role. Logographic skill is said to be mastered first for reading and then transferred to spelling once a sufficient number of words have been stored to allow children to use those letters to represent words in their own productions. Children later move on to the second phase, which is characterized by the realization that letters represent speech sounds. This stage involves an increasing command of the correspondences between phonemes and graphemes, which Frith termed alphabetic skills. According to this model, English-speaking children begin to understand that letters represent phonemes mainly through spelling, which is said to be the "pacemaker" in the alphabetic phase of literacy development. Finally, once children have mastered the basic sound-letter correspondences of the English writing system, they progress to start using orthographic skills. These entail an again automatic identification of orthographic units without the need to perform a phonological analysis. Orthographic skills would be mastered first for reading and later transferred to spelling, which continues to resort to alphabetic skills for longer. In sum, Frith's (1985) model describes three milestones in a rather U-shaped developmental route, where mastery of a particular skill is a pre-requisite to progress into the next phase, although previous skills remain available to the child if needed.

More recently, a number of investigations (e.g., Treiman, 1993; Pollo, et al., 2008; Rittle-Johnson & Siegler, 1999) have challenged the view that spelling progresses in a step-like manner, and argue that children's spellings reveal an awareness of different types of knowledge already at very early stages of development. Treiman and collaborators (e.g., Cassar & Treiman, 1997; Treiman, 1993; Treiman & Bourassa, 2000; Treiman & Cassar, 1996) showed that children apply orthotactic and morphological knowledge much earlier than described by stage-models of spelling development. In this sense, it has been suggested that use of different spelling strategies and types of linguistic awareness (phonological, morphological, etc.) may be predicted by their overall frequency of occurrence in a particular orthographic system (Pollo, et al., 2008). Spelling patterns may be thus explained from within the framework of implicit learning. From this view, children's early awareness of orthotactic constraints—such as the rules for doubling letters in English —would be the result of certain patterns being found much more often than others in the printed materials to which children are exposed. In other words, a statistically-based account of spelling development would help to explain the heterogeneity of types of knowledge children apply even in their first attempts at spelling, because they would reflect the occurrence of spelling patterns—rather than rules—in the actual input with which children are in contact.

An alternative view of spelling development that advocated for the idea that children do not learn to spell in a stage-like manner is the one proposed by Rittle-Johnson and Siegler (1999). They used the framework of the *overlapping waves* model (Siegler, 1996), which had been first proposed within the domain of the acquisition of mathematical reasoning. The main concept in the overlapping waves model is that development does not happen in stages at which one learns new strategies; rather, strategies are there from the outset, but children use them at different rates and with varying degrees of success throughout the learning process. It is important to note that the term "strategy" is taken to comprise any problem-solving resource that is used when automatic retrieval is not available. A central goal in Rittle-Johnson and Siegler's (1999) paper is the comparison of the acquisition of mathematical reasoning, an analogical domain, with spelling, a non-analogical domain. An analogical domain is one where, when the solution to a problem cannot be retrieved automatically, an efficient application of the right strategy leads to the correct answer. This is certainly the case of mathematical problems. Spelling, on the contrary, is a non-analogical domain, since—especially in the case of English inconsistent orthography—applying a spelling strategy that works in one instance may not necessarily lead to the conventional spelling of another word. For example, if a child has learned that /f/ is spelled with grapheme <f>, applying this spelling rule will not always result in an accurate spelling, as shown for words such as *photograph* or *phone.* The authors compared the spelling performance of children in their first and second year of formal literacy instruction. In spite of this crucial difference between the two domains—mathematical reasoning and (English) spelling—the results revealed that, similarly to mathematics, children in both groups (Year 1 and Year 2) resorted to the same strategies, indicating that more experience with schooling and with spelling itself did not actually introduce brand new ways of spelling words accurately. However, there were important differences, (1) in the extent to which each age group resorted to each strategy, and (2) in the extent to which a particular strategy was used efficiently. In addition, older children outperformed younger ones in their overall success rate at spelling words conventionally. Automatic retrieval increased with age, while the use of the "sounding out" strategy decreased, in favour of other spelling mechanisms. An important limitation of Rittle-Johnson and Siegler's (1999) study is that it included a very restricted age range, which restricts generalizations about the longer-span process of spelling development. Nevertheless, it provides further support to non-stage-like models insofar as they showed evidence of a number of various spelling strategies being available to the child different points in development.

To sum up, models and theories of spelling development depict it as a skill that is challenging to acquire. It is not hard to see how it may cause interference during text composition, especially during the earlier phases of development of writing skill, given that it is *the* component that defines and characterizes (together with handwriting) the written, as opposed to the oral, domain. All models and theories of spelling seem to concur that spelling would be particularly difficult for users of deep orthographies. Learners of English orthography, for example, may need to attend to a multitude of sources of knowledge (phonological, as well as morphological, syntactic, etc.), and units of analysis (phonemes, rimes, and other chunks larger than the phoneme, sometimes including whole-word patterns), or be forced to redefine the functioning of the system several times throughout their learning period, in order to become proficient users of the writing system. Such an additional burden caused by the complexity of the system may unavoidably be transferred to text-level performance. As a consequence, its interference should be evident, either directly or indirectly, for a longer period of time in comparison with the writing development in more transparent orthographies. In other words, considering the critical role of spelling within developmental accounts of writing, it follows that children learning to write in a more complex orthography will need more time to automatize this skill. This delay may have knock-on effects in delaying the development of higher-level skills and other proficiency markers of writing; this pattern should be manifest in a systematic cross-linguistic comparison of writing development. In this sense, cross-linguistic research presents itself as the ideal scenario in which to test some vital assumptions of writing development.

1.6 Research Questions and Overview of the Thesis

The main aim of this thesis is to provide empirical cross-linguistic evidence of the early stages of learning to communicate through writing. The comparison across the two orthographies in question, English and Spanish, thus constitutes the axis of the entire body of work. Having established that a central principle of current thinking about writing development is the fundamental importance attributed to low-level transcription components, comparing two orthographies at opposite ends of the orthographic depth continuum may just be an ideal natural scenario in which to test such assumption. The present studies benefited from the larger research project to which they belong, ELDEL, including access to large samples sizes and a longitudinal design.

The thesis examined microstructural aspects of the written texts produced by English- and Spanish-speaking children. Microstructural features were moreover divided into two main subgroups: word level and text level. The word level comprised text spelling skill, as well as three other aspects of writing development whose effect is mostly confined within this level. These were (1) the specification of conventional word boundaries (i.e., word segmentation), (2) the correct use of upper- and lower-case letters, and (3) marking of morphological information in writing. At the text level of performance, aspects that have been argued to contribute to text "proficiency" (in the sense of Chenoweth & Hayes, 2001) were included. These were measures of vocabulary sophistication, syntactic complexity, and elements that contribute to text structure. In addition, the text level also included productivity measures (number of words and letters).

As mentioned, the guiding research question is to elucidate the role of orthographic consistency in the early development of microstructural features of text composition. Based on the literature reviewed above, the general expectation is that the more consistent Spanish orthography will interfere to a lesser degree than English orthography with other writing processes. In addition, word-spelling skills will be automatized sooner in Spanish than in English, thus Spanish children should have more cognitive resources available to be devoted to other aspects of writing. These two aspects combined—less interference and earlier release of cognitive resources—should result in an advantage of the Spanish group in the proficiency at the word and text levels of written expression.

The remaining chapters of the thesis are organized as follows: Chapter 2 provides an overview of the main orthographic, linguistic, and pedagogical contrasts between the two language groups. Chapter 3, Study 1, explains the characteristics of the larger project that served as the framework for the studies included in the thesis. It also reports detailed descriptive statistics of the full array of microstructural features obtained from the texts, as well as data collection, transcription, and coding criteria. Chapters 4, 5, and 6, make specific research questions about the role of orthographic consistency in early writing development,

- Chapter 4-Study 2: How does the early development of microstructural writing features differ as a function of orthographic consistency?
    - Does orthographic consistency affect the emergence and early development of word-level and text-level writing features?

    - Do word- and text-level features relate to similar or different writing components across orthographies with differing degrees of consistency?
- Chapter 5-Study 3: Is spelling the bottleneck of writing development?
    - Does spelling mediate the development of word-level writing skills?
    - Does spelling mediate the development of text-level writing skills?
- Chapter 6-Study 4: What are the cognitive precursors of early writing development?
    a. What are the cognitive predictors of Year 1-Year 2 text spelling accuracy?
    b. What are the cognitive predictors of Year 1-Year 2 text-level writing proficiency?
    c. Are there common or language-specific predictors of early writing development?

The aim of this body of work was not to validate or reject the writing models *per se*, but to test the cross-linguistic implications that arise from their assumptions and formulations. Accordingly, the predictions for the research questions in Study 2, were that English participants will show a poorer performance and slower development in both word-level and text-level features. A slower development in word-level features in English was expected because word level literacy (reading and spelling) has been found to be highly sensitive to variations in orthographic consistency, with reading/spelling skills in consistent orthographies developing faster than in inconsistent ones (e.g., Caravolas & Bruck, 1993; Caravolas et al., 2013; Seymour, et al., 2003). Although word spelling and word-level *text* writing features are not identical, they are arguably closely related constructs. A slower development in text-level writing was expected for the English group because the higher demands of the more inconsistent orthography would prevent English children from attending to factors other than transcription to a greater extent than their Spanish peers.

The prediction for the research questions in Study 3 was that single word spelling skill would be a more powerful mediator in the development of writing skills in English than in Spanish. It was reasoned that, if spelling skills constrain writing development, a more complex orthography should show a larger dependence upon spelling skills than a more consistent one. In addition, across orthographies, it was reasoned that single word spelling should be a more powerful mediator of word-level text writing skills, than of text-level writing features.

Finally, Study 4 explored the cognitive predictors of writing skills, from preschool until mid-Year 2. It was hypothesized that (1) at the earlier stages of development (mid-Year 1) the same pattern of predictor variables should underlie writing development in both language groups. This assumption is based on recent findings providing robust evidence of common predictors across a number of orthographies with varying degrees of consistency (Caravolas, et al., 2012; 2013). However, (2) once Spanish children have relatively consolidated their basic spelling/encoding skills, the pattern of prediction should differ from one language group to the other; specifically, groups should particularly differ in the cognitive foundations of text-level writing skill.

Chapter 2

Relevant linguistic, orthographic, and pedagogical aspects

The purpose of the present chapter is to provide a comparison of orthographic and linguistic factors that may affect the rate and course of early writing development in English and Spanish. In addition, because several aspects of writing are sensitive to teaching methods, a summary of the results of a series of questionnaires administered to the teachers of all participating schools is provided. This comparative picture of the orthographic, linguistic, and pedagogical domains is offered not as an end in itself, but as a tool to aid in the interpretation of the results of ensuing empirical studies. The exploration showed that English orthography is much more inconsistent in its phonographic mappings than Spanish orthography. Also, the two languages exhibit relevant linguistic contrasts. In the phonological domain, English shows a more complex syllable structure and a high number of vowels as well as vowel reduction processes. From a morphological perspective, the rich morphological system of Spanish may arguably attune children from earlier on to use morphological information during writing. From a syntactic perspective, the differences between languages may possibly bias certain measures of writing performance. The report of Reception Year teaching practices revealed no differences in the relative weight of different literacy tasks across countries, though some significant differences were attested in the Year 1 and Year 2 practices. Such differences, however, did not involve the teaching of writing. Finally, teachers in both countries had similar expectations for the levels of attainment for higher-level literacy skills (e.g., text comprehension and production), but the Spanish teachers had higher expectations for the levels of attainment in word spelling and reading.

2.1 English and Spanish Orthography

2.1.1 General Characteristics

The English spelling system is typically described as consisting of 26 letters that are used to represent 24 consonant phonemes, 12 vowels, and 8 diphthongs (Received Pronunciation (RP); Roach, 2004). In Spanish there are 5 vowel phonemes, and 18 consonant phonemes (Quilis, 2010), which are to be represented by 27 letters—the same 26 of English, plus *ñ*, which corresponds to /ɲ/, a palatal nasal sound. There is a much larger disparity in the number of letters and phonemes in English, pointing to the higher complexity of its phoneme-grapheme mappings, in contrast to the quasi regularity of Spanish orthography. There are countless examples of how inconsistently the majority of English phonemes are represented by letters (see Carney, 1994, Chapter 3, for a thorough description of the English spelling system). The labio-dental, voiceless fricative, /f/, for instance, may be represented by <f>, as in *fish*, by <ph>, as in *photograph*, by <ff>, as in *coffee*, or by <gh>, as in *rough.* Vowel sounds are usually regarded as being even more inconsistent than consonants. Possible orthographic representations of the vowel /i:/, for example, include *m**e**, s**ea**, **see**, p**eo**ple, k**ey**, sk**i**, conc**e**d**e**, mach**i**n**e***, etc.

Although the number of inconsistent mappings, as well as the number of possible pairings, is considerably lower in Spanish, it is not exempt from phonographic ambiguities. Spanish /b/, for example, may be represented by <b>, as in *barco* 'ship', or by <v>, as in *vela* 'candle'; /k/ is spelled <c>, as in *casa* 'house' or *conejo* 'rabbit', but as <qu> when it is followed by a front vowel, as in *queso,* /'keso/, 'cheese', or *quinto,* /'kinto/, 'fifth'. Similarly, /θ/ may be represented by <z> or <c>, but <z> only occurs if followed by vowels /a, o, u/, while <c> must be followed by a front vowel, /e/ or /i/ *(Ortografía de la Lengua Española,* 2010). Vowels, on the other hand, are completely consistent. In short, there are important differences in the degree of consistency of the phonographic mappings in English and Spanish orthography. Proposals for quantifying consistency are reviewed in Section 2.1.2 below.

English and Spanish also differ in how they represent suprasegmental aspects, especially word stress. Both languages have free stress (Cutler, 2012); that is, any syllable may carry primary stress in the word, in contrast to languages with fixed stress, like French. Both Spanish and English show a preference for trochaic stress (Pons & Bosch, 2010); that is, usually the first syllable in a two-syllable word is the strong one. However, while Spanish three-syllable words are typically stressed in the penultimate (or medial) syllable, English shows a higher proportion of words with initial stress (Cutler & Carter, 1987; Pons & Bosch, 2010). In sum, both languages have roughly similar stress patterns, except for the larger preference in English for initial syllable stress. However, English spelling does not represent word stress in any way, while Spanish uses stress marks systematically to signal the stressed syllable in multisyllabic words. Monosyllabic words do not, as a rule, carry stress marks, although there are some homophones where one of the members of the pair takes a "diacritic stress mark" to distinguish it from the other member. Typically, diacritic stress marks are used when at least one of the homophones is a function word. For example, the word /si/ is spelled *sí* 'yes' when it refers to the affirmative adverb, but it is spelled *si* 'if', without the stress mark, when it stands for the conditional conjunction. Similarly, *té* means 'tea', but *te* refers to the second-person singular object pronoun.

Spanish uses dieresis, while this orthographic sign is not used in English. Dieresis is used in Spanish to indicate that, in the sequences <güe> and <güi>, the <u> is not silent, as it normally is. Phoneme /g/ in Spanish is represented <g> when followed by /a, o, u/, as in *gorro* 'cap'or *gusano* 'worm', or when it is followed by /ɾ, l/, if it occurs as the first consonant in a consonant cluster, as in *grande* 'big' or *globo* 'balloon'. When followed by a front vowel, however, /g/ is represented by the grapheme <gu->, as in *guerra,* /ˈgera/, 'war', and *guinda,* /ˈginda/, 'cherry'. This is because the combinations of <ge> and <gi> correspond to the velar fricative /x/ as in *genio,* /ˈxenjo/, or *girasol,* /xiɾaˈsol/. Therefore, words combining /g/+/u/+front vowel are written with <ü>, such as *pingüino* 'penguin' and *cigüeña* 'stork'. Without the dieresis, these words would read as /*pinˈgino/ and /*siˈgeɲa/, instead of the correct /pinˈgwino/ and /siˈgweɲa/, respectively[1].

Both languages make similar uses of punctuation marks, such as full-stops, commas, colons, semi-colons, etc. Question marks and exclamation points are used in

[1] Note that the Spanish dieresis only affects phoneme /w/, which is considered a semi-vowel (Quilis, 2010), rather than a semi-consonant, as in English.

functionally similar ways, but in Spanish they must appear at the beginning and end of the question or exclamation. Apostrophes are not used in Spanish, while they are an integral part of English spelling. Other contrasts in the use of punctuation marks may include some uses of the hyphen. Spanish uses a dash or hyphen, instead of inverted commas, to indicate direct speech. Also, Spanish uses hyphens quite frequently to indicate, at the end of a line, that a word continues in the following line. Surely other stylistic contrasts exist, but they are arguably not relevant to this work, considering that it deals with the very foundational stages of writing development, so children would be, for the most part, unaware of most of such conventions.

2.1.2 Orthographic Consistency

English is usually regarded as one of the most complex of all alphabetic writing systems (Share, 2008), while Spanish is usually regarded as a highly regular one (Alegría, Marín, Carrillo, & Mousty, 2003). The large number of alternative spellings for almost any phoneme has driven some to label English orthography as “chaotic” (Dewey, 1971, cited in Kessler & Treiman, 2003, p. 267). Kessler and Treiman (2003) provide a list of reasons why English spelling shows such a high degree of ambiguous phonographic mappings. One of the reasons is that established spellings remain so even if changes in pronunciation occur or if they do not fit certain dialectal variations. This is a problem for beginner spellers and readers, but it guarantees intelligibility across a wide range of English speakers. This “conservatism” (Kessler & Treiman, 2003) of spelling is shared with Spanish, which has not altered its spelling rules in spite of the growing demand for reforms to accommodate the ever increasing number of speakers of different varieties of Spanish.

A second reason for the opacity of English spelling is that loan words are usually borrowed without changing their original spelling. In contrast, the reference academy for Spanish oral and written language, *Real Academia Española,* is constantly revising new borrowings, trying to determine the best ways to adjust their spelling so that they conform to Spanish phonotactic and graphotactic constraints. For example, it is considered “improper” in Spanish spelling to use the sequence <qu-> to represent any sound other than /k/ followed by a front vowel, as in *queso,* /'keso/, ‘cheese’ and *quilo,* /'kilo/, ‘kilogram’. Therefore, Latinate words like *quorum* or proper nouns like *Qatar* are to be spelled *cuórum* and *Catar,* respectively. Similarly, it is recommended that direct loans be signalled with inverted commas or italics, so either

<*paddle/ballet*> or <pádel/balé> are considered acceptable renderings, but not <paddle/ballet> (*Diccionario Panhispánico de Dudas,* 2005).

A third reason accounting for inconsistencies in English orthography is the fact that spelling conveys information beyond the phonemic level. Indeed, lexical and etymological, as well as morphological information, many times take precedence over phonographic transparency. By favouring lexical over phonological information, spelling disambiguates many homophones of spoken English, such as *night/knight, buy/bye/by, right/write/rite,* etc. Also, the root-consistency principle preserves word roots unaltered in spelling, ignoring changes in pronunciation, but signalling quite clearly whether two words are related or not. Therefore, in keeping <c> in the spellings of both *medicine*, where <c> is pronounced [s], and *medical,* where it is pronounced [k], the relationship between the two words is evident. Conversely, from a reader perspective, the words *night* and *knightly* are easily seen as non-related (Chomsky, 1970; Kessler & Treiman, 2003, p. 270). In Spanish, in contrast, phonological transparency prevails, so that the link between *proteger,* /pɾote'xeɾ/, 'protect_INF' and *protejo*, /pɾo'texo/, 'protect_1SG_PRES_SUBJ', is more evident orally than in writing.

In spite of its "chaotic" appearance and regardless of the benefits and disadvantages of English orthography, Kessler and Treiman (2003) insist in that many assessments of the randomness of English phonographic mappings have been considerably biased. They point, especially, at the lack of regard for the position and, more generally, the phonetic environment in which many "ambiguous" phonemes occur. Taking positional and contextual constraints into account drastically reduces the number of alternative spellings for any given phoneme.

In order to provide more objective accounts of the relative transparency/opacity of different orthographies, different measurements have been proposed to quantify the degree of consistency of phonographic mappings, whether in the phoneme-to-grapheme direction (i.e., encoding) or in the grapheme-to-phoneme direction (i.e., decoding). For example, Borgwaldt, Hellwig, and De Groot (2004) performed entropy calculations of the word onsets of five alphabetic orthographies: Dutch, English, French, German, and Hungarian. These calculations involved computing the deviation from the 1:1 ideal phoneme-to-grapheme mapping (as well as the reverse direction). Therefore, the higher the entropy level, the more inconsistent the orthography. The authors concluded that English was the orthography with the highest deviation from the 1:1 ideal out of the five languages studied, in both the decoding and encoding directions. Caravolas et al. (2012) also quantified the level of inconsistency in four alphabetic orthographies: English, Spanish, Czech, and Slovak. Importantly, their calculations were not limited to word onsets, but phonemes and letters in any position were considered. Specifically, they divided the frequency of occurrence of a particular phonographic mapping by the overall frequency of occurrence of the phoneme or grapheme (depending on whether the calculation was for spelling or reading, respectively). Therefore, these consistency estimates ranged from 0-1, with higher values standing for high consistency levels in a given orthography. The consistency estimates obtained for English and Spanish were .72 and .96, respectively, for reading, and .62 and .90, respectively, for spelling (Caravolas et al., 2012, p. 679).

Even assuming that the irregularity within English spelling has been exaggerated, by all accounts English orthography is much more inconsistent for reading and, particularly, for spelling. In contrast, Spanish presents a quasi-regular orthography in the reading direction, while it is only slightly more inconsistent in the spelling direction. However, inconsistency of phonographic mappings is not the only source of complexity of written language, but the characteristics of spoken language have been argued to also impact on literacy development processes (e.g., Caravolas & Bruck, 1993; Caravolas, Bruck & Genesee, 2003; Caravolas & Landerl, 2010).

2.2 Relevant Linguistic Features of English and Spanish

Learning to write drives children to discover which elements of the spoken modality are transferred to writing and how the written code represents them. Indeed, all levels of linguistic awareness must be put into play over the course of learning to write (Ferreiro, 2009). To a certain extent, the features to be transposed, and the manner in which they are transposed, will be dependent on the phonological, morphological, syntactic, and discursive characteristics of the language to be represented. It is therefore important that, in a cross-linguistic study, relevant linguistic contrasts are considered, particularly those that may have an impact on the difficulty of the learning process, as well as on its characteristics. In what follows, the most salient contrasts between Spanish and English oral language features at the phonological, morphological, and syntactic levels are described, commenting on their potential to shape each language route towards writing development and to anticipate differences as a function of language.

2.2.1 Relevant Phonological Features

The phonological structure of oral language has been shown to influence the development of phoneme awareness, which is one of the chief predictors of later spelling skills in alphabetic orthographies (e.g., Caravolas et al., 2001; 2012; Ziegler et al, 2010). One of the most relevant phonological traits that appears to have an impact in phoneme awareness skills is syllabic complexity (Caravolas & Bruck, 1993; Carvolas & Landerl, 2010; Cossu, Shankweiler, Liberman, Katz, & Tola, 1988; Duncan, Colé, Seymour, & Magnan, 2006; Jiménez González & Haro García, 1995; Manrique & Signorini, 1994; Sprenger-Charolles & Siegler, 1997). Since children show difficulty in segmenting, as well as in spelling, consonant clusters (Bruck & Treiman, 1993), languages where clusters are abundant, and in which consonants are allowed to recombine in many ways to produce a large inventory of cluster types, may provide children with more opportunities to develop an early sensitivity to the constituent phonemes within these complex phonological structures. In this sense, syllables in English tend to present considerably more complexity than syllables in Spanish. Two- three-, and even four-consonant clusters are allowed in certain positions in English, while Spanish allows a maximum of one or two, depending on the position within the word. A maximum of two consonants may occur in word-initial position, such as *premio,* /ˈ**pɾ**e.mjo/, ‘prize’, or *clavo,*/ˈ**kl**a.bo/, ‘nail’; or in syllable coda, but only if the cluster is word-internal, as in *extraño,* /e**ks**.ˈtra.ɲo/, ‘stranger’, or *constant,* /ko**ns**.ˈtan.te/. Only one consonant is allowed in word final position. In contrast, syllable onsets in English may contain clusters of up to three consonants, which take the form /s/+voiceless stop+liquid consonant (i.e., /r/ or /l/) or approximant (i.e., /j/ or /w/; Roach, 1998). Examples of three-consonant onsets are *spring,* /**spr**ɪŋ/, *split,* /**spl**ɪt/, or *skew,* /**skj**u:/. Clusters of up to four consonants may be found in English in coda position, such as *prompts,* /prɒ**mpts**/, or *texts,* /te**ksts**/.

In addition to the number and variety of clusters, the position at which consonants and consonant clusters occur also contributes to the syllabic complexity of a language. Open syllables, that is, those with no consonant in the syllable coda, are less challenging than closed syllables—i.e., those whose rime includes at least one consonant. From a broad cross-linguistic viewpoint, closed syllables are “marked” whereas CV syllables are assumed to be present universally (Vennemann, 1988). Children speaking languages in which open syllables predominate (e.g., Czech) find it relatively difficult to spell and to orally segment consonants in coda position in contrast to children speaking languages with frequently occurring syllable codas (e.g., English, German; Bruck & Treiman, 1990; Caravolas & Landerl, 2010; Read, 1986). Again, English presents a more complex picture than Spanish in that closed syllables are quite frequent, while they are very uncommon in Spanish. Specifically, CV syllables account for more than 50% of all syllable types in Spanish (e.g., Álvarez, Carreiras, De Vega, 1992; Guirao & Manrique de Borzone, 1975; Justicia, Santiago, Palma, Huertas, & Gutiérrez, 1996; Moreno Sandoval, Torre Toledano, Curto, & de la Torre, 2006).

The number of vowel phonemes and the degree to which they are easy to identify has been claimed to affect the development of spelling and phoneme awareness skills. Spanish has a limited set of only five vowel sounds, which are perceptually stable and salient, as well as unequivocally represented in writing (Manrique de Borzone & Signorini, 1994). RP English, in contrast, presents around 20 vocalic sounds (Roach, 2004). The perceptual saliency of this large inventory of English vowels is, moreover, dependent on several factors, such as speech rate, syllable stress, number of syllables in the word, among others (Roach, 2004). Finally, vowel sounds are very inconsistently represented in writing, particularly when they occur in unstressed syllables, because of vowel reduction processes (Giegerich, 1992, p. 68-69). Not surprisingly, most spelling errors of English-speaking children involve vowels (Treiman, 1991).

English and Spanish also contrast in suprasegmental phonological aspects. Specifically, they belong to distinct speech-rhythm typologies: Spanish is, predominantly, a syllable-timed language, while English is a prototypically stress-timed language. In syllable-timed languages, syllable duration is quite regular but stress pulses are irregularly spaced, while stress-timed languages are characterized by having evenly spaced intervals between two stressed syllables, at the expense of unequal syllable duration (Abercrombie, 1967; Roach, 1982). The result of the typological distinction is the perceptual saliency of syllables in syllable-timed languages, as opposed to stressed-timed languages, which need to resort to other cues (see Jusczyk, 1999, for a review). The perceptual saliency and regularity of syllables in Spanish may be an advantage in that it may provide children with a stable correspondence unit for spelling. Indeed, the view that children resort to syllables as an early correspondence unit has been quite influential, especially in developmental studies of spelling in Romance languages (e.g., Ferreiro & Teberosky, 1979; Pontecorvo & Zucchermaglio, 1988; Tolchinsky, 2003; Vernón, 1993). It should be noted, however, that the theory of a syllabic period in spelling development has received considerable criticism (e.g. Cardoso-Martins, Correa, Lemos, & Napoleao, 2006; Pollo, Kessler, & Treiman, 2005).

In sum, different phonological variables shape language development and it is their combined effects that need to be considered to determine the facilitative effects in learning to spell/write. Although English children may more readily develop sensitivity for complex clusters (given that they are more abundant and varied in English than in Spanish), the high number of vowel sounds and its closed syllable structure are factors that affect English negatively much more than Spanish.

2.2.2 Relevant Morphological Features

English is a language with very few overt inflectional morphological markings in both nouns and verbs, while Spanish is comparatively richer in terms of verb morphology (Dressler, 2005, 2010). This typological distinction seems to be developmentally relevant, given that it has been shown that children acquire the morphology of their L1 faster if the language in question is morphologically rich (Dressler, 2010). Moreover, the productive use of bound morphology, with relatively transparent meaning, may arguably sensitize speakers of morphologically rich languages to this type of information; that is, the information conveyed by inflectional morphology may be more meaningful for them than to speakers of languages that are morphologically poor (Wijnen, Kempen, & Gillis, 2006). Therefore, Spanish children would be more sensitive to morphological information than English children. This is a crucial point, given that morphological knowledge has been claimed to constitute an important source of information to be able to spell accurately, especially when phonographic rules are insufficient to arrive at conventional spellings (e.g., Deacon & Kirby, 2004; Treiman & Bourassa, 2000; Treiman & Cassar, 1997). Certainly, morphological awareness is useful to decide between phonologically plausible spellings, as is the case of the spelling of regular past tense and participle verb forms in English. The past tense morpheme /-d/ may be realized as [-t], [-d], and [-ɪd] depending on linguistic context. In spite of the variety of allophonic variants, these are invariantly represented in writing as <-ed>. Therefore, knowing whether *past* or *passed* is the accurate spelling for the sequence [pɑ:st] requires (1) access to the corresponding morphological representation, and (2) knowledge of how the morpheme is represented in writing. The same strategy is useful in Spanish, where some verb forms— such as the imperfect past tense of verbs ending in *–ar*, e.g., *amaban,* 'loved_3PL'—, present inconsistencies that are easily resolved by applying knowledge of how the orthography represents that particular morpheme. Applying only phonographic criteria, the morpheme /-aba/ could be represented as either <-aba> or <-ava>, since phoneme /b/ is ambiguous in its mapping to spelling. However, /-aba/ is always written with <b> when representing the imperfect past-tense morpheme. In addition, morphological awareness is, perhaps, the only strategy that yields correct use of stress marks in certain pairs of monosyllabic words (diacritic stress marks; see Section 2.1.1 above).

Since access to morphological representations may be crucial to arrive at the correct spelling of phonographically inconsistent words, it is vital to elucidate which factors determine sensitivity and access to this information and, most importantly, if and when such access becomes available to writers. Although the children who participated in the studies of this thesis may be assumed to have already acquired (or be in the very last stages of acquisition of) their corresponding inflectional morphology systems, it may be argued that Spanish children would more readily use morphological information than their English counterparts.

Evidence that children learning to spell in languages with rich morphology more readily resort to this type of information to produce accurate spellings has been provided by a cross-linguistic study comparing Dutch, a language with poor morphology, with Hebrew, a highly inflected language (Gillis & Ravid, 2000; 2006). In line with the morphological-richness derived hypothesis, Hebrew children resorted to morphological knowledge as a strategy during a spelling task more often than Dutch children. It is noteworthy that this particular study did not take into account the role that consistency at the phonographic level might have had in determining the overall use of morphological information in spelling. Arguably, Hebrew children did not only resort to morphological representations to aid in the spelling of inconsistent words because Hebrew has a richer morphology, but also because its orthography is less transparent than that of Dutch. One study with Spanish children from grades 1 to 6 showed that they make use of morphological information to resolve inconsistent spellings (Defior et al., 2008). Therefore, these Spanish-based findings would give further support to Gillis and Ravid's study (2000; 2006), by presenting evidence of a richly inflected language which, in contrast to Hebrew, has a very transparent orthography.

Another contrast between English and Spanish that is partly derived from morphological typology is the average number of phonemes (and letters) in each language. This characteristic of spoken and written language may affect and even distort comparisons based on the average number of letters produced in a text writing task, an indicator of word sophistication (e.g., Riedemann, 1996; Llauradó & Tolchinsky, in press). Although English has a more varied range of possible syllable structures (see Section 2.2.1 above), which allow for complex onsets and codas to a much larger extent than in Spanish syllables, words in English tend to be monosyllabic, while in Spanish monosyllabic words are usually confined to function words and a few content words. Spanish adjectives and many nouns with gender alternations contain at least overt gender inflection, without which they would be unpronounceable (that is, they would end in a phonotactically illegal sequence), e.g., *bonit-o/a,* 'pretty_masc/fem'; *alt-o/a*, 'tall_masc/fem'; novi-*o/a,* 'groom/bride'; *niñ-o/a,* 'boy/girl'. Most Spanish words are thus multimorphemic, a feature that often results in multisyllabicity, especially when derivational processes are applied. While derivational processes are also available and productive in English, young children use them less if other word-formation processes are available, particularly, compounding (Clark, 1995). Indeed, the average word length in Spanish has been calculated to be of 8.3 phonemes, in contrast with the average of 6.94 in English (Cutler, 2012; Cutler, Norris, & Sebastián-Gallés, 2004). The difference remains stable even after controlling for the effects of frequency of use (Cutler, 2012, p. 49).

2.2.3 Relevant Syntactic Features

English and Spanish also present important syntactic contrasts. As mentioned in the preceding section, Spanish has a complex system of verb conjugations. Verb inflections give information about person (1st, 2nd, 3rd), number (singular, plural), tense (present, past, future), aspect (perfect, imperfect), and mood (indicative, subjunctive, imperative)[2]. There are no bare forms, since even non-finite forms have their corresponding inflection. The abundant information in verbal suffixes allows for the omission of overt subject pronouns, since this information is retrievable from the verbal inflection. This is known as *pro-dropping* (from "pronoun drop"), that is, subject pronouns may be dropped, since they are inferable from context. Compare, for example, the sentence below (2.2a), in Spanish, with its literal translation into English (2.2b),

2.1a	*Cuando llamé,*	*estaban*	*durmiendo.*
2.1b	*When called_1SG	were_3PL	sleeping.
	When (I) called (they) were sleeping.		

Spanish, then, frequently omits subject pronouns, both in main and subordinate clauses. In English this results in the ungrammaticality of the construction, although in certain contexts (e.g., imperatives, informal speech), subjects may also be dropped. The fact that subject pronouns are of obligatory expression in English, but not in Spanish, may have an impact on measures of lexical diversity, given that the same token (e.g., *I, he, she, we,* etc.) is likely to be repeated several times.

On the other hand, English and Spanish are usually quoted as examples of two distinct typological categories: satellite- and verb-framed languages, respectively (Slobin, 1996, 2004; Talmy, 1985). The distinction relates to the way in which events are usually coded in lexical verbs, and it must be regarded as a matter of predominance (and not absolute categories). Satellite-framed languages, like English, typically make use of an adverb or preposition to indicate the direction (*path*) of motion verbs. Most verbs in verb-framed languages, like Spanish, already include directionality as part of the semantics of the verb. Compare, for example, the following sentences,

---

[2] Note that the information conveyed by verbal suffixes is often fused into a single morpheme. For example, in the verb *trabajé*, '(I) worked', the morpheme –*e* indicates 1st person, singular, past, perfect, indicative, all at once.

2.2a *She danced into the room*

2.2b Ella entró bailando a la habitación

'She entered dancing to the room'

In sentence 2.2b, in Spanish, the verb meaning already reflects the direction of the motion, but in 2.2a, in English, preposition *into* specifies that information. Note, however, that Spanish (and verb-framed languages in general) needs to make use of a gerund (or another type of adverbial phrase) to express the *manner* in which the action takes place (*bailando* 'dancing'), while this information is already encoded in the semantics of the English verb. The possible consequence of this contrast is the influence it may have on measures of vocabulary use that take into account specific grammatical categories. For example, measures of adverb and or adjective use may be slightly biased as a result: Other things being equal, English would need to make use of more adverbs and prepositions to indicate the direction of motion verbs, while Spanish would need to make use of adverbs of manner and other adverbial expressions.

2.3 Teaching Practices in England and Spain

In contrast to oral language, the development of writing and, in general, of literacy, occurs in a formal education context. Therefore, it was important to carry out a survey of the literacy practices taking place in the classrooms the participants in all these studies were attending. Note, however, that it was not an objective of this work to manipulate the effect of teaching practices. The main goal of reporting the literacy practices that were taking place in the participating schools was to illuminate interpretations of the results obtained in ensuing chapters.

Literacy instruction in the UK starts at the outset of obligatory schooling—Reception Year—, when children are around 4 years of age. The National Literacy Curriculum (NLC) is the text of reference for the teaching of literacy across the UK. All teachers in the schools which the participants were attending adhered to the NLC. In Spain, the preschool period is divided into two stages of three years of duration each (from 0-3 years old, and from 3-6 years old), which are of optional attendance. The national government decides the general goals and achievement levels with regards to literacy instruction at preschool and primary-school level, but individual regions in Spain decide on the specific tone and methodologies to be implemented (Tolchinsky, Bigas, & Barragan, 2010). In Andalusia, the region in which the Spanish data were collected, the approach to literacy instruction in the preschool years has been defined as predominantly "functional", with teachers being advised to teach reading and writing skills in the context of meaningful social situations (Tolchinsky et al., 2010). At the outset of primary school, there seems to be a growing concern for orthographic knowledge and accuracy in reading and spelling, but preserving the general attitude of embedding literacy instruction within a broader social/discursive context (Tolchinsky et al., 2010).

2.3.1 The Questionnaires

Once each year, the teachers of all participating classrooms in the ELDEL project were invited to complete a questionnaire on their literacy teaching practices. Table 2.1 shows the number of schools and classrooms that returned their questionnaires each year in each year in Spain and the UK. Children in the UK

Table 2.1

Year 1 Teacher Practices — Summary of data sources

	UK			Spain		
	Schools	Classrooms	*n*	Schools	Classrooms	*n*
Reception	9	10	12	5	11	11
Year 1	9	10	10	6	12	12
Year 2	9	10	6	6	12	12

Note. *n* = number of questionnaires returned

typically attend the same school during, at least, the first three years of obligatory education—from Reception Year until Year 2. In Spain children might attend an infants' school for the period before Year 1 (the period up until and including the Spanish equivalent to Reception Year) and then switch to a primary school. That is why an additional school was included in the project in Spain, while in the UK the same 9 schools participated throughout the duration of the project. Another difference that should be pointed out is the size of the schools and the number of classrooms within each one. All but one of the UK schools that participated in the project only had one Reception Year/Year 1/Year 2 class, whereas the schools in Granada that took part of the project usually had two classes per school grade. In addition, although all teachers who were involved in literacy activities were asked to complete the questionnaire individually, teachers in the UK who worked at the same school usually completed them reporting the exact same information—probably suggesting that they were completing them together—, so only one questionnaire per classroom ans year was included in the present report.

The questionnaire was designed to enquire about (1) which literacy skills were taught during the year; (2) the time dedicated to literacy-related activities, such as phonics instruction, storytelling, writing, and reading; (3) and teachers' beliefs on the expected levels of achievement for different literacy skills, distinguishing between word- and text-level skills. Towards the end of each school year, the teachers were approached by the researcher or one of the research assistants in the UK and in Spain, and were invited to complete the questionnaire. No specific instructions were given, except to clarify that the general goal of the questionnaire was to survey their everyday literacy-related activities, in order to be able to better interpret the results from the assessments with children. Teachers were typically given up to about a week to return the questionnaire which they always completed without the presence of the researcher or of any member of the research project.

A copy of the Reception Year and the Year 1 questionnaires is included in Appendix 2. The Year 1 and Year 2 questionnaires were identical. Slight modifications were made to the Reception Year questionnaire before administering it again when the children were in Year 1. These modifications were mostly motivated by the need to adjust to changes in the curriculum in the transition from one year to another. A second motivation for the changes to the Year 1 questionnaire was to make it more suitable to capture quantitative differences in teachers' reported time dedicated to literacy activities, given that no significant differences were found as a function of school in the various literacy practices (see Table A3.1 in Appendix 3 for the results of these analyses). Therefore, while the Reception Year questionnaire contained *ranges* of time devoted to literacy teaching, the Year 1 and Year 2 questionnaire let the teachers introduce the average amount of time (usually in minutes per day or week) they typically dedicate to them. However, non-parametric tests within each language revealed no significant differences in Year 1 or Year 2 literacy practices either (Table A3.1, in Appendix 3).

To assess differences in teaching practices across countries, non-parametric tests were carried out, given the small sample size and the categorical nature of some items, particularly in the Reception Year questionnaire. To allow for comparability across language groups, all times reported by teachers as dedicated to the various literacy-related activities in the Year 1 and Year 2 questionnaires were transformed to minutes per day. Differences in the way individual teachers quantified their time on various tasks made it necessary to convert their raw time estimations in to percentages.

2.3.2 Report of Teaching Practices in Participating Schools

All classrooms in both the UK and Spain used a phonics-based method for teaching literacy (Table 2.2). Spanish and English teachers dedicated a similar amount of time to phonics instruction in Reception Year, U = 44.5, z = -0.44, p > .05; Year 1, U = 40.00, z = -0.04, p > .05; and Year 2, U = 10, z = -1.67, p > .05. During Reception Year, teachers in both countries estimated that time devoted to explicit phonics instruction amounted to at least 75 minutes per week—distributed into 15-minute daily sessions. Throughout the study all UK schools reported using one or a combination of the following methods for teaching phonics: *Progression in Phonics, L-S Programme,* and *Jolly-Phonics.* The most frequent combination, however, was to use *Progression...* and *L-S Programme.* In Spain, teachers were divided between those who adopted a mixed method for teaching phonics, and those who used a global approach.

Storytelling was carried out daily in all UK schools from Reception Year to Year 2. About half of the Spanish classrooms consulted in Reception Year reported also having a storytelling time every day, while the other half reported doing so about once or twice a week. This difference, however, was not statistically significant, U = 27.00, z = -2.28, p > .05. During Year 1 and Year 2, the average time devoted to storytelling was higher for the UK schools, who reported spending an average of 102 and 121 minutes per day in Year 1 and Year 2, respectively. In contrast, Spanish schools reported devoting an average of 18 and 17 minutes during the same periods. This difference was statistically significant in Year 1, U = 6.00, z = -3.31, p = .001, and Year 2, U = 0.00, z = -2.79, p = .005. It should be noted that responses to this item in the Year 1 and Year 2 questionnaires by the English teachers ranged from 40 to an exorbitant 225 minutes per day, while the Spanish teachers replies fitted within a much more restricted 5-30 minutes per day range.

Spanish schools devoted more time to reading activities (Table 2.3) in Year 1, $U = 0.00$, $z = -3.84$, $p < .001$; and Year 2, $U = 0.00$, $z = -3.26$, $p = .001$. The difference was non-significant in Reception Year, $U = 33.00$, $z = -1.34$, $p > .05$[3]. In Reception Year all Spanish classrooms—but not all UK schools—asked children to read aloud to the teacher or the class. Those schools which incorporated this practice did so often in both countries: more than once each week. Reading homework was assigned regularly in both Spain and the UK. Although some UK schools did not give reading homework explicitly, teachers explained that there was a general understanding with parents that children were to carry out reading activities over the weekend, holidays, or any leisure time.

Of special interest were those questions inquiring about the type and frequency of writing activities carried out in the classroom (Table 2.4). During Reception Year, all classrooms in the UK and Spain devoted around 15 minutes every day to letter formation activities. Spanish teachers continued teaching this skill until the end of the study period. In the UK, however, a few schools did not spend time with letter formation tasks in Year 1 and Year 2. Both groups of teachers included free and guided writing activities as early as Reception Year, and continued to do so in Years 1 and 2. Spanish teachers tended to give writing homework much more often than English teachers. Spelling instruction was also carried out by teachers in both countries. Overall, the percentage of time devoted to writing instruction did not differ across languages in Reception Year, $U = 34.50$, $z = -1.17$, p > .05; Year 1, $U = 43.00$, $z = -0.78$, p > .05; or Year 2, $U = 14.00$, $z = -1.74$, p > .05.

---

[3] Difference testing was calculated for the guided-reading item, because it was the only quantifiable one.

Table 2.2

Summary of teachers' responses about phonics instruction

	Reception	Year 1	Year 2
		UK	
Time	At least 15 minutes/day	*M:* 117'/day (min-max: 60-225)	*M:* 85'/day (min-max: 60-100)
Method	All schools used some combination of *Progression in phonics*, *L-S programme*, and *Jolly Phonics*	All schools used some combination of *Progression in phonics*, *L-S programme*, and *Jolly Phonics*	All schools used some combination of *Progression in phonics*, *L-S programme*, and *Jolly Phonics*
		Spain	
Time	At least, 15 minutes/day	*M:* 73'/day (min-max: 25-120)	*M:* 28'/day (min-max: 0-60)
Method	Phonetic or mixed global and phonetic	Mixed (6 class.) or syllable-based (4 class.)	Mixed (6 class.) or global (2 class.)

Table 2.3

Summary of teachers' responses about reading activities

	Reception	Year 1	Year 2
		UK	
Reading aloud	5/9 schools do it >1/week	—	—
Reading (guided)	Twice a week	*M*: 51'/week (min-max: 0-100)	*M*: 38'/week (min-max: 20-75)
Homework	Usually > 3 a week	More than once a week	More than once a week
		Spain	
Reading aloud	Usually more than once a week	—	—
Reading (guided)	Usually done on a daily basis	*M*: 242'/week (min-max: 150-480)	*M*: 174'/week (min-max: 120-300)
Homework	Once a week	More than once a week	More than once a week

Note. — = This question was not included at this time point.

During Year 1 and Year 2, teachers were asked about whether certain specific aspects of literacy instruction were taught (Table 2.5). Virtually all schools taught word-level literacy skills (word reading and spelling) as well as phoneme awareness and the sounds and names of letters. With time, fewer schools indicated that they kept explicitly teaching these skills, particularly within the Spanish schools, which may be taken to mean that they considered that most children had already mastered some of this abilities. In contrast, all schools continued including the teaching of higher-level literacy skills, such as reading comprehension and text writing; even some schools which, during Year 1, reported not teaching these skills systematically, were doing so by Year 2. Only a few schools in the UK and almost none in Spain taught spelling or reading of nonwords. In the rare instances in which this practice was reported, it was more likely to be reading, rather than spelling, of nonwords.

In addition to surveying what teachers did in the classroom, they were also invited to express their beliefs about the expected levels of attainment for different literacy skills. Teachers in the UK usually expected children to achieve full mastery of word-level literacy (word reading and spelling) by the end of Year 2, while Spanish teachers generally replied that they expected full word-level mastery to be achieved one year sooner. Both groups of teachers, in the UK and in Spain, had more modest prospects for children's performance in text-level literacy: reading comprehension and text production. This suggests that (1) no salient differences in levels of achievement are expected for reading as opposed to spelling; (2) differences in teachers' beliefs about children's levels of attainment seem to be rooted in the type of unit in question: word vs. text, with the latter being understood as a more protracted development than the former; (3) Spanish teachers expected a faster mastery of the word-level domain on the part of their pupils, than the UK teachers with theirs, while no major differences can be attested with regards to the text-level domain.

Table 2.4

Summary of teachers' responses about writing activities

	Reception	Year 1	Year 2
		UK	
Letter formation	≥ 15 min/day (4 schools); <15 in/day (5 schools)	*M*: 36'/day (min-max: 0-75)	*M*: 40'/day (min-max: 0-120)
Spelling	—	*M*: 50'/day (min-max: 10-100)	*M*: 17'/day (min-max: 0-40)
Text writing (autonomous)	≥ 15 min/day (6 schools); <15 in/day (3 schools)	*M*: 55'/day (min-max: 0-100)	*M*: 15'/day (min-max: 0-30)
Text writing (guided)	—	*M*: 63'/day (min-max: 25-150)	*M*: 30'/day (min-max: 0-80)
Homework	Occasionally or never	Once a week or less (usually only spelling, not writing, homework)	Once a week (usually only spelling homework)
		Spain	
Letter formation	At least 15 mins/day (usually more)	*M*: 26'/day (min-max: 10-60)	*M*: 13'/day (min-max: 10-30)
Spelling	—	*M*: 18'/day (min-max: 10-30)	*M*: 14'/day (min-max: 10-30)
Text writing (autonomous)	Usually around 15 mins/day	*M*: 13'/day (min-max: 5-30)	*M*: 13'/day (min-max: 5-30)
Text writing (guided)	—	*M*: 20'/day (min-max: 5-30)	*M*: 15'/day (min-max: 5-45)
Homework	Usually once a week	Once a week or more (spelling and/or writing homework)	More than once a week (both spelling and text writing homework)

Note. — = This question was not included at this time point.

Table 2.5

Summary of teachers' responses about the literacy skills taught in Year 1 and Year 2

	Year 1		Year 2	
	UK	Spain	UK	Spain
Letter names	Yes (all schools)	Usually (10/12 classrooms)	Yes (all schools)	Often (4/10 classrooms)
Letter sounds	Yes (all schools)	Yes (all classrooms)	Yes (all schools)	Usually (7/10 classrooms)
Phoneme awareness	Yes (all schools)	Usually (9/12 classrooms)	Usually (4/6 schools)	Often (4/10 classrooms)
Word reading	Yes (all schools)	Yes (all classrooms)	Yes (all schools)	Usually (6/10 classrooms)
Nonword reading	Often (5/9 schools)	No	Rarely (2/6 schools)	No
Word spelling	Yes (all schools)	Yes (all classrooms)	Yes (all schools)	Yes (all classrooms)
Nonword spelling	Often (4/9 schools)	No	No	No
Reading comp.	Usually (7/9 schools)	Yes (all classrooms)	Yes (all schools)	Yes (all classrooms)
Text writing	Yes (all schools)	Usually (10/12 classrooms)	Usually (5/6 schools)	Yes (all classrooms)
Spelling patterns	Yes (8/9 schools)	Yes (10/12 classrooms)	Yes (5/6 schools)	Yes (all classrooms)

Table 2.6

Summary of teachers' responses about the ideal level of attainment in Year 1 and Year 2

	Year 1		Year 2	
	UK	Spain	UK	Spain
word writing	Partial to full	Full	Full	Full
text writing	Partial to full	Basic to partial	Partial	Partial
word reading	Partial to full	Full	Full	Full
text comp.	Basic to partial	Partial to full	Partial to full	Partial to full

2.3.3 Concluding Remarks and Implications for Subsequent Studies

The assessment of teacher practices yielded two major conclusions: (1) schools within each language group/country showed very similar performances in terms of teaching to write and, more generally, in terms of literacy instruction practices; (2) across languages, similarity was also the rule, rather than the exception, particularly in the Reception Year assessment. These results strengthen the comparability of the samples across languages from the viewpoint of the amount and type of instruction they received. Importantly, the lack of significant differences across all literacy skills in the Reception Year assessment indicates an excellent starting point match. Additionally, it is worth emphasizing that, despite common beliefs about phonics instruction in Spain—which is typically considered not to occupy a predominant role in the literacy lesson and is not compulsory before Year 1—children in both countries received a comparable degree of this type of teaching approach to literacy from Reception Year onwards. Some differences with regards to the amount of reading activities and storytelling time were attested. Nevertheless, it would be highly speculative to elaborate on the implications of such subtle differences, especially in view of some methodological issues (see next section). It should be noted, however, that no differences were found across the range of activities more closely related to writing development.

2.3.4 Limitations

A series of limiting factors in the assessment of teaching practices have been identified and should be borne in mind when considering teacher effects in the development of writing. Firstly, the questionnaires were either too restricted in the options made available to teachers—as was the case with the Reception Year instrument—or, conversely, the questions were of a nature that made the answer format quite liberal and difficult to compare directly. Therefore, proportions had to be calculated to compare the estimated weighing of different literacy teaching activities. Secondly, the small sample of schools/classrooms from which students had been recruited limits generalizations on the impact of teacher practices on early writing development from a cross-linguistic viewpoint. Although it was not an aim of the present work to investigate such impact, these findings may only be applicable to the current samples.

Chapter 3

Exploratory study of microstructural features in English and Spanish early written composition

The aim of this chapter is to explain the framework for the thesis, specifying its relation with a larger research project on literacy development from a cross-linguistic perspective, and to provide a detailed exploration of the microstructural characteristics of early written productions by English- and Spanish-speaking children. Therefore I report (1) the research project that was the framework for the present series of studies; (2) participants' demographic information; (3) time points considered for the thesis; (4) the main writing task; (5) transcription and coding procedures, and coding reliability; and (6) descriptive statistics and psychometric properties of the measures. The vast majority of the text-based measures were found to be adequate for the study of early text composition in terms of their distributional properties and informativeness. Children were well off floor for many text features, while other measures seemed to be in their initial developmental stages. In spite of the multitude of cognitive demands of text composition, children managed to produce legible texts and to convey simple messages, showing awareness of several characteristics of written language.

3.1 The ELDEL Project

The data that were collected for the present doctoral thesis were obtained within the framework of the *Marie Curie* FP7-People Initial Training Network grant, entitled *Enhancing Literacy Development in European Languages* (henceforth, ELDEL, FP7-PEOPLE-2007-1-1-ITN 215961, PI: Markéta Caravolas. For a description of the project see http://www.eldel.eu/workpackage1). The project's general aim was to identify the main environmental, psycholinguistic, and cognitive factors that influence literacy development, with a strong cross-linguistic component: Participants from five different countries (Czech Republic, France, Slovakia, Spain, and the UK) took part in a series of research studies organized in Work Packages.

Work Package 1, *Establishing the Foundations of Literacy Development in European Languages* (henceforth, WP1) was a three-year longitudinal study looking at early literacy learning in the above-mentioned countries. The PhD candidate was appointed as an Early-Stage Researcher, taking part of the British strand of the project. Participants were children starting Reception Year (or its equivalent in the rest of the countries), who were followed until the end of Year 2. They were all monolingual speakers of the main languages of each country—Czech, French, Slovak, Spanish, and British English—, all of which use alphabetic orthographies. Tests batteries were designed to ensure cross-linguistic comparability and included standardized tests as well as bespoke measures to obtain estimates of, among others, the following abilities: word identification, letter and word spelling, letter knowledge, phonological awareness, morphological awareness, receptive syntax, orthographic awareness, visual attention, verbal memory span, and verbal and non-verbal IQ (for more details about the test battery, see Caravolas et al., 2012).

The studies that comprise this thesis were thus borne out of this framework, focusing on the contrast between two of the languages involved: English and Spanish. For the specific objectives of the thesis, an additional *text* writing task was carried out by these two groups of children. In the rest of this chapter I will (1) provide complete demographic information about the groups participating in this body of work; (2) describe the writing task; (3) give details of the transcription and coding criteria, and coding reliability of the text productions; (4) provide a rationale for the selection of measures obtained from the texts; and (5) report the results of descriptive statistics for each measure, time point, and language, together with a comment on relevant psychometric properties. Finally, implications for subsequent analyses will be discussed.

3.2 Participants

One hundred eighty-eight British-English children (98 boys) and 190 Spanish children (104 boys) were recruited for the English and Spanish strands of WP1, respectively. No cross-group differences were found in the number of male and female participants, $\chi^2(1) = .26$, $p = .611$. British children were recruited from ten classrooms of eight Primary schools and one Infant's school, representing a wide socioeconomic range in the areas of Yorkshire (Leeds, Doncaster, Sheffield), Cheshire (Chester), and Lancashire (Blackburn). The Spanish group was recruited from 11 classrooms in five public schools of the area of Granada, in the south of Spain. Such broad sampling reduced the likelihood of biases due to individual teachers and/or demographic factors. The inclusionary criteria specified that children were monolingual, native speakers of English (UK sample) or Spanish (Spanish sample), without any speech, cognitive, and/or hearing disability that would prevent them from completing any of the tasks.

Children in the English sample attended schools which followed the National Literacy Strategy, which adopted a phonics-based approach to literacy instruction. Children in the Spanish sample attended schools which followed the regional (i.e., Granada's) 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. The analysis of the data collected via the questionnaires has been reported in Chapter 2 above, with results showing no significant cross-linguistic differences in the teaching of writing. For this reason, teaching practices were not controlled for in the rest of the studies included in the thesis. For more details on the cross-linguistic comparison of teaching practices, please see Section 2.4 above.

Children in England enrol in Reception classes the September after their fourth birthday. In contrast, children in Spain start the equivalent class the year they turn five. In consequence, children in the Spanish group were, on average, some six months older than their English counterparts. Not surprisingly, a significant age difference was found, $t(376) = 17.11$, $p < .001$, $d = 1.76$. For this reason, age will be entered as a covariate in analyses directly comparing across languages.

Parents' education level was available for 91 and 118 English and Spanish children, respectively. The data were collected using a questionnaire to be completed by either parent. In the English sample, 49% of parents reported having attended university-level education, 43% reported having completed secondary education, and 5% reported having completed only primary education. In the Spanish sample, 53% of parents reported having attended university-level studies, 35% completed secondary or pre-university studies, and 5% completed only primary education. Parents were also asked to report the number of completed years of schooling. English parents were found to have completed a mean of 15.02 (SD = 3.57) years of schooling, whereas Spanish parents completed a mean of 12.50 (SD = 4.05) years of education. A statistically significant difference was found, $t(207) = 4.69$, $p < .001$, $d = .66$, which favoured the English group. For this reason, parents' education (in number of years) will be entered as a covariate in all analyses directly comparing across languages.

Children were assessed six times between Reception Year and the end of Year 2, twice each school year: first towards the middle of the school year (November) and then towards the end (June). However, the thesis draws text-writing data from four of those testing times (end-Reception, mid- and end-Year 1, mid-Year 2) for the English group and, due to issues with the data collection process, the Spanish data set consists of only three consecutive testing times, (mid- and end-Year1, and mid-Year 2). Note that Study 4 on longitudinal predictors of text writing (Chapter 7) includes a set of linguistic and cognitive measures obtained at Time 1 (mid-Reception Year), but no text writing task was administered prior to Time 2 (end of Reception year). All testing times were carried out simultaneously in both countries.

Table 3.1 shows the participants' characteristics at each testing point. Sample attrition was due to drop outs which, in all cases, were a result of children leaving the school (for example, because they moved to a different city). One-hundred fifty-one children in the UK sample and 131 children in the Spanish sample were present at all the testing points allowing for a direct comparison (i.e., Time 3, Time 4, and Time 5).

Table 3.1

Participants' characteristics and time points

Language		*Time 1*	*Time 2*	*Time 3*	*Time 4*	*Time 5*	*Time 6*
	School year	Reception	Reception	Year 1	Year 1	Year 2	Year 2
English	Testing month	Feb-Apr	June	November	June	November	June
(North England, UK)	n (boys-girls)	188 (98-90)	172 (89-83)	173 (88-83)	171 (89-82)	166 (87-79)	—
	Mean age (SD)	60.27 (3.67)	63.31 (3.56)	68.05 (3.62)	75.04 (3.62)	80.01 (3.65)	—
	School year	Reception	Reception	Year 1	Year 1	Year 2	Year 2
Spanish	Testing month	Feb-Apr	June	November	June	November	June
(Granada, Spain)	n (boys-girls)	190 (104-86)	179 (97-82)	167 (91-76)	161 (91-70)	170 (91-79)	—
	Mean age (SD)	66.72 (3.66)	70.88 (3.67)	75.77 (3.65)	83.04 (3.65)	87.79 (3.71)	—

Note. Cells in grey shade indicate that no text-writing data were available for that time point and/or language. No data were used in this thesis from Time 6 (end of Year 2)

3.3 The Writing Task

All children were administered the same task: a five-minute writing activity in which they had to tell recent personal events. Children were gathered in small groups of four or five and were given a blank page that was divided into two main sections: a section for writing at the top, and a section for drawing at the bottom (Figure 3.1). Children were asked to remember what they had done the day before after leaving school (alternatively, what they had done over the weekend, if testing took place on a Monday). The administrator encouraged each child to tell the group out loud what they had done. She then explained to the children that they had to write as much as they could about those things they had just recounted. Children were allowed five minutes to write. The fixed time frame was motivated by the possibility to measure writing productivity; moreover, previous studies used the same time span with good results (e.g., Berninger et al., 1992). No feedback was given to the children regarding content or spelling, and they were appraised for any efforts at writing. Once the five-minute period elapsed, they were told to draw a picture in a square box placed in the bottom half of the page, during which time the administrator could ask each child to read back what s/he had written, and transcribed the “intended” text verbatim on the back of the child’s sheet. This information was then used to interpret the written productions, especially when handwriting legibility was poor.

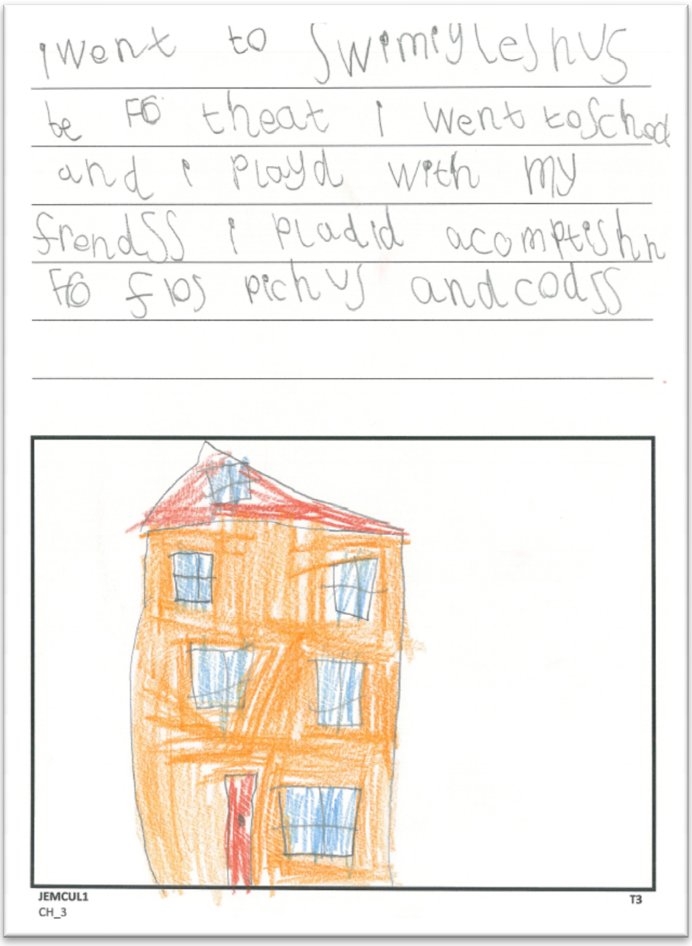


Figure 3.1. The upper half was for the child to write what s/he had done the day before. The bottom half was for the child to draw a picture, during which time the administrator could ask each child to read back to her what was on the page. This last part of the task procedure was mostly meant to elucidate illegible handwriting.

The rather open nature of the task required children to undergo all aspects of the writing process—at the very least, content retrieval and translation—, while at the same time it involved a simple prompt, thus making the task less demanding. Considering the well-established claim that spelling takes up a great deal of early writers' cognitive resources (e.g., Berninger & Swanson, 1994; Bourdin & Fayol, 1994), it was important to make content retrieval as simple as possible. A way of doing this was to elicit writing on a topic about which they were knowledgeable (McCutchen, 2006). Typically this has lead researchers to ask children to retell a well-known story (e.g., Ferreiro, Pontecorvo, Ribeiro-Moreira, & García-Hidalgo, 1996), but while such a procedure elicits a known, well organized structure, and rich content that can be used by the researcher to evaluate different aspects of the child's production, it has the disadvantage of eliciting something that the children know by rote and thus not be entirely of their own making. Asking children to report recent past events, on the contrary, has the benefit of leaving almost all aspects of the writing process up to them, but arguably facilitates content retrieval, making them focus on recent, therefore highly accessible activities.

In order to further aid in the content retrieval process, the administrator asked the children what they had done the day before, not as a means for them to "practice" what they would write or to elicit spoken versions of their texts (this part of the task was never recorded), but to help them recall their activities, so that they could make full use of the five-minute time frame provided. Children were never told in advance that they were going to be asked to write those events down, although many of them intuited they would at subsequent time points. The objective was to "activate" the events in which they had been recently involved, so that they had a clear starting point to begin the text composition task.

3.4 Transcription of Text Samples

A total of 1,180 texts were collected as described above and were transcribed in an Excel spreadsheet. First, each word, as originally written by the child (henceforth, *original text/word*), was transcribed in a single cell, applying conventional word boundaries. A string was considered a word when its bare form—i.e., without any inflectional morphemes added—corresponded to a dictionary entry (the *Concise Oxford English dictionary,* 2002 was the reference point for English, and the *Diccionario de la Lengua Española*, 2001 [Dictionary of the Spanish language] was used for Spanish). For example, *went* was counted as one word, but phrasal verbs like *pick up* or *take off* counted as two words. Similarly, *swimming pool, fútbol-sala* 'five-a-side football', or *lie-in* were transcribed as one word, but *racing car* was transcribed as two words. If the child did not segment phrases into words conventionally, such as writing <Iwent> instead of *I went,* the original text was nevertheless transcribed as two independent words (i.e., in two different cells), signalling that the word had been produced attached to the previous and/or the following one, using the symbol ~ at either or both ends of the word. For example, the string <Iwent> would be transcribed as *I~* (cell 1) and *~went* (cell 2). Accordingly, unconventional segmentations of a word (e.g., <a bout>) were transcribed in the original text cell with a low hyphen where the illegal separation was produced, e.g., a_bout. Proper names, especially names of films, games, books, or places, were written in one cell only, regardless of the number of words of which they were composed. In the case of names of places, if they included words like *city, park,* etc., these were transcribed separately.

Next to each of the original text transcriptions, the child's intended word was transcribed (henceforth, *intended text/word*). Usually the child's intended word would have been determined through his/her rereading to the administrator. Intended words were orthographically normalized. If the child included punctuation marks, these were transcribed at the end of the word next to which they appeared in the original text, but not in the corresponding intended cell.

Children's use of case in their texts was respected in the transcriptions, applying the following criteria: Letters with distinctive lower- and upper-case features (e.g., <a> and <A>) were always transcribed as written by the child. For letters in which it is mostly size that determines which case has been used (e.g., <c> and <C>), the researcher always transcribed the lowercase version, except if there was an obvious size difference, which was moreover applied systematically throughout the text. With regards to the use of case in the intended text cell, all words beginning a text were written in the intended text cell with an uppercase initial, as were proper names, names of the days of the week, months, and so on[4]. Moreover, if the child had written a full stop or any other punctuation mark after which a capital letter is required, the following intended word was written with uppercase initial, regardless of whether the use of the full stop or punctuation mark was appropriate.

The timed nature of the task meant that some children were in the middle of a sentence and were required to stop writing (see the task description and procedure above). In such cases it was possible for the child, when rereading to the administrator, to express the word or words he did not have time to write down. These words were also transcribed as intended text, but filtered out for all analyses, except when otherwise specified. They were mostly kept with a view to obtaining a better feel of the full text as planned by the child. In the same way, it became apparent

[4] Note that in Spanish names of days of the week and months should not be written with upper-case initial, except if written at the beginning of the sentence. The same happens with names of languages (e.g., *Spanish-español*) and nationalities (e.g., *British-británico*).

on occasions that children left out some words in their texts which they however did "read". These were also transcribed as intended text. In both cases—when intended words that had not been written because the child ran out of time, or those which were simply omitted—, the original text cell was left blank.

An additional note concerns the treatment of different kinds of illegible strings. The original text cells were tagged to identify cases of scribbling, letter-like symbols, and random letter strings[5]. In addition, if a word could not be determined by the administrator and the information provided by the child was insufficient to elucidate the intended word for a certain string of letters, the researcher wrote a question mark next to the best guess of both the original and intended words, whenever possible. Transcription samples are included in Appendix 1.

3.5 Microstructural Features of Early Text Writing

Writing research has fluctuated between top-down and bottom-up approaches. Top-down analyses of written text productions focus on macrostructural features of texts, involving aspects of text composition affecting the larger configuration and organization of the text, its overall coherence, richness of content, or communicative effectiveness. From this perspective, specific, local text features are seen as the result of higher-level discourse properties and the general communicative function and setting of the text (e.g., Martin, 2009; Van Dijk, 1985). Bottom-up approaches to written text analysis focus, instead, on the microstructural characteristics of a text, such as vocabulary choices, syntactic complexity, notational aspects (spelling, punctuation, capitalization), use of connectivity devices, and so on. The term *microstructural* features has been used in current literature in opposition to *macrostructural* writing features (e.g., Puranik et al., 2008; Wagner et al., 2011). Microstructural features typically comprise transcription and text generation aspects of writing (Puranik et al., 2008; Scott, 2005). Macrostructural features refer to those aspects of text composition affecting the larger configuration and organization of the text, its overall coherence, richness of content, or communicative effectiveness. The micro- and the macrostructural levels of a text are thus constantly interacting with

[5] Randomness of a given string was ascertained when no overlap between the intended word and the letters on the page could be established. Linguistic and contextual clues were taken into account.

each other, and they both seem necessary to provide a full account of written text production (Tolchinsky, 2013).

Only a few studies have looked at written composition encompassing both micro- and macrostructural levels of analysis. Malvern, Richards, Chipere, & Durán (2004) investigated a series of quantitative measures to determine the best predictors of teachers' ratings of written texts, using a large sample of texts produced by British students aged 7-14. Within a different framework, Berman and Nir (2009) proposed an analysis of expository written texts along two dimensions: *local linguistic expression* and *global text quality*. Local linguistic expression was assessed by measures of vocabulary and syntactic complexity, while global text quality was evaluated by means of a bespoke rubric assessing three dimensions: *representation and cognitive processes, structure and content,* and *discursive features.* For each of these dimensions, texts were classified into one of four different levels of quality, ranging from *minimal representation* to *structural well-formedness*, to *beyond well-formedness.* The authors reported that both the local-quantitative and the global-qualitative measures showed excellent validity and reliability.

There are difficulties with assessing macrostructural features of beginner writers. They are problematic to operationalize, usually necessitating the resort to global or holistic scores, which have shown validity issues (Berman & Nir, 2009). In addition, texts produced by beginner writers tend to be short, impacting negatively on inter-rater reliability (Berninger et al., 1992; McMaster & Espin, 2007; Puranik & AlOtaiba, 2012). In contrast to the difficulties inherent to the assessment of macrostructural aspects of written composition, text-embedded, local aspects of text composition are easier to operationalize and quantify. Moreover, recent studies on early writing development seem to overwhelmingly favour this type of approach (e.g., Puranik & AlOtaiba, 2012; Puranik et al., 2008; Wagner et al., 2011). Finally, it has been suggested that quantifiable surface forms may be used as correlates of qualitative constructs (Malvern et al., 2004; Page, 1994). Therefore, this thesis will focus exclusively on microstructural aspects and will adopt a fully quantitative approach. Importantly, microstructural features usually constitute the basis of judgements about the perceived levels of macrostructural text quality (McNamara, Crossley, & McCarthy, 2010).

The writing features obtained from the texts produced by children were selected in line with recent studies on writing development and in view of current models of writing. Therefore, a number of measures were obtained to estimate children's performance along the following dimensions: (1) *transcription;* (2) *productivity* or *text length*; (3) *text structure*; (4) *syntactic complexity*; and (5) *vocabulary*. Transcription skills were estimated through writing features which have their main impact at the word level. It thus included *spelling accuracy*, *word segmentation*, *use of case*, and *representation of morphology*. Although spelling is usually the main (or only) indicator of transcription performance, children's knowledge of conventional word boundaries (word segmentation) might prove essential for applying some orthographic rules. Correct use of case was evaluated not only because it is part of the conventions that children must master in order to communicate through writing, but also because the correct selection of the appropriate case reflects their knowledge of the notational system they are learning (Ravid & Tolchinsky, 2002). Children at these very early stages might not yet be aware of all pairs of lower- and upper-case letters. Finally, using morphological knowledge to determine correct spelling is considered an advanced strategy in stage-like spelling development models (e.g., Frith, 1985; see Pacton & Deacon, 2008 for a review on the role of morphological awareness in spelling development).

The second dimension, productivity, has been assessed in most studies on writing development. A simple measure of the *number of words* produced has been shown to be an excellent indicator of children's text composition abilities (e.g., Berman & Verhoeven, 2002; Connelly, Dockrell, Walter, & Critten, 2012; Mackie & Dockrell, 2004; Puranik et al., 2008; Scott & Windsor, 2000). Because the average word length differs cross-linguistically (Section 2.2.2), productivity was also measured as the *number of letters* that children produced.

The third dimension of text structure should not be confused with an evaluation of text organization or coherence; instead, in the present thesis it refers to a construct of specific writing features that can be argued to contribute to such coherence and organization. The layout of words on the page (henceforward, *text layout*) was evaluated to determine whether children integrated the different sentences into a larger text unit (such as a paragraph) or whether they left each sentence or idea to stand on its own (as in a list). In this sense, it was assumed that constructing larger units of integrated discourse was more demanding than representing isolated, unconnected text parts. The text structure dimension also included a measure of children's use of connectives (conjunctions and discourse markers), and punctuation marks. The use of punctuation marks is usually regarded as a convention in text writing studies (e.g., Puranik et al., 2008). Nevertheless, punctuation marks are a fundamental piece in text construction which do not merely serve to fulfil a convention. They are crucial to marking text boundaries, the connections between text units, and they are essential to resolving potentially ambiguous sentences. The very scarce research on the early uses of punctuation marks seems to suggest that children begin by using full stops prior to any other punctuation mark. In addition, they appear to use full stops to mark the end of the text before using them to signal the end of a sentence (Ferreiro & Pontecorvo, 1999). These findings suggest that children are using punctuation marks in a non-random way and they moreover use them to mark text boundaries, with larger units being identified prior to smaller ones. It is in this sense that punctuation marks were taken to contribute to the structure of written texts.

A fourth dimension of syntactic complexity was assessed through two measures: (1) a measure of syntagmatic syntactic complexity of the average number of words produced per clause; and (2) a measure of syntactic embeddedness, which evaluated the percentage of subordination in children's text productions. Syntactic complexity has been shown to be a good marker of text quality (e.g., McNamara et al., 2009) and to be affected in the writing of children with dyslexia or children with language impairments (e.g., Mackie & Dockrell, 2004).

A fifth dimension of vocabulary choices was included. It included three measures: (1) a measure of lexical diversity, obtained by calculating type-token ratios; (2) a measure of lexical density (e.g., Malvern et al., 2004), which evaluates the proportion of content words (nouns, verbs, adjectives, and adverbs) to the total of words produced, based on the idea that content words add to the degree of informativeness (i.e., density) of the texts;(3) a measure of the number of optional and evaluative word tokens, estimated through the proportion of adjectives and adverbs. These word categories add to the precision and richness of the contents being reported. Finally, (4) a measure of lexical sophistication was included, which was calculated as the average length in letters of all content words. This last measure is based on the assumption that longer words tend to be less frequent than shorter words as a result of general grammaticalization processes in language (Bybee, 2007). The average length of content words may then be used to estimate the average lexical frequency or, more generally, the average degree of lexical sophistication. Infrequent lexical items are usually more precise in meaning (Biber, 1995), are acquired later, and may also reflect the adoption of a higher register on the part of the writer.

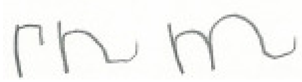
3.6. Coding Criteria

Coding of the corpus aimed to provide a thorough description of children's written productions. Each word in children's texts was assessed for (a) its level of representation, thus evaluating whether the child was attempting to represent the phonological structure of the word; (b) spelling accuracy; (c) word segmentation accuracy, that is, showing conventional word boundaries; (d) case, evaluating the appropriate use of lower- and upper-case letters; (e) representation of morphology; (f) text length or productivity; (g) text layout, that is, paragraph-like or other formats; (h) punctuation; (i) connectives, that is, the child's use of conjunctions, discourse markers, and other lexical means of text organization; (j) syntactic complexity; (k) vocabulary-related measures. In what follows I describe the full coding criteria.

3.6.1. Level of representation. Each word was classified into ten different categories according to the type or level of representation of the target words it displayed (Table 3.2). Scribbling, letter-like strings, and random letter strings, each formed a separate group, and coded as categories 0, 1, and 2, respectively. Illegible words in children's texts were coded in category 3, while letter strings of which the corresponding intended text could not be determined were coded in category 4. Category 5 grouped words that were intended by the child, but which were not written down because the child ran out of time. Words omitted for other reasons were grouped in category 6. Category 7 accommodated loan words, usually names of foreign video games and films, of which the target phoneme-grapheme

Table 3.2

Categories in the assessment of Level of Representation

Categories	Description	Original word	Intended text
0	Scribbling		They were very poor
1	Letter-like string		My daddy...
2	Random string		I went to...
3	Illegible string		?
4	Intended word cannot be determined (insufficient information)		?
5	Word intended, but out of time	—	[friends]
6	Word omitted	—	of
7	Loans		Wii
8	Numbers and symbols	*7*	seven
9	Fully analyzable words		played

mapping could not be determined because they are pronounced in different ways even by adult monolingual speakers. This category was mostly meant for the Spanish sample, where these loans are frequent. Nevertheless, it served to accommodate some names (e.g., of websites, *CBeebies*) in the English corpus as well. Category 8 included numbers or symbols that children wrote instead of using the corresponding words (e.g., <7> for *seven,* <&> for *and*). Finally, category 9 included all *analyzable words*; that is, words bearing at least some relation to the intended text, which, in turn, was unambiguously determined through the rereading, and where a conventional phoneme-grapheme mapping could be established.

3.6.2. Spelling accuracy. Each original word was scored as correct or incorrect against its intended word, according to the orthographic conventions in each language. Diacritic omission in Spanish was not penalized, in order to avoid underestimating these children's knowledge of phonographic correspondences, while leaving orthographic marking of suprasegmental aspects out of the analysis. Similarly, in the English corpus, incorrect use or omission of the apostrophe was not penalized. Level categories 0 through 2—scribbling, letter-like and random letter strings—were given credit and scored as incorrect, the underlying reason being that the child made an attempt at writing those words. Level categories 3 through 8 were excluded from this analysis.

3.6.3. Word segmentation. Children's awareness and knowledge of word boundaries in writing was assessed classifying the written productions into four possible categories. The classification criteria were taken from Ferreiro and Pontecorvo (2002), and adapted for a word-by-word analysis. The categories were (1) hyposegmentation, e.g., <Iwent>, where both words were penalized as being hyposegmented; (2) hypersegmentation, e.g., <a bout>; (3) mixed hypo- and hypersegmentation, e.g. <com preun> '(I) bought a', where the verb *compré* is unconventionally segmented in the middle, while at the same time its last syllable is attached to the following article *un* 'a_masc'; and (4) accurately segmented words. Compound words, so frequent especially in English, presented a difficult case. As mentioned above (Section 3.4), they were treated as one word if they constituted an independent dictionary entry (e.g., *Concise Oxford English Dictionary,* 2002). However, whether they should be written as one continuous string (e.g., *bedroom*), two (or more) hyphenated words (e.g., *lie-in, hide-and-seek*), or two words without a hyphen (e.g., *swimming pool*) is for the most part an arbitrary decision. For this reason, children were not penalized for the way they segmented these units, and one- and two-word versions, with or without a hyphen, were accepted as correct. Following Ferreiro and Pontecorvo's (2002) criteria, unconventional segmentations that occurred at the end of a line were not penalized either. The presence of a hyphen anywhere else in the text, together with full stops and commas, were taken as indicators of a child signalling a word boundary. Words classified in categories 0 through to 8 in the Level of Representation assessment were excluded from this analysis.

3.6.4. Case. Children's knowledge of the appropriate contexts to use lower- and upper-case letters was evaluated comparing each original word to its corresponding intended word, which had been transcribed following the criteria explained in Section 3.4 above. If the child used a capital letter in the same position as the intended word, it was scored as "correct", with any different uses scored as "incorrect." There had to be full overlap between the case of the original and the intended text; that is, if a word had to be written with a capital initial (e.g., a person's name) and the child wrote all capital letters (e.g., <JOSE>, <PHOEBE>) it was coded as incorrect. Conventional orthography was not confounded in this analysis, therefore <Tose> (for *José*) and <Feebee> (for *Phoebe*), for example, would have been scored as correct in this analysis. Words classified in categories 0 through to 8 in the Level of Representation assessment were excluded from this analysis.

3.6.5. Representation of morphology. Given the cross-linguistic differences in morphology (Section 2.2.2), it was not possible to assess its written representation in an identical way in both languages. For the English texts, the main objective of the morphology measure was to determine whether the child was representing morphology or not in an obligatory context. An obligatory context was defined as all instances in which (1) the intended word is multi-morphemic, and one of the morphemes is inflectional; (2) the inflectional morpheme in question has both a phonetic realization —e.g., as is the case of the plural morpheme in English nouns, /s/, in contrast to the singular morpheme, /ø/— and it is moreover represented in writing (in the above example, *-s* or *–es*). In this way, the relevant morphemes analyzed were, (a) *-ed* (past simple or participle); (b) *-(e)s* (plural, genitive, or the contraction of verbs forms *is* and *has*); and (c) *-ing.* Irrelevant words, that is, those which did not constitute an obligatory context, including also scribbling, and letter-like and random strings, received a score of 0. A score of 1 was given to words in obligatory contexts where the corresponding morpheme was not represented (e.g., <woch> for target *watched*). A score of 2 represented words in which the inflectional morpheme had been written phonetically (e.g., <watch**t**>). A score of 3 was given to words in which morphology was only partially represented (e.g., <watch**d**>). Finally, a score of 4 was given to words with a conventional orthographic representation of the pertinent morphological ending (e.g., <watch**ed**>).

In the case of Spanish, where the vast majority of words are multi-morphemic, morphological endings cannot be avoided neither in speech nor in writing, since the result is usually an unpronounceable sequence of sounds. Substitution errors do occur, but only in earlier stages of oral language development and in language-impaired children (e.g., Bedore & Leonard, 2001, 2005). Spanish-speaking children (and even some educated adults) nevertheless struggle to represent morphology in a conventional manner in those cases where phonographic correspondences are insufficient to lead to the correct spelling (Defior et al., 2008). For example, typical spelling errors involving morphemes are the omission of <h> in forms of the auxiliary verb *haber* 'there is/are_INF', such as **e ido* instead of *he ido.* In the present analysis, only multi-morphemic words which involved an inflectional morpheme that, in addition, presented some inconsistency in phoneme-to-grapheme mappings were considered. For example, some forms of the imperfective past tense involving phoneme /b/, which is inconsistent in Spanish, and can be represented both by <v> and <b>. However, the imperfective past is always written with <b>. Accordingly, words regarded as relevant for the representation of morphology analysis in Spanish included (a) final *-s,* (plural and some verb inflections), which is usually omitted in the region of Spain where these data were collected (Defior et al., 2008), constituting a case of 'silent' morphology; (b) all conjugated regular verb forms ending in -*aba(n)/(s)*; and (c) the presence of letter "h" in all forms of the verb *haber* when used as an auxiliary, as in *he ido* '(I) went'*.* Non-relevant words, together with scribbling, letter-like, and random strings, were given a score of 0. A score of 1 was given to words in which the inconsistent morpheme was represented with an unsuitable letter or letters (e.g., <llamavan> 'were_calling_3PL', instead of the correct *llamaba*). Finally, a score of 2 was afforded to words with orthographically inconsistent morphemes that were written conventionally.

In the case of both English and Spanish, word segmentation and spelling accuracy of the root were not taken into consideration in this analysis. Also, note that some cases presented an overlap of categories. These cases made it impossible to distinguish between a score of 2 (phonetic representation) or 3 (partial orthographic representation) in some English morphemes, such as the past tense of *play,* which takes the form [-d]. If a child wrote <playd> it was not possible to determine whether <d> was a phonetic representation of phoneme [d] or if it was a partial representation of the conventional orthographic spelling: *-ed.* In such cases, the lowest score was afforded.

3.6.6. Text length. The length of the texts children produced was assessed in two ways: (1) by counting the number of words written by each child in the five-minute task, following the definition in Section 3.4 above; and (2) by counting the number of letters in the intended texts automatically in Excel. Measuring text length in letters was motivated by the different syllabic structure and average word length in English and Spanish (see Chapter 2). Punctuation marks and other symbols (e.g., numbers) were not counted. In both word and letter counts, words classified in categories 0 through 6 in the Level of Representation assessment were excluded from this analysis.

3.6.7. Punctuation. All punctuation marks produced by the children were transcribed next to the original text cell. The presence of punctuation marks was assessed both quantitatively and qualitatively. For the quantitative assessment, the total number of punctuation marks was divided by the total number of words written by the child (i.e., words in Level of Representation categories 7, 8, and 9) and multiplied by 100, to obtain the average percentage of use of punctuation marks. Motivation for using a percentage of punctuation marks over the total number of words, instead of using the raw number of punctuation marks directly, was rooted in the idea that longer texts might give more opportunities for punctuation mark use. Hyphens and apostrophes were not counted as instances of punctuation marks. Note that some punctuation marks are conventionally used in "pairs," such as parentheses or inverted commas; in Spanish, this extends to question marks and exclamation points, which are to be written at both the beginning and end of the sentence they refer to. In all these cases, only one instance was counted.

For the qualitative assessment, each text was classified according to whether it included the following punctuation marks: (1) a full stop at the end of the text; (2) a full stop (or full stops) within the text; and (3) other punctuation marks (question mark, commas, parentheses, etc.). Such coding of the presence/absence of different types of punctuation marks resulted in three binary variables. Dummy variables of this kind are useful to evaluate the presence of a feature when its occurrence is not expected to have a substantial quantitative variation (e.g., Puranik & Lonigan, 2011). In this sense, previous studies on the emergence of punctuation marks found units other than the full stop to be rare at these and even later stages (e.g., Ferreiro & Pontecorvo, 1999). In the same vein, no evaluation of the appropriateness of use of the punctuation marks was performed and only the presence or absence of marks was taken into account.

3.6.8. Layout. This measure was scored using an ordinal variable, classifying children's productions according to the "appearance" of the texts, evaluating whether they used the width of the paper to its full extent, thus having a "paragraph-like" layout. This text feature was evaluated qualitatively according to whether the text presented the following layout types (Figure 3.2): Texts with strings of letters making use of the entire length of the lines on the paper, thus having a paragraph-like appearance, were given a score of 3 (Figure 3.2, d); sentences represented in isolation, starting in a new line to write each separate sentence or event, with the resulting text looking like a list, were scored as 2 (Figure 3.2, c); productions of one word or unfinished sentence per line—like breaking down a sentence into different lines—was given a score of 1 (Figure 3.2, b); finally, texts consisting of only one clause or a clause which took up only one line were scored as 0, since it was not possible to assess the use of space with such short texts (Figure 3.2, a). Children who produced scribbling, random letter strings, or letter-like symbols were not excluded from this analysis, since it was possible to evaluate their conceptions about text appearance from the way in which they tried to "imitate" it.

3.6.9. Connectives. All conjunctions and discourse markers were identified in the texts. They were further tagged and classified into one of the following groups: basic connectives, which included exclusively the lexical items *and, or,* and *but* (*y, o,* and *pero*, in Spanish); subordinating conjunctions (e.g., *that, when, if*); and other connectives (e.g., then, later, first). Note that only inter-clausal connectives and

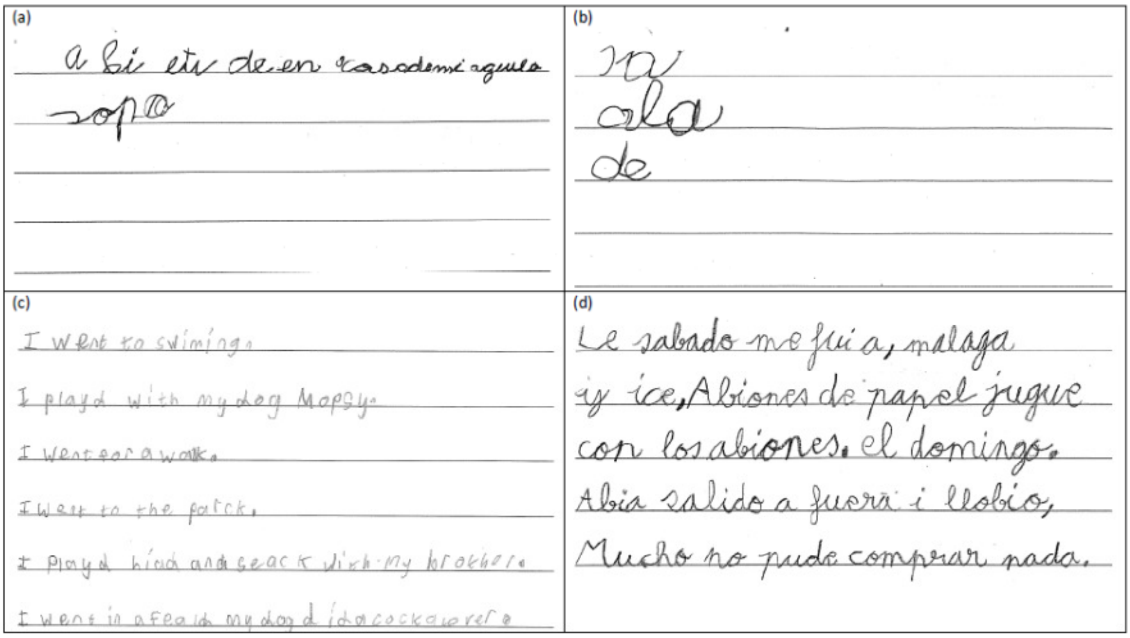


Figure 3.2. Samples of text layout types
(a) Text sample with insufficient writing to determine intended layout; (b) text sample with isolated words or phrases; (c) list-like layout sample; (d) paragraph-like layout sample.

discourse markers were counted, while inter-phrasal conjunctions, such as *and* in *My mum <u>and</u> my dad took me to the movies* were not counted, because the aim was to estimate connectivity at the text level. The total number of connectives in the text was then divided by the total number of analyzable words (Level of Representation categories 8 and 9) and multiplied by 100, to obtain a percentage of use for each connective type. In addition, the distribution of the different connective types over time in each language was explored. The average percentage of each type of connective over the total number of connectives used at any given time point was calculated. For both estimates (percentage of connectives out of total number of words produced and percentage of each type of connective over total number of connectives), only connectives classified in the Level of Representation categories 8 and 9 were considered.

3.6.10. Words per clause. This aspect of syntactic complexity was calculated by dividing the total number of words (Level of Representation categories 7, 8, and 9) over the total number of clauses in the text, which were manually counted. The working definition for *clause* was that of a grammatical unit that consists of a subject and a verb-phrase predicate, whether it be finite or non-finite. The clause types considered were, thus, simple clauses (example 3.1), subordinate clauses (example

3.2), and non-finite isolated clauses (example 3.3). Fragments and isolated noun phrases were excluded from all counts.

Example 3.1 *I went to the park*

Example 3.2 *Last night my granddad and my grandma came round for tea* *because he is a blind maker he took ours down because they was falling apart*

Example 3.3 *Ver la televisión un paseo y por último dormir la siesta*
'Watch TV a walk and at last take a nap'

3.6.11. Percentage of subordination. In order to estimate the amount of subordination in the texts, a subordination index was created by dividing the number of nested clauses (i.e., subordinated clauses) over the total number of clauses produced by the child. No distinction between types of subordinated clauses—for example between relative clauses and completive clauses—, was made. Given that the texts were quite short, pooling them together resulted in a more meaningful estimate of complex syntax in early written productions.

3.6.12. Lexical diversity. The ratio of different words (types) to the total number of words (tokens) in each text was calculated as an estimate of the lexical diversity in the texts. For example, in the sentence *I went home with my mommy and my daddy* there are 9 tokens and 8 types (*my* is repeated twice), thus giving a type-token ratio (TTR) of .889. Any modification to a word base, such as adding or subtracting inflectional endings or derivational morphemes, was interpreted as a different type. For instance, in *I played with my toys. My favourite toy is a teddy,* the tokens *toy* and *toys* were counted as two types, following the general criterion for calculating TTR (see Malvern et al., 2004 for an extended review).

3.6.13. Percentage of adjectives and adverbs. All adjectives and adverbs in the texts where tagged according to the intended text. They were summed and divided by the total number of analyzable words (Level of Representation categories 7, 8, and 9) and multiplied by 100. The resulting measure was the percentage of adjectives and adverbs out of all words in the text. It should be noted that only "individual" adjectives were identified and counted and not adjectival phrases, since any other element in the adjectival phrase, such as adverbs, was counted as another token.

3.6.14. Lexical density. Lexical density was estimated by dividing the number of content words (nouns, verbs, adjectives, and adverbs) over the total number of words in the text and multiplied by 100 (Ure, 1971). The resulting measure was thus the percentage of content words in the text.

3.6.15. Average content-word length. The average length of each content word in the text was calculated by counting the number of letters in all words in a text, filtering out function words. Thus, only nouns, adjectives, verbs, and adverbs were counted. Proper nouns were also excluded from this analysis. The resulting number was then divided by the number of analyzable words (Level of representation category 9) in each text. Therefore, this measure is to be interpreted as the average number of letters per content word.

3.7 Coding Reliability

In order to assess the coding reliability of the measures obtained from the texts, agreement between the PhD candidate and two external coders, blind to the objectives of the studies, was calculated. The external coders were a graduate student of Psychology (for the English sub-sample) and of Linguistics (for the Spanish sub-sample), and they were both native speakers of English and Spanish, respectively. They received a document with the coding criteria used for obtaining each measure and were trained with an initial sample of 10 texts. After training, they 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 was evaluated. Agreement between these raters' and the PhD candidate's coding was 80% or higher for all measures—spelling accuracy (96% for English, 99% for Spanish), segmentation, (97% for English, 99% for Spanish), use of case (95% for English, 97% for Spanish), number of words (94% for English, 99% for Spanish), punctuation (81% for English, 94% for Spanish), text layout (94% for English, 89% for Spanish), and number and type of clauses (90% for English, 96% for Spanish).

For measures based on the tagging of texts for part-of-speech—i.e., 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 PhD candidate's manual tagging and an automatic part-of-speech analyzer (CLAWS4, for English, Garside, 1987; MACO, for Spanish, Civit et al., 2003) on a quasi-random selection of 10% of the texts in each language. Agreement between the two types of counts always exceeded 80%.

Reliability estimates for other measures included in the study were unnecessary, since they were the result of calculations based on the former subset of measures/counts. This affected the following measures: type-token ratio, and number of letters, and average content word length. See coding criteria (Section 3.6) for more details.

3.8. Results

In this section I present the descriptive statistics of the measures obtained from the text productions at each time point in the two languages. Unless stated otherwise, mean percentages and standard deviations are reported for each language and time point. Observations about the similarities and differences between the two language groups will be made tentatively, where possible trends in the data seem to emerge. The effects of (growth over) time and language (or orthography), as well as their interaction, are the subject matter of Study 2 (Chapter 4) below.

The objective of this first set of analyses was thus twofold: On the one hand, they aimed to provide a detailed description of the nature of the text productions; on the other, their purpose was to identify the best and most informative measures to guide subsequent studies. Of the final set of measures, and given that some of them had no precedent in the literature for these age groups, an account of their psychometric properties is also provided.

3.8.1 Level of representation

This variable was created to categorize each word as originally produced by the child in relation to the intended word to determine whether some level of phonographic mapping between the two could be established, and to consider whether the word was analyzable. Table 3.3 shows the distribution of words in each language into the different categories.

The number of fully analyzable strings increased with time in English, $F(1.72, 240.88) = 56.27$, $p < .001$, $\eta_p^2 = .288$, and in Spanish, $F(1.53, 228.05) = 54.83$, $p < .001$, $\eta_p^2 = .269$. The lowest percentage of analyzable words achieved was in the Time 2

English sample, with only 69.53% of words analyzable, but from Time 3 to Time 5 both language groups show high mean percentages of analyzable words, ranging from 85% to 98%. Instances of scribbling were rare throughout the study, in both languages. At Time 5 in English there is a rather high 8.13% mean percentage of scribbling; it was the result of a single child whose "re-reading" (i.e., intended text) was very long. Thus, each intended word was categorized as "scribbling". Note that this extreme case only affects the distribution of this variable, since scribbling was filtered out of most variables. Letter-like strings never constituted a substantial proportion of the corpus, with means always under 4%, and of 0% by Time 5. Words consisting of a random string of letters were the second most important category for English at Times 2 and 3, and for Spanish at Time 3. The proportion of this type of representation decreased with time in English, $F(1.96, 272.16) = 19.16$, $p < .001$, $\eta_p^2 = .121$, and Spanish, $F(1.00, 130.04) = 5.51$, $p = .020$, $\eta_p^2 = .041$, and was almost negligible by Time 4 in both languages.

Categories 3 and 4—unintelligible and undecipherable words, respectively—were extremely rare across time points and languages, and were always below 2% in both language groups. These items were the result of either extremely illegible handwriting (Category 3) or words for which a match could not be established, because the child did not remember what he had intended to write, or because it was impossible for the administrator and the researcher to establish any matches with certainty (Category 4). This is reassuring, in the sense that the vast majority of the words in the texts were interpretable.

Words written "out of time" (Category 5) were very rare as well. Omitted words—i.e., words children pronounced when "reading" back to the administrator, but that were absent from the written production (Category 6)—showed means lower than 1% in both languages by Time 5.

Categories 7 (loans) and 8 (symbols or numbers) never showed means higher than 2% and 1%, respectively in either language, at any time point. Words in these two categories were not really unanalyzable, since their meaning could be easily established. However, their characteristics were incompatible with some analyses. For example, it was not sensible to evaluate a Spanish-speaking child's rendering of a foreign name such as *Playstation* when there are many alternative pronunciations even among educated adults. As for symbols or numbers, it was not possible either to assess sound-to-letter correspondences or the use of case, morphology, etc. However, words belonging to these two categories were included in other analyses in which sound-to-letter correspondences were not relevant (e.g., numbers were counted as adjectives).

Table 3.3

Distribution of words according to their level of representation (mean percentages and SDs)

Categories in Level of Representation	*Time 2*		*Time 3*		*Time 4*		*Time 5*	
	M	SD	*M*	SD	*M*	SD	*M*	SD
					English			
0.Scribbling	0.58	(7.63)	0.16	(1.27)	0.00	(0.00)	8.13	(30.24)
1.Letter-like	3.57	(18.42)	0.00	(0.00)	0.59	(7.65)	0.00	(0.00)
2.Random	**17.77**	(34.66)	**9.80**	(26.20)	2.81	(15.49)	1.53	(10.97)
3.Unintelligible	0.08	(1.09)	0.20	(1.23)	0.09	(0.65)	0.02	(0.29)
4.Undecipherable	1.15	(8.16)	1.13	(3.33)	1.76	(5.74)	0.43	(1.86)
5.Out of time	2.71	(10.20)	1.12	(5.17)	0.46	(2.11)	0.20	(1.16)
6.Omitted	4.60	(13.22)	2.19	(5.10)	1.19	(4.03)	0.71	(2.42)
7.Loans	0.00	(0.00)	0.14	(1.39)	0.11	(1.14)	0.10	(0.79)
8.Symbols/numbers	0.02	(0.32)	0.09	(0.88)	0.19	(0.90)	0.38	(1.09)
9.Analyzable	**69.53**	(37.20)	**85.17**	(27.22)	**92.27**	(20.35)	**95.99**	(13.54)
					Spanish			
0.Scribbling	—		0.00	(0.00)	1.53	(13.98)	2.47	(12.99)
1.Letter-like			1.20	(10.91)	0.00	(0.00)	0.00	(0.00)
2.Random			4.59	(19.68)	0.00	(0.00)	0.02	(0.21)
3.Unintelligible			0.62	(4.00)	0.33	(1.43)	0.31	(1.20)
4.Undecipherable			1.10	(4.37)	0.47	(1.60)	0.28	(0.93)
5.Out of time			0.54	(3.35)	0.44	(3.94)	0.00	(0.00)
6.Omitted			2.97	(8.10)	0.19	(1.08)	0.17	(0.86)
7.Loans			1.53	(6.08)	0.72	(2.35)	0.78	(1.90)
8.Symbols/numbers			0.09	(0.85)	0.25	(1.50)	0.38	(1.27)
9.Analyzable			**87.42**	(24.20)	**97.60**	(5.17)	**98.05**	(3.07)

Categories 0 through to 6 were excluded from all analyses, except for spelling accuracy, where categories 0 to 3 were taken into account, as they do constitute an attempt on the part of the child to represent sounds, and scored as "incorrect". Categories 7 and 8 were included in most analyses, unless otherwise stated.

In sum, except for Time 2 in English, where a non-trivial proportion of words had to be left out of most analyses—usually due to these words being random strings of letters—, the mean percentage of fully analyzable words was very high from Time 3 to Time 5 in both language groups. This means that analyses were carried out using the vast majority of children's productions, while strings that were unclear were excluded to avoid potential over- or underestimations.

3.8.2 Spelling accuracy

Children's average scores of text spelling accuracy appear to increase with time (Table 3.4). At every time point the Spanish group obtained higher mean percentages of correctly spelled words than the English group, displaying mastery levels of 85.13% mean accuracy by Time 5.

Table 3.4

Spelling accuracy: Mean percentage and SD of conventionally spelled words

	Time 2		*Time 3*		*Time 4*		*Time 5*	
Language	*M*	SD	*M*	SD	*M*	SD	*M*	SD
English	33.30	(25.31)	46.92	(23.88)	62.85	(21.43)	69.07	(19.17)
Spanish	—		54.60	(26.62)	80.99	(12.68)	85.13	(10.16)

3.8.3 Segmentation

Children's awareness and knowledge of conventional boundaries between words appeared to improve considerably over time (Table 3.5). The Spanish group showed higher mean percentages than the English group at every time point, reaching near-ceiling performance by Time 5. Analyses of word segmentation error patterns were conducted for each language group separately, to detect language-specific trends. Repeated measures ANOVAs showed a significant effect of error type at all time points and language groups—for English, $F(2, 342) = 156.92$, $p < .001$, $\eta_p^2 = .48$; $F(2, 344) = 205{,}41$, $p < .001$, $\eta_p^2 = .54$; $F(2, 340) = 104.53$, $p < .001$, $\eta_p^2 = .38$; $F(2, 330) = 57.87$, $p < .001$, $\eta_p^2 = .26$, at Times 2, 3, 4, and 5, respectively; for Spanish, $F(2, 332) = 166.36$, $p < .001$, $\eta_p^2 = .50$; $F(2, 320) = 73.23$, $p < .001$, $\eta_p^2 = .31$; $F(2, 3378) = 50.77$, $p < .001$, $\eta_p^2 = .23$, at Times 3, 4, and 5, respectively. Planned contrasts showed that, in all cases, hyposegmentations were more frequent than both hypersegmentation and mixed hypo- and hypersegmentation errors. Hypersegmentations comprised less than 2% of all words across time points and language groups, and “mixed” error type never reached an average of 1% in either language. Given that error patterns were very similar in both languages, only the percentage of correctly segmented words will be considered in subsequent analyses.

Table 3.5

Word segmentation: mean percentages and SDs

Type of segmentation	*Time 2*		*Time 3*		*Time 4*		*Time 5*	
	M	SD	*M*	SD	*M*	SD	*M*	SD
					English			
Conventional	36.82	(38.13)	47.34	(37.16)	67.97	(33.39)	79.31	(26.81)
Hypo.	32.34	(32.91)	36.42	(32.33)	22.27	(27.19)	14.64	(22.86)
Hyper.	0.85	(3.24)	0.76	(2.56)	1.50	(2.78)	1.83	(3.05)
Mixed	0.27	(1.59)	0.82	(3.26)	0.50	(1.52)	0.34	(1.16)
					Spanish			
Conventional	—		53.80	(34.37)	87.94	(14.01)	91.92	(10.38)
Hypo.			33.02	(31.80)	8.75	(12.08)	5.14	(8.52)
Hyper.			1.23	(2.99)	0.92	(2.47)	0.85	(2.54)
Mixed			0.66	(3.07)	0.22	(1.08)	0.13	(0.71)

Note. Hypo. = hyposegmentation; Hyper. = hypersegmentation; Mixed = mixed hypo-and hyper-segmentation.

3.8.4 Case

Children’s knowledge of the different units of their notational system, that is, upper- and lowercase, and of their contexts of use also appeared to increase with time (Table 3.6). Spanish children showed near-mastery levels (exceeding 80%) by Time 4, while their English counterparts approached mastery in this skill by Time 5.

3.8.5 Morphology

Although reporting past events elicited past tense formation in both languages, the open-ended aspect of the task made it impossible to control for the number of morphologically relevant words in each text. Consequently, the final number of relevant items for this analysis was low. Table 3.7 shows that the percentage of relevant words was similar across time points and languages (ranging from 10% to 13%) but, considering the average number of words written at each time point (see Section 3.7.6 below), the average number of morphology-relevant contexts amounted to only a few words at most.

Some trends are nevertheless apparent from the data reported here. At Time 2 more than half of all relevant inflectional suffixes were simply omitted. The percentage of omitted suffixes appeared to decrease with age, mostly in favour of phonetic representations, rather than of partial orthographic representations. The percentage of full orthographic representations of morphological endings, in contrast, appeared to increase slowly. This seems to suggest that children may go through at least three stages to represent inflectional morphemes in writing: Firstly, they omit the morpheme altogether; secondly, they represent it using plausible phoneme-to-grapheme correspondences; finally, they represent these morphemes conventionally. An intermediate phase between a phonetic and a conventional (full orthographic) representation would be one where children write only some of the conventional letters (e.g., <wachd>), thus showing awareness of the fact that phonographic correspondences are not sufficient to represent those parts of words. However, the data presented here do not seem to indicate a pattern by which children would necessarily go through such a stage.

In the case of the Spanish group, omission was not an option, and instead children needed to learn about the inconsistent letter-sound mappings of underlying morphological spellings; therefore, morphology-relevant words were either correctly (conventionally) or incorrectly (usually, phonetically) written. While at the earliest time point (Time 3) children were writing only approximately half of the inconsistent suffixes conventionally, by Time 5 over 80% of them were written correctly.

Clearly, the data were insufficient to carry out more complex analyses to elucidate children's awareness of morphological marking during composing. A more focused investigation of the development of the representation of morphology in text production would require an experimental design aimed to elicit the target forms. For these reasons, the measure of representation of morphology will not be kept in subsequent analyses.

Table 3.6

Mean percentages and SDs of words with conventional use of case

	Time 2		*Time 3*		*Time 4*		*Time 5*	
Language	*M*	SD	*M*	SD	*M*	SD	*M*	SD
English	49.00	(32.50)	64.84	(26.74)	74.57	(22.88)	77.13	(18.97)
Spanish	—		64.15	(32.85)	90.83	(12.63)	94.20	(5.43)

Table 3.7

Mean percentages and SDs of the representation of morphology

	Time 2		*Time 3*		*Time 4*		*Time 5*	
Type of representation	*M*	SD	*M*	SD	*M*	SD	*M*	SD
				English				
Not represented	**50.77**	(46.45)	**37.03**	(40.25)	17.90	(29.28)	16.35	(28.78)
Phonetic	29.24	(40.32)	31.52	(36.27)	27.46	(34.45)	30.36	(35.55)
Partial orthographic	7.60	(22.24)	12.04	(24.55)	16.00	(25.85)	10.34	(20.41)
Full orthographic	12.40	(28.43)	19.41	(31.31)	**38.65**	(39.36)	**42.96**	(38.03)
Relevant contexts[1]	10.66	(14.58)	13.86	(9.69)	10.02	(6.97)	10.83	(6.20)
				Spanish				
Phonetic			48.32	(40.75)	18.25	(33.06)	17.38	(26.57)
Full orthographic	—		51.68	(40.75)	81.75	(33.06)	82.62	(26.57)
Relevant contexts[1]			12.45	(11.99)	11.85	(9.00)	10.57	(7.28)

Note. Sample sizes were reduced because not all children produced morphologically relevant contexts in their texts. Sample size for English at Time 2 = 98, at Time 3 = 152, at Time 4 = 148, and at Time5 = 159. Sample sizes for Spanish at Time 3 = 116, at Time 4 = 135, and at Time 5 = 152. [1] = Percentage of words presenting a relevant context for the assessment of the representation of morphology.

3.8.6 Text length

With time children's texts in both language groups appeared to become longer (Table 3.8). There was a steady increase in the average length of the texts they were producing, which was attested by word and letter counts. Recall that only analyzable words were counted. Both word and letter counts appeared to offer very similar trends across time points and language groups. The correlations between them were extremely high in English and Spanish, with Pearson's r values ranging from .94 to .98, $p < .001$. This means that the number of words and number of letters measures seem to be tapping the same construct. Although the effect of orthographic consistency will be tested in Chapter 4 (Study 2), the mean number of words produced at each time point appears to be remarkably similar in both language groups (Table 3.8).

Table 3.8

Text length: mean number and SDs of words and letters

	Time 2		*Time 3*		*Time 4*		*Time 5*	
Unit counted	*M*	SD	*M*	SD	*M*	SD	*M*	SD
				English				
Words	7.58	(6.42)	15.49	(8.65)	24.85	(13.72)	31.69	(15.90)
Letters	27.83	(24.24)	55.82	(31.19)	91.12	(50.03)	117.35	(59.78)
				Spanish				
Words	—		13.98	(8.33)	23.84	(8.27)	34.06	(13.23)
Letters			52.65	(29.14)	88.49	(28.95)	126.64	(49.60)

3.8.7 Punctuation

The use of punctuation marks in children's texts was analyzed in two ways. First, I looked at the overall use of punctuation marks. To control for the influence that the length of the text might have on this count, the total number of punctuation marks was divided over the number of words produced in each text and multiplied by 100. Descriptive statistics showed that the use punctuation marks was very low in both language groups, but especially in Spanish (Table 3.9). However, and controlling for the number of words in the texts, their use appeared to increase with time in both languages. These apparent trends will be examined statistically in the next chapter.

Texts were also analyzed qualitatively as for the presence/absence of types of punctuation marks in the texts produced by each language group. The use of the different types of punctuation marks—full stop and the end of the text, full stop(s) within the text, and other punctuation marks—were recoded as dummy variables.

The aim of this analysis was to fine-grain the more global quantitative analyses, looking at the patterns of use of specific types of punctuation marks at each time point and language. Table 3.10 shows that, similarly to the quantitative descriptive results, the overall frequency of use of all types of punctuation marks appears to increase with time in both language groups. While at Times 2 and 3 the majority of the texts showed no punctuation marks at all, by the end of Year 1 (Time 4) most texts (around 60%) in both languages included at least one instance of punctuation marks. It is also apparent from Table 3.10 that children almost never used punctuation marks other than the full stop, with only a handful of texts at Time 5 including other types of punctuation marks—usually, commas. The developmental pattern in the use of punctuation marks was assessed next. Non-parametric McNemar tests with Yates' continuity correction were run to compare the preference in each language group and time point for using a full stop at the end of the text as opposed to within the text. In Spanish, children more readily used using a full stop to mark the end of the text, rather than to mark within-text boundaries at Time 4, χ^2 (1, N=161) = 30.67, p < .001, and at Time 5, χ^2 (1, N=170) = 15.41, p < .001, but not at Time 3, where no significant differences were observed. The use of a full stop to signal the end of the text prior to the use of full stops within the texts is a finding in line with prior studies in Spanish and other Romance languages (e.g., Ferreiro & Pontecorvo, 1999). No such distinction was attested for English, where children used within-text and end-of-text full stops to a similar extent from the earliest time points. It should be noted, however, that McNemar tests do not take into account ties for the calculation of statistical significance; that is, texts where children used *both* full stop(s) within and at the end of the texts were excluded from the formula, as were cases in which no full stops were used. A detailed look at these instances, however, reveals that they constituted the vast majority at most time points in both language groups: The number of ties out of all texts in English was 153/172, 135/173, 119/171, and 96/166 at Times 2, 3, 4, and 5, respectively. In Spanish, the figures were 152/167, 92/161, and 95/170, for Times 3, 4, and 5, respectively.

Table 3.9

Punctuation (quantitative assessment): mean percentages and SDs of punctuation marks over total of analyzable words

	Time 2		*Time 3*		*Time 4*		*Time 5*	
Language	*M*	SD	*M*	SD	*M*	SD	*M*	SD
English	3.89	(16.61)	5.56	(10.85)	5.78	(6.55)	6.57	(6.60)
Spanish	—		2.02	(9.23)	4.53	(5.68)	5.47	(4.99)

In sum, children at the initial stage of writing development seem to make very scarce use of punctuation marks. The combined quantitative and qualitative picture reveals that near-floor effects characterize the use of punctuation at Times 2 and 3 on both languages, but also indicates that most children are using punctuation marks by the end of the study (Table 3.10). In any case, these preliminary observations call for caution in the interpretation of subsequent analyses. These results also show that punctuation use at the initial stages of writing appeared to be restricted to the full stop. This overview of the early use of punctuation marks supports adopting a quantitative approach, especially in light of the lack of distinct preferences for the use of full stops to signal different text units, particularly in English. Further analyses will thus be carried out with the total number of punctuation marks, controlling for text length as explained above. The rationale for adopting a fully quantitative approach, without concern for the appropriateness of the different punctuation marks, follows from the understanding that, at these stages, the mere production of a punctuation mark in the texts should be regarded as an indication of a child's awareness of these items as part of the composition process. A normative approach, that is, one which assessed the use of punctuation marks penalizing unconventional uses, would most likely yield inconclusive results given the small amount of punctuation marks overall.

Table 3.10

Punctuation (qualitative assessment): Frequencies

	Time 2	*Time 3*	*Time 4*	*Time 5*
	% of texts (n)	% of texts (n)	% of texts (n)	% of texts (n)
		English		
No PMs	83.10 (143)	61.30 (106)	38.60 (66)	24.10 (40)
End-of-text FS	11.60 (20)	30.10 (52)	45.60 (78)	56.00 (93)
Within-text FS	11.00 (19)	24.30 (42)	45.60 (78)	53.60 (89)
End + within	5.8 (10)	16.20 (28)	30.40 (52)	33.70 (56)
Other PMs	0.00 (0)	0.60 (1)	1.80 (3)	4.80 (8)
		Spanish		
No PMs		86.80 (145)	41.60 (67)	21.80 (37)
End-of-text FS		9.60 (16)	51.60 (83)	64.70 (110)
Within-text FS	—	3.60 (6)	15.50 (25)	32.40 (55)
End + within		6.60 (11)	22.40 (36)	44.10 (75)
Other PMs		1.80 (3)	5.00 (8)	11.80 (20)

Note. PMs = punctuation marks; FS = full stop; End + within = full stop(s) at the end *and* within the text.

3.8.8 Layout

The arrangement of words in children's texts was assessed by looking at the percentage of texts that were classified into the different layout types at each time point and language group (Table 3.11). It was assumed that a paragraph-like arrangement would be the optimal layout, as it reflects the conceptualization, on the part of the child, of the text as a unit. However, other types were possible, especially the report of events in a list format. The resulting three-point scale thus ranged from children writing isolated words or fragments (lowest end); writing events as a series of isolated sentences, as in a list (mid-point of the scale); to children writing in a paragraph-like manner (top-end). Children who wrote less than two lines of text did not allow for a proper evaluation of the chosen layout type. Texts in this category accounted for 3% or less of all texts by Time 5. Other types of layouts, which were impossible to classify in the above-mentioned categories, were extremely rare at all time points. Fisher's exact Chi-square tests were run to detect above-chance distributions of texts among the different layout types (categories 1, 2, and 3) at each time point and language group. These tests were significant for English—Time 2: χ^2

(2, N=122) = 27.52, p < .001; Time 3: χ^2 (2, N=163) = 42.6, p < .001; Time 4: χ^2 (2, N=163) = 64.69, p < .001; Time 5: χ^2 (2, N=157) = 92.82, p < .001—and Spanish—Time 3: χ^2 (2, N=143) = 29.34, p < .001; Time 4[6]: χ^2 (1, N=158) = 57.51, p < .001; Time 4: χ^2 (1, N=164) = 125.64, p < .001. A quick look at the frequencies of layout types, indicated that children were divided between the sentence list (category 2) and paragraph-like (category 3) formats. Therefore, 2x2 Chi-square tests with significance levels set at p < .01 were conducted to test for this particular preference. In English, the distribution of children's text layout types revealed no significant preference for one particular layout at Times 2 or 3, but they showed an above-chance preference for the paragraph layout at Time 4, χ^2 (1, N=159) = 14.73, p < .001, and Time 5, χ^2 (1, N=156) = 40.71, p < .001. The Spanish group showed a strong preference for the paragraph layout from the earliest time point: Time 3, χ^2 (1, N=127) = 10.27, p = .001; Time 4, χ^2 (1, N=158) = 57.51, p < .001; and Time 5, χ^2 (1, N=162) = 77.03, p < .001.

Table 3.11

Frequencies in each language group and time point of different text layout types

	Time 2	*Time 3*	*Time 4*	*Time 5*
Layout types	% of texts (n)	% of texts (n)	% of texts (n)	% of texts (n)
		English		
One sentence/line only	27.90 (48)	4.60 (8)	2.30 (4)	3.00 (5)
Isolated words/fragments	4.70 (8)	5.80 (10)	2.30 (4)	0.60 (1)
List of sentences	**34.90** (60)	**56.10** (97)	26.90 (46)	15.10 (25)
Paragraph	31.40 (54)	32.40 (56)	**66.10** (113)	**78.90** (131)
Other layouts	1.20 (2)	1.20 (2)	2.30 (4)	2.40 (4)
		Spanish		
One sentence/line only		14.40 (24)	1.90 (3)	0.60 (1)
Isolated words/fragments		9.60 (16)	0.00 (0)	1.20 (2)
List of sentences	—	22.80 (38)	10.60 (17)	5.90 (10)
Paragraph		**53.30** (89)	**87.60** (141)	**89.40** (152)
Other layouts		0.00 (0)	0.00 (0)	2.90 (5)

[6] A 3x2 Chi-square table could not be tested because one of the cells had a frequency lower than 5 (layout type 1 = isolated words/fragments). The 2x2 table was ran instead and shown here for consistency.

These results seem to support the validity of the scale; in particular, they support the initial assumption that children progressively adopt the paragraph-like layout to report recent past events. For this reason, this variable will be recoded as a dummy variable (determining whether the child followed a paragraph-like layout or not), so that it can more easily used in a broader range of statistical analyses.

3.8.9 Connectives

Table 3.12 shows the mean percentages for each type of connective over the total number of words. The most frequent type of connective were the basic conjunctions *and, or,* and *but* (*y, o, pero* in Spanish, respectively), of which *and* comprised virtually all occurrences. The preference for *and* (*y,* in Spanish) was tested by comparing its overall production rate with that of all other connectives, which were collapsed into a single category. The preference for conjunctions over other connectives held for all time points and language groups: for English at Time 2, $t(146) = 2.71$, $p = .008$; Time 3, $t(166) = 5.17$, $p < .001$; Time 4, $t(164) = 7.07$, $p < .001$; Time 5, $t(163) = 4.83$, $p < .001$; and for Spanish at Time 3, $t(159) = 4.80$, $p < .001$; Time 4, $t(160) = 10.01$, $p < .001$; Time 5, $t(169) = 10.03$, $p < .001$.

Table 3.12

Mean percentages and SDs in the use of connectives

Types of connectives	*Time 2*		*Time 3*		*Time 4*		*Time 5*	
	M	SD	*M*	SD	*M*	SD	*M*	SD
				English				
and, or, but	2.63	(8.97)	3.33	(6.15)	4.47	(5.24)	5.28	(5.05)
Other	0.38	(2.01)	0.64	(2.07)	0.88	(2.36)	1.96	(3.11)
Sub. conj.	0.33	(1.80)	0.27	(1.39)	0.42	(1.29)	0.97	(1.95)
				Spanish				
y, o, pero			4.37	(5.57)	6.99	(5.12)	7.28	(4.51)
Other	—		1.47	(4.10)	1.95	(3.65)	2.30	(2.93)
Sub. conj.			0.28	(2.11)	0.28	(1.07)	0.80	(1.75)

Note. Other = other connectives; Sub. conj. = subordinating conjunctions.

Table 3.13

Distribution in the use of connectives

Types of connectives	*Time 2*		*Time 3*		*Time 4*		*Time 5*	
	M	SD	*M*	SD	*M*	SD	*M*	SD
				English				
and, or, but	72.15	(41.02)	76.60	(36.23)	77.48	(34.00)	62.45	(35.50)
Other	13.38	(29.61)	13.78	(27.70)	14.24	(28.25)	22.95	(31.04)
Sub. conj.	14.47	(34.70)	9.62	(27.53)	8.28	(20.99)	14.60	(26.31)
Total (n)[1]	100 (38)	—	100 (65)	—	100 (104)	—	100 (134)	—
				Spanish				
y, o, pero	—		76.24	(34.78)	79.99	(28.61)	71.61	(28.24)
Other			20.38	(32.08)	17.21	(27.59)	20.18	(24.58)
Sub. conj.			3.38	(16.33)	2.80	(10.39)	8.21	(18.15)
Total (n)[1]			100 (84)	—	100 (132)	—	100 (157)	—

Note. Other = other connectives; Sub. conj. = subordinating conjunctions.

[1] Sample sizes used for this analysis are smaller because only children who had produced at least one connective were included.

The distribution of the different types of connectives was explored to see whether the relative weight of each specific kind of connective changed over time in each language. Table 3.13 shows that conjunctions *and, but, or* were preferred the most across time points and language groups. Similarly to the overall production of connectives, conjunctions were compared to all remaining categories of connectors. Here, however, the relative distribution of connectives into the two categories (basic conjunctions vs. other connectives) was compared only for children whose texts included connectives of some sort (see the actual sample sizes in the "Total" row on Table 3.13). Results showed that conjunctions were the most popular category in English at Time 2, $t(37) = 9.34$, $p = .001$, Time 3, $t(64) = 15.28$, $p < .001$; Time 4, $t(103) = 20.84$, $p < .001$; and Time 5, $t(133) = 17.76$, $p < .001$. A similar picture emerged for Spanish at Time 3, $t(83) = 16.77$, $p < .001$; Time 4, $t(131) = 27.78$, $p < .001$; and Time 5, $t(156) = 27.68$, $p < .001$. Note that the number of children that made use of connectives was higher at each subsequent time point in both languages—from only 38 children at Time 2 to 134 at Time 5, in English, and from 84 children at Time 3 and 157 at Time 5, in Spanish.

In sum, both languages showed an increase in the use of connectives with time and displayed a similar distribution of the different types of connectives at every time point, revealing no clear language-specific trends. Similarly to punctuation, then, the mere occurrence of connectives (more than their type or distribution) seems to be a valid way of measuring this aspect of text construction. For this reason, subsequent analyses involving connectives will use the percentage of connectives over the total number of analyzable words.

3.8.10 Words per clause

Table 3.14 shows that the Spanish group ranged from a minimum average of 5.61 to a maximum of 6.00 words per clause across time points. The English group ranged from 5.06 to 5.89 during the same period (Time 3 through to Time 5). This trend appears to show stability in this measure, rather than radical changes over time. The trend, as well as cross-linguistic differences, will be tested statistically in the next chapter.

3.8.11 Percentage of subordination

Very few children produced subordinated clauses in their texts. At Time 2 in English, only 19 texts included subordinated clauses. With time, however, more children appeared to incorporate them in their written productions. The number of children who did so at each time point was remarkably similar across language groups. In English, 36, 46, and 75 children produced subordinated clauses at Times 3 through to 5. In Spanish, the figures were 37, 47, and 79. The subordination index was thus at floor or near floor levels, particularly at the earlier time points (Table 3.14). Because the percentage of subordinated clauses out of all clauses produced provides an in-depth analysis of syntactic complexity, complementary with the linear assessment provided by the word-per-clause measure, it will be retained in subsequent analyses. However, the scarcity of the data should be taken into account in the interpretation of the results.

Table 3.14

Means and SDs for the syntactic complexity measures

	Time 2		*Time 3*		*Time 4*		*Time 5*	
Measures	*M*	SD	*M*	SD	*M*	SD	*M*	SD
Words/clause[1]	4.85	(1.67)	5.08	(1.35)	5.79	(1.94)	5.60	(1.40)
Subordination (%)[2]	4.66	(13.53)	5.84	(11.66)	6.26	(11.66)	11.93	(15.05)
Words/clause[1]	—		5.17	(2.72)	5.96	(3.19)	5.72	(1.53)
Subordination (%)[2]			7.37	(15.64)	7.84	(13.65)	12.54	(14.92)

Note. Subord. index = Subordination index. [1]Number of analyzable words (Level of representation categories 7, 8, and 9) over total number of clauses. [2] Percentage of subordinated clauses over total number of clauses.

3.8.12 Lexical diversity

The ratio of unique words (types) to all words produced (tokens) decreased with time in both languages (Table 3.15)— $F(2.84, 396.87) = 50.14$, $p < .001$, $\eta_p^2 = .264$, in English; $F(2, 260) = 52.75$, $p < .001$, $\eta_p^2 = .289$, in Spanish. This result is counter-intuitive, because one would expect an increase in children's diversification of vocabulary choices and not the reverse. However, the short length of the texts might have limited the sensitivity of this measure to detect variations in lexical diversity, given its dependence on text length (e.g., Malvern et al., 2004). Short texts usually result in high ratios since, the shorter the text, the less the opportunities to repeat words; contrarily, longer texts typically show lower values: the more words, the more room for repetition. Children in this study appear to have increased the number of tokens they produced with time, but the rate of lexical diversification did not seem to progress at the same rate.

In addition, independent-samples *t*-tests with the Bonferroni correction for multiple comparisons applied (significance level set at $p < .025$) showed that the Spanish group obtained higher type-token ratios than the English group across time points: $t(323.85) = 8.23$, $p < .001$, $d = .91$, for the Time 3 comparison; $t(324) = 5.64$, $p < .001$, $d = .63$, for the Time 4 comparison; and $t(334) = 2.70$, $p = .007$, $d = .29$, for the Time 5 comparison. This could be indicating a better performance in vocabulary choices by the Spanish children, but it could also be a byproduct of language-specific phenomena (Berman, 2008). Indeed, the non-pro-drop parameter of English (Section

2.2.3) could be biasing the results. In these short personal accounts of recent events, children will typically need to refer to the protagonist of these events (usually, themselves), as well as to other animated and unanimated participating entities. In Spanish, subject pronouns may be omitted in most circumstances, with information about the subject of an action being retrievable from the verbal inflection; that is, bound morphology not constituting a separate word and, thus, not counting as another token. In English, in contrast, the same information needs to be repeated—for example, by constantly mentioning the first-person pronoun *I*—, thus inflating the final value of the denominator in the type-token ratio calculation.

The type-token ratio, however, may be useful in within-language analyses, despite its potential for cross-linguistic biases. For this reason, the measure will be used in within-language analyses.

Table 3.15

Means and SDs of vocabulary measures in each language group and time point

Vocabulary measures	*Time 2*		*Time 3*		*Time 4*		*Time 5*	
	M	SD	*M*	SD	*M*	SD	*M*	SD
				English				
Lex. div.	.851	(.156)	.735	(.153)	.706	(.131)	.695	(.113)
Lex. dens.	39.91	(24.34)	49.64	(16.95)	43.22	(12.17)	45.67	(9.22)
Adj. + Adv.	1.81	(4.79)	2.36	(4.39)	4.70	(5.69)	6.12	(5.36)
Word length	4.21	(2.32)	4.74	(1.34)	4.86	(1.15)	4.85	(0.91)
				Spanish				
Lex. div.			.856	(.117)	.780	(.107)	.727	(.107)
Lex. dens.	—		49.53	(19.17)	44.50	(8.53)	43.97	(6.79)
Adj. + Adv.			2.03	(3.77)	3.63	(4.69)	4.60	(4.24)
Word length			5.08	(1.39)	5.40	(0.69)	5.35	(0.53)

Note. Lex. div. = Lexical diversity (type-token ratio); Lex. dens. = lexical density (percentage of content words); Adj. + Adv. = percentage of adjectives and adverbs; Word length = average content-word length in letters.

3.8.13 Lexical density

The average percentage of content words (nouns, verbs, adjectives, and adverbs) out of all words in a given text is presented in Table 3.15 for each language and time point. Both language groups appear to be characterized by stability in the proportion of content words used in their texts. The interest of this measure may then lie in its relationships to other measures in the study, particularly those that tap into vocabulary. Therefore, and given that no evident floor or ceiling effects were attested, this measure will be kept in subsequent analyses.

3.8.14 Adjectives and adverbs

Adjectives and adverbs were produced very scarcely in children's texts throughout the study (Table 3.15). The small percentage of adjectives and adverbs in children's early compositions is, however, not markedly different from quantities reported in other studies. For example, Ravid and Levie (2010) reported an average of 0.12 adjectives per clause in the narrative written texts produced by 4th grade, Hebrew-speaking children. Since their texts were of an average length of 18.44 clauses or 74.19 words, the estimated mean percentage of adjectives per text was thus 2.98%—only slightly higher than the means reported here for younger participants. Llauradó and Tolchinsky (in press) reported 3% of adjective use in 1st grade Catalan-speaking children across six different text types. The slightly higher percentage found by Llauradó and Tolchinsky might then be attributable to the fact that the figure of 3% corresponds to an average across 6 text types, which might elicit adjectives and adverbs to a different extent than the prompt used in the present thesis and which were moreover produced with no time constraints. In sum, although the absolute percentages of adjectives and adverbs are very low, they are very much in line with the overall rate of occurrence reported for other languages. Therefore, this measure will be retained in subsequent studies.

3.8.15 Average content-word length

The average content-word length, measured in letters (Table 3.15) was characterized by stability rather than change over time: English texts ranged from an average content-word length of 4.21 letters to 4.86 letters, while in Spanish the range was 5.08-5.40. These results also indicate that average content-word length was slightly longer for Spanish than for English. However, more than a real cross-linguistic difference in text production abilities, this result might be reflecting the average content-word length in each language. As mentioned in Section 2.2.2 above, Spanish contains a much higher proportion of polysyllabic words, while in English monophthongs are frequent, with long, multimorphemic words belonging to formal registers. Although English also has a more complex syllabic structure than Spanish—which could arguably have an impact on the average number of letters in the written domain—, it does not seem sufficient to balance out the shorter average word length. The results of subsequent analyses involving this measure should be interpreted with caution, particularly the interpretation of cross-language comparisons.

3.9 Psychometric Properties of the Measures Used in Subsequent Analyses

The general objective of this subsection is to report the results of the psychometric properties assessment of the measures that will be used at some point during the thesis. This means that some measures (e.g., representation of morphology or use of commas) were not included in this assessment (see Table 3.16 for the final list of measures). The assessment was carried out in two steps: First, outlier scores were identified and corrected; second, the normality of the distribution for each variable was evaluated. A total of 42 variables were assessed: all variables in Table 3.16 (excluding the dummy variable for layout), at Times 3 through to Time 5, which are the times in which cross-linguistic comparisons may be performed.

Following Kline (2011, p. 54-55), the outliers assessment consisted in looking for cases that were > 3.0 standard deviations (SDs) above or below the mean for all relevant variables. One hundred and forty-four scores were identified across the 36 variables, M = 4.00, SD = 3.27, range: 0-17. In other words, taking into account the sample size, outliers amounted to 1.06 % of all scores (range: 0-4.49%).

It was important to distinguish between univariate outliers (i.e., cases with only one extreme score) and multivariate outliers (i.e., cases with multiple extreme scores), in order to determine whether there were cases that may not belong to the same population as the rest. Therefore, the distribution of extreme scores within single cases was explored. Ninety-eight cases showed at least one extreme score, most of whom (n = 73) affected a single variable; that is, 73 children produced only one extreme score across all variables and time points. Of the remaining cases, 17 showed 2 extremes scores, and 8 showed three or more extreme scores. It was possible that this last group of 8 children with multiple extreme scores belonged to a different population. However, an exploration of each of these children's linguistic and cognitive profiles—which had been obtained in the context of the larger WP1 project—it was determined that there was no clear evidence suggesting that these children belonged to a different population, especially with regards to their spelling and writing profiles. In addition, the scope of the larger project, as well as of the present thesis, was to recruit an unselected sample, representative of the real population found at schools (Caravolas et al., 2012). For these reasons, no child was excluded from subsequent analyses on this basis.

In order to avoid extreme scores from distorting the distribution of values within a given variable, scores were brought within the ±3 SDs range; that is, scores deviating 3 SDs or more from their language group mean were replaced by the value equivalent to exactly ±2.99 SDs, the negative or positive value depending on their original place in the distribution.

The second step in the assessment of the psychometric properties of the text writing measures was to check for the normality of the distribution for each variable after having corrected extreme scores as specified above. Given the large sample sizes, significant testing of the normality of distribution (e.g. Shapiro Wilk or Kolmogorov-Smirnov tests) are considered to be too sensitive even to minor deviations from normality (Field, 2009). Plots and histograms, on the other hand, though valuable to perform an initial assessment, are deemed as a highly subjective means for evaluating deviations from normality (Field, 2009). Kline (2011) suggests looking for skewness values ≥ 3.0, and for kurtosis values ≥ 10. None of the selected measures (Table 3.16) showed skweness or kurtosis values outside the above-mentioned cut-off points. Skewness values ranged from -1.72 to 1.86, while kurtosis values ranged from -1.58 to 4.63. In short, no major deviations from normality were observed for the final set of measures.

Table 3.16

Final set of variables selected for subsequent analyses

Variable name	Description	Calculation
Transcription		
TSPELL	Conventional text spelling accuracy	% of words spelled correctly
SEG	Conventional word segmentation accuracy	% of words showing conventional boundaries
CASE	Accuracy in the use of upper and lowercase letters	% of words showing a conventional use of upper and lowercase use
Productivity		
WORDS	Text length in words	Raw number of analyzable words
LETT	Text length in letters	Raw number of letters
Text structure		
LAYOUT	Classifies the text as to whether it has a paragraph layout or not	Dummy variable where 1=paragraph layout and 0=other layouts.
PUNCT	Use of punctuation marks	Number of punctuation marks over total number of words multiplied by 100
CONN	Use of connectives (conjunctions and discourse markers)	% of all inter-clausal connectives over total number of words
Syntactic complexity		
WCL	Measure of linear syntactic complexity	Ratio of words per clause
SUBD	Measure of syntactic embeddedness	% of subordinated clauses over total number of clauses
Vocabulary		
LEXDIV	Measure of lexical diversity	Ratio of unique words (types) to total words (tokens)
LEXDEN	Use of content words (nouns, verbs, adjectives, and adverbs) in texts	% of content words over the total number of words
ADJADV	Use of optional lexical items that contribute to the detail and precision in texts	% of adjectives and adverbs over total number of words
WLEN	Estimate of content-word sophistication	Average content-word length in letters

3.10 Discussion of Descriptive Results and Implications for Subsequent Studies

One of the aims of this chapter was to provide the reference framework within which the origin and scope of this thesis should be interpreted. In this sense, the strong points of this series of studies are its large sample sizes, and its longitudinal and cross-linguistic nature. It is particularly this last aspect—the cross-linguistic component—what enabled the exploration of current thinking about writing development, which is to a large extent based on English-only studies, to determine whether it may be extended to other orthographies.

In addition, a further aim was to provide a thorough description of the characteristics of the main task and procedures that were followed to obtain the target text-based writing measures. It was stressed that the focus of the thesis is on microstructural, rather than macrostructural, text features. This choice was founded on the fact that microstructural features are easier to operationalize and quantify in an objective, systematic way. Moreover, microstructural writing features are of interest from the point of view that they are usually the means through which macrostructural writing quality is achieved.

A number of word- and text-level variables was explored, to determine the most informative ones for the developmental period under scrutiny. Given that children generally produced very short texts, subgroups or sub-classifications were avoided. Therefore, instead of distinguishing, for example, between subtypes of connectives (e.g., between conjunctions and discourse markers), subsequent studies will consider all connective types together. Similarly, although different types of word segmentation errors were identified, only percentage accuracy in showing conventional word boundaries will be taken into account. This procedure resulted in adequate distributional properties of the final set of measures, with most extreme scores limited to a single case and variable over the three time points under consideration. In short, the initial exploration of the texts produced over the first year of schooling (year and a half, for the English sample) supported the adoption of a fully quantitative approach to the study of early text composition skills.

What the descriptive results reported in this chapter have shown indisputably is that, toward the middle of Year 1 (or the end of Reception Year, in the case of the English group) children have already acquired a wealth of knowledge about text construction, well beyond the spelling of single words. The vast majority of the texts collected were interpretable and there were sufficient phoneme-to-letter correspondences to enable the mapping of the children's strings of letters to their intended meanings. The level of representation assessment also showed that unanalyzable letter strings were negligible, thus validating the task procedure.

Children's conventional spelling accuracy was well off-floor in both language groups, and so was their knowledge of conventional word segmentation and of the appropriate use of case. The qualitative analysis of word segmentation errors indicated that the developmental pattern described for Romance languages—Spanish and Italian (Ferreiro and Pontecorvo, 1999)—also holds for English: Children began by producing continuous letter strings, with few or no word boundaries, as shown by the pervasiveness of hyposegmentation errors. Other types of errors—namely, hypersegmentation, and mixed hypo- and hypersegmentation—were rare. These findings, however, contrast with those obtained for Portuguese by Correa and Dockrell (2004), who found that hypo- and hyper-segmentation errors were about equally frequent in the texts of Year 1 children, though not in those of Year 2 or Year 3. With both language groups in this thesis displaying the same pattern of word segmentation strategies, only the percentage of correct segmentations at each time point will be considered in subsequent studies.

The assessment of the representation of inflectional morphemes in early written texts was limited in its scope and potential for generalizations, since the average percentage of morphology-relevant words was never higher than 13%, involving and average range of 2 words per text. Perhaps the reduced time span in which children were allowed to produce texts impacted on overall text length and, consequently, on the number of morphology-relevant contexts (see Section 3.11 on Limitations below). It is also possible that children adopted avoidance strategies and thus a more controlled task environment might be required to elicit the representation of inflectional morphology.

Children's texts showed a tendency to become longer with time. In spite of the languages differing in their typical syllabic structure and mean word length (Section 2.2.2 above), children's average number of words or letters at each time point was remarkably similar. The next series of statistical tests will clarify these preliminary trends. In addition, whether length is measured in words or letters did not alter the pattern of results obtained so far.

The extent to which children were able to produce cohesive, well-structured texts was evaluated via (1) their use of punctuation marks, (2) their use of connectors, and (3) the type of text layout chosen at each time point (isolated words, lists of sentences, or paragraph). In the case of punctuation marks, the descriptive statistics indicated that only by T4—end of Year 1—does the use of punctuation marks become generalized in both languages. By mid-Year 1 (Time 3) the majority of children were not including punctuation marks in their texts at all in either language. This means that we are witnessing the very emergence of this feature in text composing in both languages. No language-specific trends were observed in the pattern of use of punctuation marks, which was characterized by an overwhelming preference for the full stop, and the general sparse use. In view of these results, the mere occurrence of punctuation marks in a text was considered to suffice to assess this aspect of text composing. Future research should aim at discovering patterns of incorrect use of punctuation marks, for which the elicitation of longer texts may be essential. Children's use of connectives of all types to link the different text units was productive from the earlier time points in each language group and it showed a tendency to increase with time. An analysis of preferences in the use of the different types of connectors showed no language-specific trends: in both languages conjunction *and* (*y* in Spanish) prevailed over the rest of possible connectives. These results coincide from findings from early oral narratives (e.g. Berman & Slobin, 1994). Again, a quantitative approach seems most adequate to assess children's ability to link text units in writing, without making distinctions between different types of connectivity devices. Finally, the layout chosen by children to report recent past events also showed a developmental pattern by which children initially report them in lists of isolated sentences to only later produce a more narrative-like text using a paragraph layout. Results supported the hypothesis that children move from reporting events in a list format to progressively adopt a paragraph-like format.

With regards to syntactic complexity, this first exploration appears to indicate that children at these stages write mostly in simple clauses which vary little in their average length both over time as across language groups. Indeed, subordination was almost absent from their productions, especially in the earlier time points. In spite of these near-floor effects, both variables showed adequate distributional properties, and will therefore retained in ensuing studies.

Lexical aspects of children's texts were evaluated by (1) a measure of lexical diversity, the type-token ratio; (2) a measure of lexical density, defined as the percentage of content words; (3) a measure of the amount of evaluative comment in the text, represented by the percentage of adjectives and adverbs; and (4) an estimate of vocabulary frequency and sophistication, measured by the average content-word length in letters. All measures showed adequate psychometric properties, though different trends where observed among this rather reduced set of measures. On the one hand, the analysis of lexical diversity showed the reverse trend one might expect: Average type-token ratios decreased over time in both languages. However, it has been pointed out that this may well be a product of the lack of sensitivity of the measure itself in short texts, rather than a developmental finding. In addition, the cross-linguistic difference favouring Spanish should be interpreted with caution as the pro-drop nature of Spanish subjects may have biased results. The descriptive results for average content-word length may also be biased as a consequence of language-specific trends; in particular, the higher average word length in Spanish. Lexical density (percentage of content words) showed a tendency to stability across time points and languages. An increase in lexical density would be expected on the basis that texts should progressively be more informative and precise in their report of events, which requires a more profuse use of these word categories (i.e., nouns, verbs, adjectives, and adverbs). However, it is possible that even if such increase took place, it would be offset by the introduction of specific functional elements, such as connectors (see Section 3.8.9 above). Llauradó and Tolchinsky (in press) reported no clear developmental pattern in the lexical density of written texts composed by over 2,000 Catalan speakers, from kindergarten to the end of high-school[7]. No significant differences as a function of age were attested in the same time span as the one presented here (from preschool to Year 2). All in all, lexical density and diversity, as well as the average length of content words may require a longer span between time points to show reliable differences (Llauradó & Tolchinsky, in press). Finally, and in spite of their seemingly low incidence, the percentage of adjectives and adverbs was found to be in line with findings reported in studies with similar populations (e.g.,

[7] Catalan is a Romance language spoken in parts of Spain, especially, in the region of Catalonia. In terms of lexical aspects, it is comparable to Spanish.

Llauradó & Tolchinsky, in press; Ravid & Lavie, 2010) while showing a tendency to increase, which will be explored further in the next chapter.

To sum up, most of the measures obtained from the texts were found to be adequate for the study of early text composition in terms of their distributional properties and the reliability of measurement. Children were well off the floor for many text features, while other measures seemed to be in their initial developmental phases. Despite the multitude of additional demands of text writing, in contrast to single-word spelling, children managed to produce legible texts, and to convey at least simple messages, showing awareness of several untaught characteristic of written language.

3.11 Limitations

This chapter served to identify a number of limitations that are likely to affect subsequent studies. First, some key characteristics of the sample are either missing or incomplete, so one should be cautious before generalizing findings throughout the thesis to the general population. Parents' education level was reported for only about half the sample, and no information on socioeconomic status (SES) was available. SES is a major factor in shaping literacy outcomes (e.g., Bowey, 1995; Duke, 2000; Duncan & Seymour, 2000), and although maternal education may be considered a reliable proxy for SES (e.g., Raag, Kusiak, Tumilty, Kelmen, Bernheimer, & Bond, 2011; Stevens, Lauinger, & Neville, 2009), inability to fully control for its effect is likely to limit the scope for generalizations based on findings in this thesis.

It has been suggested that lefthandedness may vary geographically, with a higher prevalence among Asian people, although no reliable country-based rates exist (Raymond & Pontier, 2004). Left-handed children—particularly boys— are more likely to experience reading and writing difficulties (Latham, 2002). Therefore, it would have been optimal to control for left-handedness in these studies to lessen its direct or indirect effect on writing development.

Another limitation of the present design is related to the consequences of having a time-limited prompt. Certainly, texts produced by the children in both groups were very short, potentially restricting children's range of skills, as well as the analyses that could be performed on measures obtained from the texts—as was the case, for example, of the representation of morphology measure. As a result, further findings in this body of work are confined to this text type and prompt.

Finally, the lack of cross-linguistic differences in the distribution of male and female participants should not be taken to mean that this variable does not have an impact in writing achievements. Previous studies have indicated that gender accounts for a non-trivial portion of variance in writing outcomes (e.g., Olinghouse, 2008). The within-language effect of gender, however, was beyond the objectives and scope of the work included in this thesis, and adding it as an additional factor might lessen the statistical power of the key factors of the studies. In short, this series of limiting factors should be factored in during the interpretation of the results reported in the present and subsequent chapters.

Chapter 4

Study 2: The role of orthographic consistency in early writing development

Study 2 explored children's early text writing development as a function of the degree of consistency of the orthography they were learning. Current thinking about the development of written language has been determined, for the most part, on the basis of how it occurs in English, an inconsistent orthography. Therefore, a systematic comparison of English-speaking children to children learning to write in orthographies with different degrees of consistency should help determine the extent to which English-based notions may be extended to other alphabetic orthographies. The development of a set of microstructural features was thus explored over a period of one year, starting at mid-Year 1 and until mid-Year 2. Spanish children were clearly superior in transcription skills across the above-mentioned period, and their advantage over the English children only seemed to be enhanced with time. However, children in both language groups were remarkably similar in most other text features. Additionally, this study explored the microstructural dimensions underlying early writing. Exploratory factor analyses corroborated that word-level and productivity dimensions were well established from the earliest time point in both language groups. In addition, support for an emergent vocabulary dimension was found. However, an inconsistent pattern of relationships was found underlying the other text-level dimensions. In conclusion, learning to write in a consistent orthography does not result in an advantage in the composition of written text beyond the word-level of performance, at least in the early stages of learning to write. Moreover, transcription and productivity aspects seem to be more established at this stage of literacy development, in contrast to other text-level features.

4.1 Introduction

From very early on in their academic lives, children are expected to be able to produce written texts and to express themselves mainly through written language. Poor writing is likely to have knock-on effects on potentially all school subjects, given that writing quickly becomes the main or sole means both for accessing to information, as well as for evaluating a child's performance; it is harder (if not impossible) for teachers to assess a child's knowledge if his/her (written) communicating skills are weak (Graham & Perin, 2006). Mastering text composition is thus of vital importance in fostering a child's academic progress. Nevertheless, most studies of early literacy do not go beyond the word level, both for reading and writing. The intense focus on (word) spelling at the earlier stages of literacy learning finds its roots in the assumption that some degree of mastery or automatization of spelling is necessary for other aspects of the writing process to emerge and develop (e.g., Berninger, et al. 1991; 1992; Berninger & Swanson, 1994; Bourdin & Fayol, 1994; Juel, et al., 1986). Current models of writing place transcription skills (spelling and handwriting) at the base of the writing process (e.g., Berninger & Amtmann, 2003; Berninger & Winn, 2006), considering them the necessary platform that allows for the generation of content.

There is considerable evidence suggesting that high-level writing skills and some microstructural aspects of text construction depend upon low-level writing skills. Many studies have reported that beginner writers produce shorter texts, make more grammatical mistakes, show a weaker knowledge of punctuation and capitalization rules, and are less aware of writing conventions than children attending mid-primary school (Abbott et al., 2010; Puranik et al., 2008; Wagner, et al., 2011). Children compose longer and better texts when they are asked to dictate to an adult, as opposed to having to write the text themselves (e.g., Hayes & Berninger, 2009; Graham, 1990). These findings are consistent with the *cognitive load* hypothesis (Bourdin & Fayol, 1994), which states that the written modality is cognitively more costly than the spoken modality. Therefore, until transcription skills (spelling and handwriting; Berninger & Swanson, 1994) have reached a stage in which they have been fairly well automatized, they require a considerable amount of the finite pool of cognitive resources, with very little left for other aspects involved in text production[8]. On these assumptions, then, it is predicted that until children attain automatized spelling skill, the rest of their text composing abilities will be seriously compromised (e.g., Juel et al., 1986).

Most studies from which these ideas have spread were carried out with English-speaking participants, whose orthography is considered to be an "outlier", given the high degree of inconsistency of its phonographic mappings (Share, 2008). Cross-linguistic comparisons of *word* spelling and reading achievements have indicated that there are important differences as a function of orthographic consistency, where more consistent orthographies are learned more quickly than less consistent ones (e.g., Bruck & Caravolas, 1993; Caravolas et al., 2013; Seymour, et al., 2003). Arguably, if spelling skill presents the cognitive bottleneck of writing development, then children learning a simpler orthography should show higher levels of achievement in non-transcription aspects of writing sooner than children learning more a more inconsistent orthography; that is, spelling would become automatized earlier, thus freeing cognitive resources to be devoted to other aspects of text construction. Conducting cross-linguistic comparisons should help clarify whether the varying difficulty of the orthography influences the development of text composition at the word- (e.g., spelling) and text-level of performance. The main aim of this study was thus to establish the differences and similarities in the first stages of writing development between monolingual English- and Spanish-speaking children that might be attributable to the nature of the orthography that each group of children was learning. In addition, a secondary aim was to explore the dimensions underlying early text composition across languages, distinguishing between word-level/transcription and other text-level dimension, in order to establish at what level of performance cross-linguistic differences were found.

4.1.1 Global vs. Multidimensional Approaches to Writing

Since the writing process involves multiple skills and may be evaluated from several viewpoints and analyzed at various levels of performance, a great deal of effort

[8] But see Bourdin & Fayol (2002) for evidence that, even in adults, spelling still is quite cognitively demanding under certain conditions.

has been put to either structuring such complexity or simply trying to reduce it analytically. Conceptualizations of writing thus range from proposals like the *simple view* (Juel, et al., 1986), which conceives of writing as the product of spelling plus ideation, to a multi-faceted approach where several features of writing products (e.g., Puranik et al., 2008) and processes (e.g., Berninger et al., 1991) are considered. Notwithstanding the theoretical framework in use, the assessment of writing has typically been performed via global or holistic measures of *text* or *compositional quality* (e.g. Berninger et al, 1992; 1994; Graham, 1990; Graham et al., 1997), with differing degrees of success in achieving acceptable levels of inter-rater reliability. Even when reliability is high, it is still questionable whether the evaluation is a valid estimate of the target ability, particularly for beginner writers (Puranik & AlOtaiba, 2012; see also Berman & Nir, 2009, for a review of text quality evaluations). It has also been suggested that the type of assessment (holistic/global vs. multiple-features approach) of the text productions may have an impact on the type of model that best summarizes the data (Wagner, 2012). In this sense, it seems that unidimensional conceptualizations are usually validated when writing performance has been assessed globally (e.g., Mehta et al., 2005; Vellutino, Tunmer, Jaccal, & Chen, 2007), whereas multidimensional models are supported when text writing has been evaluated using multiple indicators (e.g., Puranik et. al, 2008; Wagner et al., 2011). In line with Study 1 (Chapter 3), a multiple-feature approach will be adopted in the present study.

4.1.2 Microstructural Features

In light of the main purpose of the present study, the measures obtained from the texts needed to comply with a number of criteria; these included (1) similarity across both languages, that is, tapping into the same skills on the basis of the same criteria; (2) quantitative measures, avoiding subjective ratings and evaluations that depend to a large extent on raters' writing style or that could not be applied systematically; and (3) comparability with measures used in research with older participants, so that findings from this study may be linked to other studies of writing across the life-span.

Individual microstructural features were classified into two broad categories: *word level* and *text level* of writing performance. Word-level writing skills were conceptualized as including writing features which have their main impact at the word and sub-word level and are typically studied in single-word spelling research. These included spelling, word segmentation, and use of case. In contrast, text-level writing skills were assumed to have an impact beyond the word level and to be key contributors to overall text quality. In addition, both word- and text-level features were organized into the following dimensions: ***transcription, productivity, text structure, syntactic complexity,*** and ***vocabulary***. Transcription included *spelling accuracy*, *word segmentation*, and *use of case*. Although spelling is a privileged indicator of transcription performance, children's knowledge of conventional word boundaries (word segmentation) might also prove essential for applying some orthographic rules. Correct use of case was evaluated not only because it is part of the conventions children that must master in order to communicate through writing, but also because the correct selection of the appropriate case reflects their knowledge of the notational system they are learning: children at these very early stages might not yet be aware of all pairs of lower- and upper-case letters (Ravid & Tolchinsky, 2002).

Productivity was assessed by the length in number of *words* produced as well as in number of *letters*. It has been repeatedly claimed that the amount of text produced is highly sensitive to children's transcription skills (e.g., Berninger et al., 1998; Graham, et al., 1997). The two different transcription units were used to ensure that language-specific traits, such as average word length, did not bias the results.

Text features that have been argued to contribute to text structure were also assessed. Individual indicators of this dimension were the use of use of *punctuation,* and *connectors.* Punctuation marks are typically regarded as part of the conventions of writing that children must learn (e.g., Puranik et al., 2008); however, punctuation marks signal the boundaries of different discourse segments and there is evidence that children acquire them in a predictable order (Ferreiro & Pontecorvo, 1999). Connectives (conjunctions and discourse markers) are fundamental to showing the links among parts of the text and thus contributing to text cohesion (Hickmann, 2003).

Syntactic complexity has been regarded as a key dimension of discourse elaboration and essential to the characterization of text composition skills across the life-span (Berman, 2008). Two measures of syntactic complexity were included: a syntagmatic measure of clause complexity, *words per clause,* and a measure of the degree of syntactic embeddedness, *percentage of subordination.*

Finally, an analysis of the vocabulary choices in children's written compositions serves to evaluate the lexical precision and the degree of elaboration of the written products. Based on preliminary findings (Study 1, Section 3) the measures included as indicators of this sub-component were *lexical density*— that is, the percentage of content words out of all words in a text—, based on the assumption that a greater proportion of semantically charged tokens results in a richer and more informative text (Halliday & Hassan, 1985; Malvern et al., 2004); *proportion of adjectives and adverbs,* based on the assumption that these optional lexical elements add to the elaboration and precision of the message (Ravid & Levie, 2010); *lexical diversity,* given that children may increase the number of words they produce, but this may not be accompanied by diverse and accurate vocabulary choices (Malvern et al., 2004); and *average content-word length,* which is used as a proxy for lexical sophistication and use of low-frequency tokens, based on the assumption that longer words are less frequent (Bybee, 2007). Note that these last two measures—lexical diversity and average content-word length—were used only in within-language analyses (correlations and exploratory factor analyses), due to the intrinsic bias they presented for cross-linguistic comparisons (see Section 3.8.12 and 3.8.15; also 3.10, for a discussion).

A secondary aim of this study was to examine whether the individual measures analyzed were reliable correlates of the various writing dimensions. Puranik et al. (2008) analyzed a number of microstructural features in the texts written by English-speaking children from 3rd to 6th grade. They found support for an *accuracy* category, which included spelling errors, conventional use of punctuation and capitalization, and a measure of the proportion of grammatical t-units. The accuracy category is then similar to the proposed dimension of transcription, with exception of the role of punctuation, which has been argued, here, to contribute to text organization—rather than constituting a mere convention. A second category that surfaced in Puranik et al.'s (2008) study was productivity, which included number of words, ideas, t-units, and clauses. These variables were found to be very highly correlated across grades, suggesting that a simple measure of number of words is also a reliable estimator of text or content generation. The authors found support for a third category of *complexity*, which included measure of syntactic density and embeddedness; thus, akin to the proposed dimension of syntactic complexity. In a later study of similar characteristics, but with 1st and 4th graders, Wagner et al. (2011) corroborated Puranik et al.'s (2008) findings. In this sense, the present study thus expands on previous research on the dimensionality underlying early writing skills by testing similar categories of microstructural features from a cross-linguistic viewpoint. Moreover, a previously untested dimension of vocabulary is added. Finally, this study differs from Puranik et al.'s (2008) and Wagner et al.'s (2011) in its longitudinal design.

4.1.3 Objectives and Hypotheses

Multiple indicators of each skill and sub-skill were considered, consistent with the multi-dimensional approach. First, the role of orthographic consistency in the early development of written text features is explored, comparing developmental patterns as a function of the orthography. For this purpose, a series of mixed Analyses of Variance (ANOVAs) were performed, with orthography as the between-subjects factor and time as the within-subjects factor. Given the robust evidence suggesting that (1) orthographic consistency affects the development of spelling and that (2) spelling is considered the bottleneck of early writing development, the working hypotheses were (a) that Spanish children would significantly outperform their British peers in all word-level measures, given that these have been shown sensitivity to variations in orthographic consistency (e.g., Seymour et al., 2003). In addition, hypothesis (b) stated that the simpler Spanish orthography would facilitate the emergence and development of text-level features in the Spanish group, again leading to their better performance on all text-level measures relative to the British group. Second, this study aimed to examine the relationships between groups of microstructural features in English and Spanish early writing development. Therefore, Exploratory Factor Analyses (EFA) were carried out for each language group and time point separately, to corroborate whether they validated the hypothesized dimensions.

4.2 Method

4.2.1 Participants

One-hundred eighty-eight English-speaking children from the North of England and 190 Spanish-speaking children from Granada, Spain, were recruited to participate in this study (see recruitment details and other criteria in Section 3.2 above). Children were assessed twice per year, around the middle and end of each school year, on several literacy measures and key literacy-related skills for a total of three years, thus yielding a total of six time points. In this study data are reported of the mid-Year 1 (Time 3), end of Year 1 (Time 4), and mid-Year 2 (Time 5) time points. Only children who were present at all relevant times points (Time 3 through to Time 5) are included in this study. Therefore, the sample used in subsequent analyses consisted of 151 English-speaking children (79 boys, mean age at Time 3: 68.12 months, range: 61-74 months) and 131 Spanish speaking children (73 boys, mean age at Time 3: 75.93 months, range: 70-82 months).

4.2.2 Task and Procedure

Children in both language groups were administered a text writing task. The task required children to report recent past events in writing. The administrator gathered children in groups of 5-6 and asked them about what they had done the day before after they had left school. The aim of this prep-talk was to naturally establish the topic of the task and to activate relevant events in memory, given that the task was time limited. After all the children briefly shared what they had done, they were told they would have to write about those events in a page that had been previously distributed (see Figure 3.1 for a sample). Children had five minutes to write and were encouraged to write as much as they could during the process. However, no advice or more specific prompts were given. They were praised for any efforts and told to write "as they could" if they manifested difficulty or asked for help with writing a certain word.

Once the five-minutes elapsed, children were instructed to draw a picture illustrating their text. The administrator used that time to go around asking each child to read back the text he had produced. The aim of this procedure was to be able to elucidate the intended message the child wanted to convey, especially for children with illegible handwriting or very poor spelling skills.

4.2.3 Scoring of Text-based Features

The texts were transcribed in Excel just as they were written by the children, using the administrator's report obtained from the reading back as a guideline. Most of the measures obtained from the texts resulted from contrasting the intended text against the actual written string. The full list of measures and the tallying procedure is explained in the next sub-section.

4.2.3.1 Word-level measures.

4.2.3.1.1 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.

4.2.3.1.2 Segmentation. The percentage of words that showed conventional boundaries was used as an index of children's word separation skills. When two words were written as one word (e.g., <Iwent>) both tokens were penalized.

4.2.3.1.3 Case. Children's knowledge and use of lower- and upper-case letters was assessed calculating the percentage of words that showed conventional use of case. Words in which the upper-case version of a letter was used in an inappropriate context, e.g., within a word as in <hoMe>, were counted as errors. The appropriate contexts for upper-case letter use were assumed to be proper nouns, names of days of the week and months (English only), pronoun *I*, etc. A capital initial was also expected at the beginning of the text and after any full stop within the text.

4.2.3.2 Text-level measures.

4.2.3.2.1 Number of words. Children's text production skills were assessed by counting the total number of words written in the allotted time. Any word for which a parallel could be established with a word in the rereading was counted. Words were counted automatically in Excel.

4.2.3.2.2 Number of letters. Because average word length is slightly different in each language (Section 2.2.2), children's text production skills were also assessed by counting the number of letters they wrote. Only actual letters and numbers were counted. 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.

4.2.3.2.3. Connectivity. Children's use of connectors (discourse markers and conjunctions) was assessed by identifying all inter-clausal connectors in the texts and then calculating the percentage these represented out of the total number of words in the text. Connectors were identified manually and counted automatically in Excel.

4.2.3.2.4 Punctuation. All types of punctuation marks—e.g., full stops, exclamation points, commas, question marks, etc.—were counted manually directly from the original texts. In order to control for the effect of overall text length, whereby longer text may contain more punctuation marks, the total number of punctuation marks was divided by the number of words in the text and multiplied by 100.

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

4.2.3.2.6 Subordination index. This measure consisted of calculating the percentage of subordinated clauses out of the total number of clauses, thus constituting an index of the amount of embeddedness in children's texts.

4.2.3.2.7 Adjectives and adverbs. Children's use of optional word classes to provide detail and enrich the report of events was estimated by manually identifying all adjectives and adverbs in the texts, which were then counted automatically in Excel. Adverbs of obligatory expression, such as those in phrasal verbs, e.g., take off, were not counted. The final measure consists in calculating the percentage of adjectives and adverbs out of all words in the text.

4.2.3.2.8 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. All words in the texts were labelled for grammatical class. Then, all content words were counted and the resulting number was divided over the total amount of words in the texts and multiplied by 100.

4.2.3.2.9 Lexical diversity. The type-token ratio was calculated for each text as an indicator of lexical diversity in written composition. Any modification to a word base, such as adding or subtracting inflectional endings or derivational morphemes, was interpreted as a different type. This variable presented a bias in favour of Spanish, probably due to its pro-drop nature. As a result, it will only be used in the within-language analyses in this chapter—i.e., correlations and EFA. See section 3.8.12 for more details.

4.2.3.2.10 Average content-word length. The average number of letters of all content words in children's texts was calculated as an estimator of vocabulary frequency and sophistication. This variable presented a bias in favour of Spanish, probably due to its higher average word length. As a result, it will only be used in the within-language analyses in this chapter—i.e., correlations and EFA. See section 3.8.15 for more details.

4.3 Results

4.3.1 Mixed ANOVAs

Two-way repeated-measures analyses of variance (ANOVAs) were conducted for each measure to determine the effects of time (within-subjects factors) and language/orthographic consistency (between-subjects factor) in the expression of word- and text-level writing features[9]. Mauchly's tests were usually significant, meaning that the sphericity assumption was not met; therefore, corrected Greenhouse-Geisser F values and degrees of freedom, with their associated p values, are reported. Because improvements in scores were expected over time, planned contrasts were conducted by choosing the Repeated option for the within-subjects factor, time, to compare the first level to the second—mid-Year 1 (Time 3) vs. end-Year 1 (Time 4)—, and the second level to the third—end-Year 1 (Time 4) vs. mid-Year 2 (Time 5). A significant result for the effect of time should always be interpreted as an increase in the use of a given measure (e.g., higher percentages of conventionally written words, rate of punctuation mark use, adjectives and adverbs, subordinated clauses, etc.). The direction of the language effect is explicitly reported in each relevant case. Any significant interaction between time and language was followed up by independent samples t-tests, to check for the simple effect of language at each level of the within-subjects factor (Time 3, Time 4, and Time 5).

[9] Although a significant difference was found for age, which favoured Spanish participants, and parents' education, which favoured English participants (Section 3.2), these variables showed no significant associations with the dependent variables included in this study, after controlling for the main factor: orthographic consistency. Partial correlations are reported in Appendix 4. The full set of parallel analyses—that is, controlling for age and for age and parents' education—is nevertheless reported in Appendix 4. For the most part, the pattern of results was unaltered. Discrepancies with the results reported in this study are also examined in Appendix 4.

Table 4.1

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

	English			Spanish		
	Time 3	*Time 4*	*Time 5*	*Time 3*	*Time 4*	*Time 5*
	M (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*M* (SD)
TSPELL	47.34 (23.77)	64.08 (21.15)	69.48 (19.37)	55.30 (26.66)	81.16 (12.28)	85.25 (9.58)
SEG	48.41 (36.58)	68.58 (33.37)	79.02 (21.17)	54.88 (34.22)	87.79 (13.79)	92.13 (10.19)
CASE	65.43 (26.31)	75.28 (21.57)	77.19 (18.15)	64.71 (32.65)	91.70 (8.36)	94.10 (5.53)
WORD	15.36 (8.52)	25.08 (12.57)	31.90 (15.89)	14.36 (8.41)	24.04 (8.55)	34.47 (13.29)
LETT	56.60 (31.21)	92.15 (46.15)	117.79 (57.82)	53.64 (29.83)	88.79 (30.35)	129.07 (50.46)
PUNCT	5.41 (8.68)	6.03 (6.53)	6.53 (6.28)	1.29 (4.49)	4.07 (5.43)	5.45 (4.99)
CONN	4.07 (6.27)	5.73 (5.98)	8.13 (5.72)	6.42 (7.32)	9.34 (6.63)	10.55 (5.89)
WCL	5.08 (1.30)	5.83 (1.90)	5.54 (1.20)	4.97 (2.52)	5.79 (2.03)	5.66 (1.57)
SUBORD	5.77 (11.42)	5.78 (10.70)	12.43 (14.97)	6.59 (14.05)	9.04 (14.31)	12.58 (14.78)
ADJADV	2.26 (4.15)	4.73 (5.71)	6.18 (5.26)	1.94 (3.58)	3.69 (4.84)	4.83 (4.34)
LEXDEN	46.48 (16.78)	43.96 (11.05)	46.02 (7.91)	48.78 (17.29)	44.17 (8.15)	46.02 (7.91)

Note. TSPELL: spelling accuracy (percentage correct); SEG: word segmentation (percentage correct); CASE: use of case (percentage correct); WORD: number of words; LETT: number of letters; PUNCT: number of punctuation marks divided by no. of words and multiplied by 100; CONN: percentage of connectors out of all words; WCL: words per clause; SUBORD: percentage of subordinated clauses; ADJADV: percentage of adjectives and adverbs; LEXDEN: lexical density (percentage of content words).

Similarly, one-way repeated measures ANOVAs tested the simple effect of time in English and Spanish, separately. The Bonferroni correction was selected in all cases where multiple comparisons were conducted, including both planned contrasts as well as simple effects tests. Effect sizes are reported for the main effect of language, for planned contrasts, and for simple effects. In all cases *r* was calculated, using Formula 4.1 with the within-subject planned contrasts and the main effect of language. Formula 4.2 was used with the independent-samples t-tests.

Formula 4.1 $$r = \sqrt{\frac{F(1,\mathrm{df_R})}{F(1,\mathrm{df_R}+\mathrm{df_R})}}$$

Formula 4.2 $$r = \sqrt{\frac{t^2}{t^2+\mathrm{df}}}$$

Table 4.1 displays the descriptive statistics for all dependent variables. Table 4.2 shows the exact *F, p,* and *r* values of the planned contrasts for the effect of time aggregating across groups. Table 4.3 shows the results of the *t*-tests to follow up the simple effect of language. Finally, the results of the one-way repeated-measures ANOVAs to test the simple effect of time are shown in Table 4.4, for English, and 4.5, for Spanish. Unless otherwise stated, nonsignificant results yielded effect sizes $r \leq .10$.

4.3.1.1 Transcription. This dimension included the assessment of spelling, word segmentation, and use of case. For spelling accuracy, there was a significant main effect of time, $F(1.53, 427.43) = 278.64$, $p < .001$, with planned contrasts revealing that the effect size was strongest in the Time 3-Time 4 comparison (Table 4.2). There was also a significant main effect of language, $F(1, 280) = 48.80$, $p < .001$, $r = .39$, indicating that Spanish children outperformed English children. There was a significant Time x Language interaction of small size, $F(1.53, 427.43) = 8.80$, $p = .001$, $r = .03$ (Figure 4.1). Simple effects *t*-tests showed that the language effect became stronger with time (Table 4.3). On the other hand, repeated measures ANOVAs revealed that English, $F(1.71, 256.65) = 143.07$, $p < .001$, and Spanish children, $F(1.4, 182.06) = 135.08$, $p < .001$, improved significantly with time, and both did so more strongly during the Time

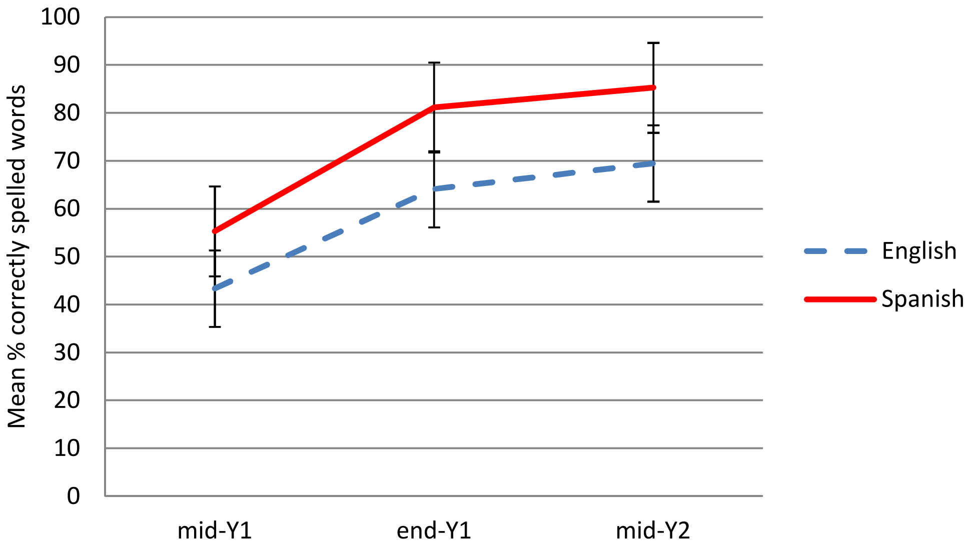


Figure 4.1. Graph of the significant Language x Time interaction for the measure of text spelling accuracy (percentage of words spelled conventionally).

4-Time 5 interval (Table 4.4, for English; Table 4.5, for Spanish). These results indicate that children's spelling skills improved to a greater extent during Year 1 across languages than in the transition from the end of Year 1 to mid-Year 2. Moreover, the role of the orthography became more important with time, with the Spanish children showing a larger advantage over their English peers in the later time points.

Children's ability to segment words in writing conventionally improved with time, $F(1.63, 455.61) = 219.99$, $p < .001$. Planned contrasts revealed that the improvement was larger from Time 3 to Time 4 than from Time 4 to Time 5 (Table 4.2). A significant main effect of language (Spanish > English) was found, $F(1, 280) = 22.24$, $p < .001$, $r = .27$. Main effects were however modified by a significant Time x Language interaction, $F(1.63, 455.61) = 7.02$, $p = .002$, $r = .03$, of very small size (Figure 4.2). Simple effects *t*-tests indicated that the language effect was nonsignificant at Time 3 (Table 4.3), but a significant for both Time 4 and Time 5. Effect sizes indicated that differences as a function of language became more important with time. One-way repeated measures ANOVAs were significant for English, (Table 4.4), and Spanish, (Table 4.5). Similarly to

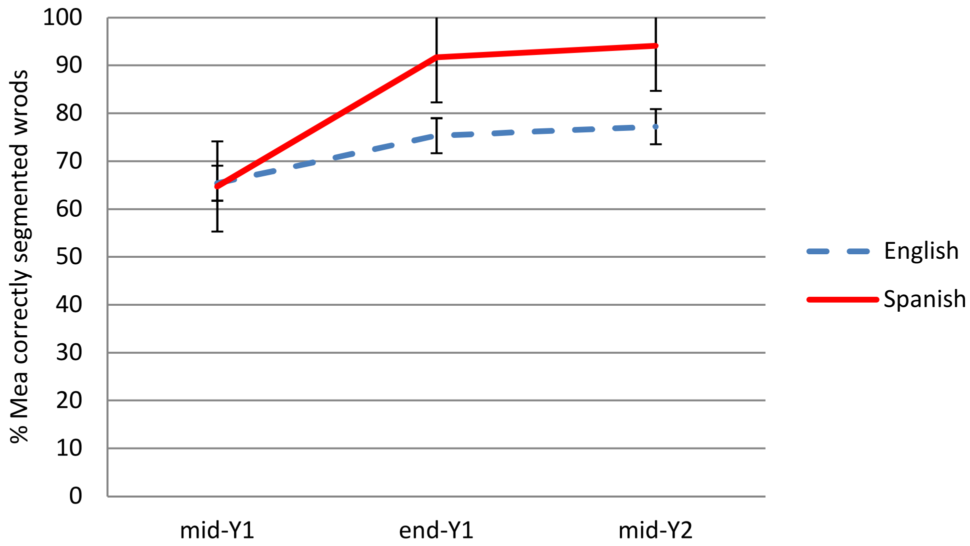


Figure 4.2. Graph of the significant Language x Time interaction for the measure of word segmentation accuracy (percentage of words segmented conventionally).

the case of spelling, the largest improvements in word segmentation occurred during the first interval, mid- to end-Year 1.

The analyses of children's use of case during text composition showed a significant main effect of time, $F(1.35, 376.69) = 119.27$, $p < .001$, which was strong in the Time 3-Time 4 contrast and small in the Time 4-Time 5 interval (Table 4.2). There was also a medium-size, significant main effect of language, $F(1, 280) = 33.57$, $p < .001$, $r = .33$, by which Spanish children outperformed their English peers; however, this was modified by a significant Time x Language interaction, ($F(1.35, 376.59) = 23.52$, $p < .001$ (Figure 4.3). Simple main effects *t*-tests confirmed that the language effect was significant at Time 4 and Time 5, but nonsignificant at Time 3 (Table 4.3), again suggesting that the advantage of Spanish children over their English peers becomes more important with time. One-way repeated measures ANOVAs were significant for English (Table 4.4) and Spanish, (Table 4.5), and follow up contrasts showed that the stronger effect of time during Year 1 held in both language groups. Just like conventional spelling and word segmentation, children's appropriate use of case seems to be a feature that develops considerably during Year 1, while improvement continues, although at a slower pace, during Year 2.

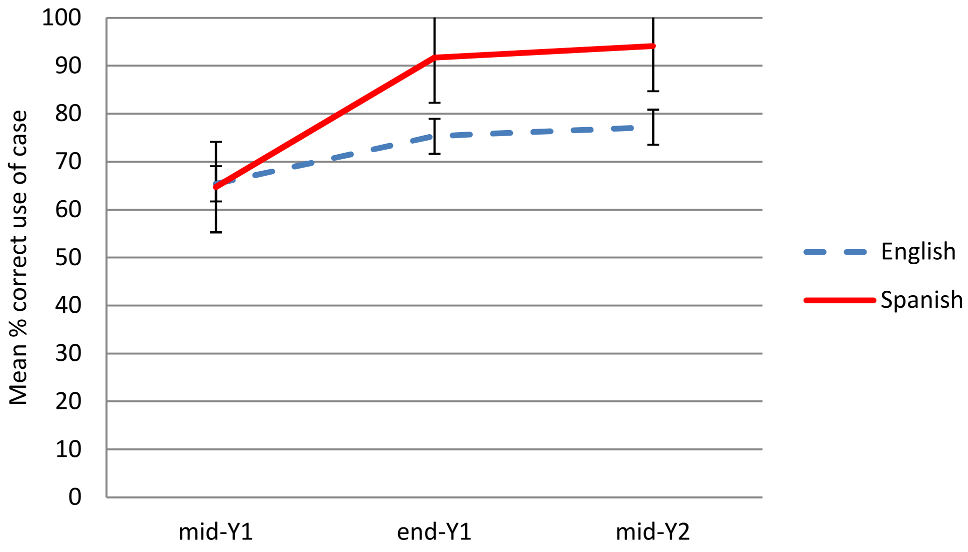


Figure 4.3. Graph of the significant Language x Time interaction for the measure of accuracy in the use of case accuracy (percentage of words conventional use of lower- and upper-case letters).

Table 4.2

Planned contrasts for the main effect of Time

	Time 3-Time 4			Time 4-Time 5		
Measures	*F*[1]	*p*[2]	*r*	*F*[1]	*p*[2]	*r*
TSPELL	266.23	< .001	.698	36.60	< .001	.341
SEG	214.44	< .001	.660	34.98	< .001	.333
CASE	120.23	< .001	.549	7.08	.008	.158
WORD	222.21	< .001	.665	134.18	< .001	.569
LETT	239.82	< .001	.679	143.03	< .001	.581
PUNCT	9.60	.002	.182	5.24	.023	.134
CONN	20.82	< .001	.263	16.00	< .001	.232
WCL	23.67	< .001	.279	2.17	.142	.089
SUBORD	1.47	.226	.071	20.41	< .001	.261
ADJADV	32.21	< .001	.321	10.73	.001	.192
LEXDEN	11.27	.001	.197	2.96	.086	.100

Note. TSPELL: spelling accuracy; SEG: word segmentation; CASE: use of case; WORD: number of words; LETT: number of letters; PUNCT: punctuation; CONN: connectors; WCL: words per clause; SUBORD: proportion of subordinated clauses; ADJADV: proportion of adjectives and adverbs; LEXDEN: lexical density.

[1]Df = 1; df (error) = 280.

[2] Significance level (corrected for multiple comparisons) < .025

4.3.1.2 Productivity. A significant main effect of time was found on the length of the texts children wrote, whether it was measured in words, $F(1.91, 533.94) = 285.85$, $p < .001$, or letters, $F(1.87, 523.52) = 332.28$, $p < .001$. Contrasts showed that the effect was strong at both the Time 3-Time 4 and Time 4-Time 5 intervals (Table 4.2). In contrast to the transcription measures, the effect of language was nonsignificant for words $F(1, 280) = 0.01$, $p = .941$, $r < .01$, or for letters $F(1, 280) = 0.16$, $p = .686$, $r = .03$. There was, however, a significant Time x Language interaction for number of words, $F(1.89, 529.09) = 4.36$, $p = .015$; and letters, $F(1.87, 523.52) = 4.95$, $p < .001$(see Figures 4.4 for a graphic representation of this interaction for the number of words measure). One-way repeated measures ANOVAs indicating skill growth over time were significant for number of words and letters in English (Table 4.4) and Spanish (Table 4.5). Planned contrasts showed that the effect of time was consistently of large size for Spanish, ranging from $r = .66$ to .72, across number of words and letters (Table 4.5). In the case of the English group, the effect size of time in the end-Year 1 to mid-Year 2 (i.e., Time 4-Time 5) contrast is somewhat smaller than in the Year 1 interval (i.e., Time 3-Time 4), ranging from .47 to .65 (Table 4.4).

In sum, although the orthography did not have a direct impact on the average text length children produced, the significant Time x Language interaction revealed that the Spanish group experienced consistent improvement in text productivity across time points, while English children experienced a subtle deceleration in the end- Year 1 to mid-Year 2 interval.

Table 4.3

Results for tests of simple effects of language *(independent-samples* t*-tests)at each time point*

	Time 3			Time 4			Time 5		
	t(df)	p^1	r	t(df)	p^1	r	t(df)	p^1	r
TSPELL	-2.65 (280)	.008	.156	-8.42 (246.35)	< .001	.473	-8.47 (225.74)	< .001	.491
SEG	-1.57 (280)	.117	.093	-6.42 (205.84)	< .001	.408	-5.49 (196.90)	< .001	.364
CASE	0.17 (248.97)	.866	.011	-8.60 (199.71)	< .001	.520	-10.88 (181.68)	< .001	.628
WORD	1.26 (280)	.210	.075	0.83 (265.52)	.408	.051	-1.48 (279.64)	.140	.088
LETT	0.81 (280)	.419	.048	0.73 (261.77)	.465	.045	-1.75 (279.99)	.081	.104
PUNCT	5.09 (231.48)	< .001	.317	2.76 (279.52)	.006	.163	1.61 (277.90)	.109	.096

Note. TSPELL: spelling accuracy; SEG: word segmentation; CASE: use of case; WORD: number of words; LETT: number of letters; PUNCT: punctuation. A negative t value indicates that the English group obtained lower mean scores.

[1] Significance level (corrected for multiple comparisons) < .017

4.3.1.3 Text structure. This dimension evaluated children's use of punctuation marks, controlling for text length, and the use of connectors, calculated as the percentage of inter-clausal connectors over the total number of words. In the case of punctuation, there was a significant main effect of time, $F(1.81, 505.45) = 14.50$, $p < .001$, which focused contrasts revealed to be small for both the Time 3-Time 4 and Time 4-Time 5 intervals (Table 4.2). There was also a significant, small main effect of language, $F(1, 280) = 24.07$, $p < .001$, $r = .28$, by which the English children included more punctuation marks on average. These effects were modified by a significant Time x Language interaction, $F(1.81, 505.45) = 4.92$, $p = .010$. Simple main effects t-tests showed that the English children's advantage at Times 1 and 2 was nonsignificant by Time 5 (Table 4.3).

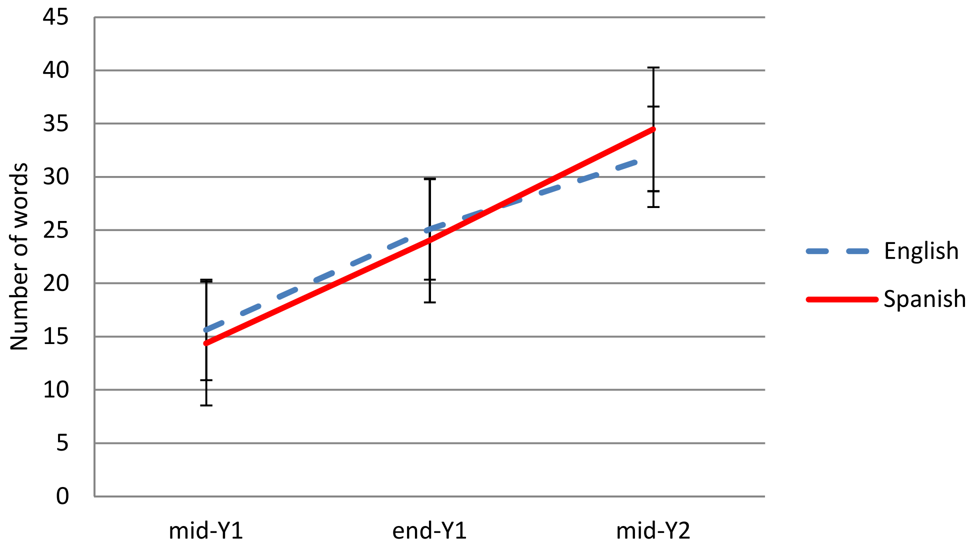


Figure 4.4. Graph of the significant Language x Time interaction for the measure of number of words.

Table 4.4

Results for the simple effects tests of Time (one-way repeated measures ANOVA) of the English group

	Main effect of Time		Planned contrasts: Time 3 vs Time 4			Planned contrasts: Time 4 vs Time 5		
	F(df, dfe)	*p*[1]	*F*(1, 150)	*p*[1]	*r*	*F*(1, 150)	*p*[1]	*r*
TSPELL	143.07 (1.71, 256.65)	< .001	136.52	< .001	.690	26.06	< .001	.385
SEG	87.02 (1.80, 270.49)	< .001	77.45	< .001	.584	25.57	< .001	.382
CASE	27.02 (1.70, 254.58)	< .001	29.77	< .001	.407	1.92	.168	.114
WORD	125.53 (1.91, 258.68)	< .001	104.95	< .001	.642	42.46	< .001	.470
LETT	134.69 (2, 300)	< .001	110.26	< .001	.651	45.71	< .001	.484
PUNCT	31.44 (2, 260)	< .001	0.50	.481	.055	0.62	.433	.063

Note. TSPELL: spelling accuracy; SEG: word segmentation; CASE: use of case; WORD: number of words; LETT: number of letters; PUNCT: punctuation.

[1] Significance level (corrected for multiple comparisons) < .017

One-way repeated measures ANOVAs to test the simple effect of time were significant for Spanish (Table 4.5), but not for English (Table 4.4). In Spanish, focused contrasts were significant for the Time 3-Time 4 and the Time 4-Time 5 comparisons, but with very small effect sizes in both cases. (Table 4.5) In short, the earlier start by the English children in the use of punctuation marks resulted in their not showing any significant improvement in the use of this feature of writing. Spanish children, who lagged significantly behind their English peers in the use of punctuation marks by mid-Year 1, caught up during this period of time, showing no difference with the English group by the middle of Year 2.

Table 4.5

Results for the simple effects tests of Time (one-way repeated measures ANOVA) of the Spanish group

	Main effect of Time		Planned contrasts: Time 3 vs Time 4			Planned contrasts: Time 4 vs Time 5		
	F(df, dfe)	p^1	$F(1, 130)$	p^1	r	$F(1, 130)$	p^1	r
TSPELL	135.08 (1.40, 182.06)	< .001	130.32	< .001	.473	12.39	< .001	.295
SEG	142.82 (1.31, 170.66)	< .001	135.60	< .001	.408	11.70	< .001	.288
CASE	91.74 (1.11, 144.83)	< .001	86.05	< .001	.520	9.19	.003	.257
WORD	194.18 (1.86, 241.40)	< .001	122.11	< .001	.051	98.76	< .001	.657
LETT	207.13 (1.77, 229.87)	< .001	140.55	< .001	.045	101.95	< .001	.633
PUNCT	31.44 (2, 260)	< .001	23.02	< .001	.163	7.69	.006	.237

Note. TSPELL: spelling accuracy; SEG: word segmentation; CASE: use of case; WORD: number of words; LETT: number of letters; PUNCT: punctuation.

[1] Significance level (corrected for multiple comparisons) < .017

The average percentage of connectors used was affected significantly by time, $F(2, 560) = 35.47$, $p < .001$, with small effects in both contrasts (Table 4.2). There was also a significant main effect of language of medium size, $F(1, 280) = 31.16$, $p < .001$, $r = .32$, by which the Spanish children included more connectors, on average, than their English peers. The Time x Language interaction was nonsignificant, $F(2, 560) = 1.06$, $p = .348$. The use of connectivity devices developed at a slow but steady rate: children across languages showed an increasing awareness of the need of connectors to link the different parts of the short pieces of writing they were producing. Contrarily to the use of punctuation marks, where the English group showed an advantage at least in the first two time points, Spanish children made a larger use of connectors when composing written texts.

4.3.1.4 Syntactic complexity. Children's syntactic complexity in written text production was assessed, using the indicators: number of words per clause and the percentage of subordinated clauses. For the number of words per clause measure, a significant main effect of time was found, $F(1.94, 544.29) = 14.85$, $p < .001$. However, focused contrasts showed that the effect to be small and significant only for the Time 3-Time 4 contrast (Table 4.2). There was no significant main effect of language, $F(1, 280) = 0.01$, $p = .931$, and the Time x Language interaction was also nonsignificant, $F(1.94, 544.29) = 0.29$, $p = .745$. In short, children only slightly increased the average number of words they included in a clause over the Year 1 time points, but the ratio words/clause was otherwise quite stable in the time frame under consideration. Furthermore, orthographic consistency did not impact the development of this feature of text composing.

The percentage of subordinated clauses that children included in their texts also increased significantly as a function of time, $F(1.95, 546.48) = 18.42$, $p < .001$. Focused contrasts revealed that the effect was only significant, and small, in the Time 4-Time 5 contrast (Table 4.2). Similarly to the words-per-clause ratio, language was not a significant factor, $F(1, 280) = 2.13$, $p = .146$, $r = .09$, and the interaction was not significant, $F(1.94, 544.29) = 1.10$, $p = .334$. In sum, while the words-per-clause aspect of syntactic complexity showed a slight but significant increase during the Year 1 time points, the percentage of subordination increased especially during Year 2. On the other hand, neither syntactic measure seemed to be affected by orthographic consistency.

4.3.1.5 Vocabulary. Children's vocabulary choices during text composition were evaluated by two measures: the percentage of adjectives and adverbs, that is, of optional lexical elements; and lexical density, which was calculated as the percentage of content words (nouns, main verbs, adjectives, and adverbs) of all words in the texts. The average percentage of adjectives and adverbs children included in their texts increased significantly with time, $F(2, 560) = 41.68$, $p < .001$. Focused contrasts indicated that the effect was of medium-small size in the Year 1 interval and smaller in the Year 1 to mid-Year 2 contrast (Table 4.2). The analyses also revealed a significant, small, main effect of language, $F(1, 280) = 6.33$, $p = .012$, $r = .15$, with English children including more adjectives and adverbs than the Spanish group. The interaction between language and time was nonsignificant, $F(2, 560) = 0.96$, $p = .382$. In short, children slowly incremented their use of optional lexical elements, even in the very short period under consideration. Finally, the small effect of orthographic consistency went in the opposite direction of the predictions, with the English children showing some advancement over their Spanish counterparts.

Lexical density showed a significant effect of time, $F(1.51, 423.47) = 7.81$, $p = .002$, with focused contrasts revealing that the effect was again only significant, and small, in the mid-Year 1 to end-Year 1 (Time 3-Time 4) contrast, but not in the end-Year 1 to mid-Year 2 contrast, when it was nonsignificant (Table 4.2). These results indicate that children very slightly increased the proportion of content words in their texts, thus making them more informative. No effect of orthographic consistency was attested, $F(1, 280) = 0.55$, $p = .814$, $r < .01$, and the Time x Language interaction was also nonsignificant, $F(2, 560) = 2.45$, $p = .103$. In short, the results for the vocabulary dimension point to a very slow rate of growth in the proportion of open word classes, in general, and more specifically, in the proportion of optional lexical elements.

4.3.2 Concurrent Correlations

In order to examine the pattern of relationships between the multiple indicators, Pearson product-moment correlations were conducted for each language and time point. Tables 4.6, 4.7, and 4.8 show the concurrent correlations among all text-based measures at Times 3, 4, and 5, respectively. Transcription or word-level measures (text spelling, segmentation, and case) 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 medium to large correlations, with the exception of the correlation between case and segmentation at Time 3 (Table 4.6). Therefore, although correlations differed in slightly in size, both language groups showed these three word-level indicators to be tightly related at the early stages of learning to write.

Productivity measures, number of words and number of letters, showed very high correlation values throughout the study ($r > .90$). The close relationship between different measures of productivity has been found before (e.g., Berninger et al., 1992; Puranik et al., 2008; Wagner et al., 2011). In this sense, it has been suggested that number of words constitutes the most representative of these measures and, thus, the most widely used in studies measuring productivity (Puranik et al., 2008). It is also evident from Tables 4.6, 4.7, and 4.8, that number of words and number of letters were not only extremely highly correlated, but also the pattern of correlations with other measures was virtually identical.

The rest of the text-level dimensions, text structure, syntactic complexity, and vocabulary, did not show such a clear pattern of relationships. With a few exceptions, correlations between the two indicators of each dimension were in the low range and were usually nonsignificant. In the case of the text structure dimension, correlations between use of punctuation and percentage of connectives were always nonsignificant, except in Spanish at Time 5 (Table 4.8). Interestingly, both languages showed a tendency towards a small, negative correlation, suggesting that children in both languages groups used fewer punctuation marks if they used more connectives.

Table 4.6

Correlations between text-based measures at Time 3

	1	2	3	4	5	6	7	8	9	10	11	12	13
1.TSPELL	--	.70*	.77*	.60*	.53*	.16	.35*	-.05	.12	.13	.25*	-.14	-.27*
2. SEG	.43*	--	.59*	.46*	.42*	.30*	.37*	.01	.05	.16	.13	-.09	-.14
3. CASE	.32*	.21	--	.54*	.51*	.14	.29*	-.03	.06	.20	.20	-.07	-.16
4. WORD	.57*	.29*	.33*	--	.96*	.09	.29*	.06	.22*	.12	.34*	-.53*	-.19
5. LETT	.54*	.33*	.34*	.96*	--	.09	.32*	.08	.19	.13	.31*	-.49*	-.02
6. PUNCT	.08	.18	-.02	.13	.15	--	.08	.00	-.01	.12	.07	.06	.06
7. CONN	.35*	.27*	.22	.39*	.40*	.11	--	-.15	.14	.03	-.01	-.11	-.04
8. WCL	.23*	-.01	.08	.39*	.35*	.01	.07	--	-.10	-.10	.12	.05	.01
9.SUBORD	-.05	-.05	-.02	.02	-.01	-.06	-.04	.07	--	.02	.31*	-.02	-.22*
10.LEXDEN	-.03	.28*	.14	-.20	-.10	-.02	-.15	-.34*	-.24*	--	.21*	.16	-.06
11.ADJADV	.21	.25*	.11	.23*	.30*	.07	.21	.24*	.01	.05	--	-.01	-.32*
12. LEXDIV	-.31*	-.09	-.10	-.70*	-.61*	-.07	-.23*	-.33*	-.01	.33*	-.01	--	.03
13. WLEN	-.13	.04	.06	.03	.21	-.01	.04	.12	-.11	.04	.06	.09	--

Note. Correlations for the English and Spanish samples appear on top and below the diagonal, respectively. TSPELL = text spelling; SEG = segmentation; CASE = use of case; WORD = text length in words; LETT = text length in letters; PUNCT = punctuation; CONN = connectors; WCL = words per clause; SUBORD = percentage of subordination; LEXDEN = lexical density; ADJADV = percentage of adjectives and adverbs; LEXDIV = lexical diversity; WLEN = average length of content words.

*$p < .01$.

Table 4.7

Correlations between text-based measures at Time 4

	1	2	3	4	5	6	7	8	9	10	11	12	13
1.TSPELL	--	.68*	.71*	.54*	.52*	.25*	.27*	-.09	.22	.18	.20	-.24*	-.15
2. SEG	.43*	--	.63*	.41*	.40*	.23*	.19	-.02	.20	.24*	.25*	-.08	-.06
3. CASE	.35*	.32*	--	.53*	.52*	.21	.27*	-.06	.19	.28*	.37*	-.25*	-.12
4. WORD	.29*	.13	.39*	--	.98*	.06	.29*	-.01	.27*	.35*	.33*	-.53*	-.08
5. LETT	.28*	.18	.40*	.94*	--	.07	.29*	.01	.27*	.36	.35*	-.48*	.05
6. PUNCT	.09	.14	-.09	-.12	-.03	--	-.16	-.07	.01	.22	.29	.05	-.09
7. CONN	.01	-.01	-.02	.14	.08	-.18	--	.03	.07	-.01	.01	-.21	.13
8. WCL	.06	-.11	-.19	-.05	-.11	-.03	-.32	--	-.01	-.09	-.13	.08	.11
9.SUBORD	.10	-.01	-.18	.30*	.26*	-.18	-.05	-.21	--	.03	.07	.05	-.01
10.LEXDEN	-.22*	-.08	.12	.15	.15	-.11	-.05	-.15	.18	--	.47*	-.03	-.02
11.ADJADV	-.03	.09	.29*	-.06	.03	-.02	-.14	.20	.09	.47*	--	.08	-.21
12. LEXDIV	-.18	-.07	-.19	-.60*	-.54*	.16	-.15	.01	-.03	.21	-.03	--	.10
13. WLEN	-.18	.13	-.01	-.17	.03	.16	-.13	.10	-.19	-.08	-.17	.11	--

Note. Correlations for the English and Spanish samples appear on top and below the diagonal, respectively. TSPELL = text spelling; SEG = segmentation; CASE = use of case; WORD = text length in words; LETT = text length in letters; PUNCT = punctuation; CONN = connectors; WCL = words per clause; SUBORD = percentage of subordination; LEXDEN = lexical density; ADJADV = percentage of adjectives and adverbs; LEXDIV = lexical diversity; WLEN = average length of content words.

**p* < .01.

Table 4.8

Correlations between measures at Time 5

	1	2	3	4	5	6	7	8	9	10	11	12	13
1.TSPELL	--	.65*	.65*	.42*	.41*	.15	.25*	.11	.09	.29	.34*	-.06	-.18
2. SEG	.61*	--	.56*	.36*	.36*	.18	.21	.11	.18	.21	.32*	-.01	-.05
3. CASE	.53*	.43*	--	.38*	.37*	.11	.21	.07	.15	.35*	.44*	-.03	-.23*
4. WORD	.03	.09	.13	--	.98*	-.26*	.29*	.25	.31*	.47*	.21	-.56*	-.23*
5. LETT	.05	.11	.13	.97*	--	.24*	.28*	.25*	.30*	.46*	.22*	-.54*	-.10
6. PUNCT	.04	-.02	-.08	-.21	-.18	--	-.20	-.06	-.03	-.06	.21	.29*	.24*
7. CONN	.09	.10	.12	.10	.12	-.28*	--	-.01	.00	.06	.06	-.23*	-.04
8. WCL	-.09	-.21	.07	.04	.01	-.19	-.30*	--	-.03	.23*	-.17	.00	.02
9.SUBORD	.08	.03	.22	.22	.21	-.11	-.04	-.24*	--	.22*	.06	-.08	-.13
10.LEXDEN	-.06	-.06	-.25*	.01	.04	-.04	-.10	-.10	.01	--	.44*	.12	-.11
11.ADJADV	.28*	.35*	.10	-.10	-.03	.09	.01	-.42*	.10	.46*	--	-.07	-.24*
12. LEXDIV	.14	-.03	-.04	-.63*	-.56*	.11	-.09	-.10	-.02	.34*	.22	--	.09
13. WLEN	-.05	-.05	.17	-.10	-.00	-.07	-.10	.19	-.09	-.14	-.19	.10	--

Note. Correlations for the English and Spanish samples appear on top and below the diagonal, respectively. TSPELL = text spelling; SEG = segmentation; CASE = use of case; WORD = text length in words; LETT = text length in letters; PUNCT = punctuation; CONN = connectors; WCL = words per clause; SUBORD = percentage of subordination; LEXDEN = lexical density; ADJADV = percentage of adjectives and adverbs; LEXDIV = lexical diversity; WLEN = average length of content words.

*$p < .01$.

Syntactic measures were, as a rule, not significantly correlated with each other throughout the study and across language groups. Finally, individual measures of the vocabulary dimension were mostly unrelated at Time 3 (Table 4.6), with the exception of content-word length, which was negatively associated with the proportion of adjectives and adverbs in English; and, in Spanish, lexical diversity was positively associated with lexical density. This pattern was also observed at Time 5, but not at Time 4. At Times 4 and 5, however, both language groups showed moderate correlations between the proportion of adjectives and adverbs and lexical density (Table 4.7 and 4.8). 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 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 into aspects of written vocabulary (Malvern, et al., 2004).

4.3.3 Exploratory Factor Analyses

The hypothesized dimensions of word-level/transcription, productivity, text structure, syntax, and vocabulary 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 has been used in previous studies investigating the dimensionality of developmental writing (Puranik et al., 2008). 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, cited in Kim & Mueller, 1978), as well as those with communalities < .50 (Field, 2009), and the analysis was re-run with the remaining variables. In all cases, factors extracted were those with eigenvalues > 1.

4.3.3.1 Dimensionality of writing in English. Table 4.9 shows the standardized regression coefficient factor loadings for English at Times 3, 4, and 5. At Time 3 a four-factor solution was found, which explained 76% of the variance. At Times 4 and 5, in contrast, a three-factor solution was found, which explained 80.08%

and 79.27% of the variance, respectively. At all time points, spelling accuracy, segmentation, and case loaded on a common factor. Also across time points, both measures of text length, number of words and letters, loaded on a common factor, together with lexical diversity (type-token ratio). It should be noted that, although initially thought of as an estimate of vocabulary-related aspects, lexical diversity was always strongly—and negatively—associated with productivity measures, meaning that longer texts were characterized by less diverse vocabulary (Section 3.9). At Times 4 and 5 two measures of vocabulary, percentage of adjectives and adverbs and lexical density, loaded on a common factor. As predicted, punctuation and connectivity loaded on a common factor, though only at Time 3; in addition, and as noted in the description of the correlation patterns, their loadings were of opposite signs. Percentage of subordinated clauses and percentage of adjectives and adverbs loaded on a common factor at Time 3 only. Words-per-clause and average content-word length were always dropped from the analyses, given that they contributed poorly to the models and/or their communalities did not equal or exceed .50.

In short, two factors were consistently obtained across time points in English: a word-level/transcription factor, composed by spelling accuracy, segmentation, and case; and a productivity factor, composed by number of letters, number of words, and lexical diversity (or type-token ratio). A vocabulary factor emerged at Time 4, composed of lexical density and percentage of adjectives and adverbs. Other factors were a one-off occurrence. The word-level and productivity factors showed correlations of -.286, -.353, and -.210 at Times 3, 4, and 5, respectively. The word-level factor showed a correlation of .372 with the vocabulary factor at Times 4 and 5, while the correlation between the productivity factor and the vocabulary factor was -.154 at Time 4 and -.137 at Time 5.

4.3.3.2 Dimensionality of writing in Spanish. Table 4.10 shows the standardized regression coefficient factor loadings for Spanish at Times 3, 4, and 5. At Times 3 and 4, a two-factor solution was found, which explained 81.74% and 76.69% of the variance, respectively. Spelling accuracy and segmentation, but not case, loaded heavily on the first factor. Only at Time 5 did case load on the word-level factor. Similarly to English, number of words, letters, and the lexical diversity loaded on another factor across time points. At Time 5 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. In comparison to English, more variables did not contribute to explaining variance in writing development as components of any of the dimensions identified. These were the syntactic complexity measures, percentage of subordination and words per clause; the text structure measures, connectivity and punctuation; and the average length of content words. Similarly to English, lexical diversity consistently loaded on the productivity factor, instead of being associated to a vocabulary factor, as hypothesized.

The word-level and the productivity factor showed correlations in Spanish of .297, .220, and .043 at Times 3, 4, and 5. The vocabulary factor showed a correlation of -.108 with the word-level factor, and of .052 with the productivity factor.

In sum, word-level or transcription measures, as well as the productivity dimension, were stable constructs in both English and Spanish early text writing development. A vocabulary factor was also identified in both language groups, composed of adjectives and adverbs and lexical density. The examination of the correlation pattern (Section 4.3.2) revealed some tendencies for other text-level dimensions, but these were characterized by the unreliability of their patterns. This suggests that some aspects of text-level composition may not be completely mature or reliable, an issue that should be taken into consideration in subsequent studies.

Table 4.9

Standardized regression coefficient factor loadings for English at Times 3, 4, and 5[†]

	Time 3				Time 4			Time 5		
	word level	*productivity*	*text structure*	*SUBORD + ADJADV*	*word level*	*productivity*	*vocabulary*	*word level*	*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%
TSPELL	.821	--	--	--	.897	--	--	.873	--	--
SEG	.853	--	--	--	.932	--	--	.913	--	--
CASE	.755	--	--	--	.792	--	--	.727	--	--
WORD	--	-.731	--	--	--	-.763	--	--	-.843	--
LETT	--	-.722	--	--	--	-.786	--	--	-.852	--
PUNCT	--	--	-.800	--	--	--	--	--	--	--
CONN	.665	--	.569	--	--	--	--	--	--	--
SUBORD	--	--	--	.830	--	--	--	--	--	--
LEXDEN	---	--	--	--	--	--	.767	--	--	.790
ADJADV	--	--	--	.771	--	--	.908	--	--	.875
LEXDIV	--	.941	--	--	--	.909	--	--	.877	--

Note. Correlations for the English and Spanish samples appear on top and below the diagonal, respectively. TSPELL = text spelling; SEG = segmentation; CASE = use of case; WORD = text length in words; LETT = text length in letters; PUNCT = punctuation; CONN = connectors; SUBORD = percentage of subordination; LEXDEN = lexical density; ADJADV = percentage of adjectives and adverbs; LEXDIV = lexical diversity.

[†] Loadings < .50 are not displayed

Table 4.10

Standardized regression coefficient factor loadings for Spanish at Times 3, 4, and 5†

	Time 3		Time 4		Time 5		
	word level	*productivity*	*word level*	*productivity*	*word level*	*productivity*	*vocabulary*
variance explained by factor	60.08%	21.66%	51.57%	25.12%	31.80%	27.03%	19.02%
TSPELL	.689	--	.801	--	.869	--	--
SEG	.931	--	.887	--	.825	--	--
CASE	--	--	--	--	.765	--	--
WORD	--	.912	--	.947	--	.982	--
LETT	--	.858	--	.905	--	.964	--
LEXDEN	---	--	--	--	--	--	.800
ADJADV	--	--	--	--	--	--	.881
LEXDIV	--	.908	--	-.782	--	-.719	--

Note. Correlations for the English and Spanish samples appear on top and below the diagonal, respectively. TSPELL = text spelling; SEG = segmentation; CASE = use of case; WORD = text length in words; LETT = text length in letters; LEXDEN = lexical density; ADJADV = percentage of adjectives and adverbs; LEXDIV = lexical diversity.

† Loadings < .50 are not shown

4.4 Discussion

Study 2 set out to establish the role of orthographic consistency in the early text writing development of children learning to read and write in English as opposed to those learning to write in Spanish. For this purpose, a systematic, longitudinal comparison of children's performance on a number of text-based measures was carried out, making a distinction between word-level and text-level measures. This distinction was supported by the data, in that each group of features showed different rates of development and different degrees of fluctuation as a function of orthographic consistency. Word-level measures, reflected a high rate of improvement especially during the Year 1 time points (Times 3 and 4), revealing this school year as a critical period in the development of word-level writing skills. Interestingly, the role of orthographic consistency was not as important in mid-Year 1 as it would become at the end of Year 1 and, particularly, during Year 2. All three measures involved in this level of analysis—spelling, segmentation, and case—seemed to go hand in hand during the first years of formal schooling in both language groups. Text-level measures, in contrast, improved at a slower pace, without either Year 1 or Year 2 standing out as crucial stages for the development of these features. In addition, text-level measures showed little or no sensitivity to language or orthographic consistency, which tended to have a reduced or inexistent effect—depending on the specific measure. In general, text-level measures showed a small- to medium-size effect of time for both the Time 3-Time 4 and Time 4-Time 5 contrasts, together with marginal, if any, effects of language. In sum, word- and text-level measures differed from each other in at least two ways: first, in terms of their rate of development—where word-level features developed very quickly, especially during Year 1, while text-level features did so at a slow and steady pace. Second, word- and text-level measures were differentially affected by the consistency of the orthography: It is assumed that the main factor moderating development at the word level was orthographic consistency, but no such effect was evident at the text level of analysis.

Possibly, the differences in the rate of development of each level of analysis (word- vs. text-level) are driven by the relative specificity of each skill to the written domain: While word-level features (spelling, word segmentation, capitalization) strictly belong to the transcription component of writing, most text-level features exist in the oral modality as well. This combination of broad component skills in the writing domain is analogous to the 'spelling' and 'ideation' components proposed for reading in the Simple View of writing (Juel et al., 1986). In addition, to the extent that the start of formal schooling implies a sudden shift in attention from the oral to the written domain, it follows that those aspects of writing that are inherent and exclusive to the written modality undergo the most substantial improvement.

Orthographic consistency was shown to moderate those aspects of writing which are specific to transcription, rather than those which exist in other modalities. Most text-level features did not vary as a function of the complexity of the orthography, but they could be determined by factors such as discourse genre, register, or communicative purpose (e.g., Beers & Nagy, 2009; 2010; Berman & Verhoeven, 2002).

Text length (whether measured as number of words or letters) showed a strong effect of time and had a steady rate of improvement, which was significantly more constant for Spanish, than for English. The subtle deceleration of the English group—which nevertheless showed a strong effect of time—, in contrast to the Spanish group, resulted in a significant interaction. Therefore, although in absolute terms the groups performed remarkably similarly on these measures (as can be seen in Table 4.1 and in Figure 4.2), this small interaction might be indicating that orthographic and/or language contrasts might consolidate later on in development.

Text-level measures were characterized by slow but stable growth over time, as well as by small to nonsignificant effects of orthographic consistency. It is important to understand, however, those instances in which even a small language/orthography effect was found. This concerns the differences found in the use of punctuation marks, the percentage of connectives, and the percentage of adjectives and adverbs.

English children produced significantly more punctuation marks than Spanish children, particularly at the earlier time points (Time 3 and Time 4), while the reverse trend was found for the use of connectors. Moreover, the two measures showed a tendency to be negatively correlated in each language. Both connectives and punctuation contribute to text structure and cohesion and are valid alternatives to signal the boundaries of text units. I would like to suggest that, at these early stages, punctuation (which predominantly equates to using a full stop) and connectives (which predominantly equates to using conjunction *and*) are in free variation, meaning that, when one occurs, the other does not. In a sense, the tendency of these two measures to be negatively correlated seems to indicate that using connectives to define and link discourse units (e.g., sentences or clauses) limits the use of punctuation marks, given that they may be functionally redundant. Certainly, the more profuse appearance of punctuation marks in English at Time 3 and Time 4 could also be linked to instruction effects. It is more difficult, however, to explain the higher rate of connectives in Spanish on the same basis. The Spanish children's preference for connectors could be related to the type of layout usually chosen by this group; namely, the paragraph-like layout (see Chapter 3). It might well be the case that linking a series of related events using a paragraph layout requires the use of connectors to a larger extent than punctuation marks, to make such links evident.

An orthographic consistency/language effect was obtained for measures of the percentage of adjectives and adverbs used, such that the English writers produced a higher number of these word types. A plausible explanation for this difference relates to the satellite-framed nature of English—as opposed to Spanish, which is predominantly verb-framed (Slobin, 2004; Talmy, 1985). That is, the direction of English motion verbs is usually expressed by means of a satellite adverb or preposition (e.g., *go* *in/out*, *climb* *up/down*), and this might naturally give rise to an increased adverb count in English. In contrast, in Spanish the verb direction would typically be encoded in the semantic root of the verb (e.g., *salir* 'go out'/*entrar* 'go in'; *ascender* 'climb up'/*descender* 'climb down'; see Section 2.2.3 for a brief explanation of this syntactic contrast between English and Spanish), creating a lesser need for adverbs of direction. Note, however, that this typological difference would only affect the frequency of adverbs, not adjectives. The possible reasons behind the slight advantage of the English group in comparison to the Spanish group must be considered with caution because the proportion of this type of lexical element was small, as was the size of the effect—which was only significant at Time 5. This group effect might more suitably represent a trend rather than a robust cross-linguistic difference.

A secondary aim of this chapter was to test the hypothesized dimensions of writing; namely, word-level/transcription, productivity, text structure, syntactic complexity, and vocabulary. Results have shown that the dimensions of word-level and productivity are well established constructs already by mid-Year 1. This finding is in line with other studies which measured similar construcs: Wagner et al. (2011) and Puranik et al. (2008) both found support for a productivity construct and an accuracy construct, which is akin to the transcription/word-level construct because in all studies it included measures of spelling accuracy and capitalization.

Results also supported an emergent vocabulary factor in English already at Time 4 and at Time 5 in Spanish. Percentage of adjectives and adverbs and lexical density were the variables loading on this factor across languages. This is not surprising, considering that the percentage of adjectives and adverbs is a subset of the percentage of content words, measures by the lexical density measure, but the fact that they were not consistently related in English until the end of Year 1, and in Spanish until mid-Year 2, suggests that some constructs of text-level performance may not consolidate until later stages of writing development. Of interest is the finding that lexical diversity, calculated as the type-token ratio, was not associated with any other vocabulary measure, but loaded heavily on the productivity factor across time points and language groups. Other studies observed a similar behaviour of 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 which poorly captures vocabulary-related aspects of text construction.

In contrast to previous studies assessing the dimensionality of early writing, the present one 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 this type of measures. 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.’s (2008), who obtained support for a complexity factor, may be related to the fact that participants in their study were 3rd and 6th graders 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; Salas, 2010), which is moreover moderated by discourse genre (Beers & Nagy, 2011). Finally, 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 complexity latent factor, may be related to (1) the slightly different types of syntactic measures; (2) the use of a different prompt (explaining which animal they would like as a pet for their class), considering the effect of genre in syntactic complexity (Beers & Nagy, 2009, 2010); and (3) that children were given twice as much time to produce the texts.

In short, the findings of the present study corroborate that learning to write a consistent orthography provides an advantage in the rate of growth in word-level or transcription skills; crucially, spelling. However, the hypothesized impact of orthographic complexity on text-level writing was not fully supported by the data. Most text-level variables did not show variation as a function of orthographic complexity and, when they did, cross-linguistic differences were better interpreted as deriving from non-orthographic factors. The findings for text length, estimated in number of words and letters, suggest that cross-linguistic differences could consolidate later in development, as shown by a small deceleration in the rate of growth for the English group, in contrast to the seemingly linear trend found for the Spanish group. In addition, this study showed that early writing in English and Spanish is composed of similar components. A word-level component and a productivity component were well-established in both language groups from the earliest time point.

The striking cross-linguistic similarity in the average number of words/letters produced still begs the question of whether productivity in written composition is driven by the same cognitive skills in both languages. In other words, even if both groups of children showed virtually identical performances in terms of the amount of text they are able to produce in the same period of time, each language group could be using different strategies and/or resorting to different underlying skills. The extent to which text length, as well as other microstructural features of text composition, is dependent on spelling skills will be the topic of the next section (Study 3, Chapter 5). Finally, the cognitive underpinnings of text-based measures will be the subject matter of Chapter 6 (Study 4).

4.5 Limitations

This study served to identify a number of limitations. First, the cross-linguistic design aimed for participants across countries to be matched with regards to the amount of formal literacy instruction, but 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, while a discussion over whether age or a child's first language is a first-order factor is beyond the scope of this thesis. Future research should endeavour to compare writing development across populations which start literacy instruction at the same chronological age—e.g., comparing Spanish children to English-speaking children from education systems where literacy instruction begins later. Second, estimates of SES—such as parents' education—were only available for about half the sample in each language group. Although this variable was controlled for and shown that it had no impact on the outcome variables (see Appendix 4), it is not possible to determine whether it would have had an effect had it been available for the complete sample of children. These limitations narrow the scope for generalizations of findings in this study to the general population.

Chapter 5

Study 3: The role of spelling in early writing development

Current developmental models of writing assign a very prominent role to spelling and, more generally, to transcription skills, in the emergence and development of text writing abilities. The findings from Study 2 already downplayed the role of spelling in writing development for both consistent and inconsistent orthographies, showing that a similar level of writing competence is observed across language groups in text features beyond the word level, in spite of the marked contrasts in performance on spelling and other word-level writing features. The present study aimed to expand on those findings by testing the claim that writing features are dependent on spelling ability. Even in the absence of differences at the text-level of writing performance, the expression of text-level features might still be dependent on spelling skills to a differing extent in each language group. A novel way to look at the spelling-writing relationship is put forward, proposing a simple-mediation model to test the hypothesis that word spelling influences writing indirectly, via its effect on text spelling. The same participants of Study 1 took part in this study. Spelling was assessed during the text production task and in a single-word dictation task, which was used as an independent predictor, uninfluenced by the demands of the writing task. Results showed support for the mediation model. Word spelling skill was causally related to word-level writing (word segmentation and use of case) and to written productivity through its influence on text spelling. The indirect effect of word spelling on word-level writing was robust and consistent in both language groups across time points. Spelling effects were attested for English until mid-Year 2, while in Spanish the predictive power of spelling was lost by the end of Year 1. The models failed to account for other text-level measures on the basis of their relationship to spelling. The discussion emphasizes (1) the differences between word and text spelling; (2) the distinction between word- and text-level writing; (3) the importance of cross-linguistic, longitudinal, and multiple-feature research on writing development. Finally, the implications of the findings for current models of writing are considered.

5.1 Introduction

Over the last 20 years of writing research, we have witnessed a progressive shift in the attention given to the role of low-level skills in text composition processes. Two of the most influential models of writing, Hayes and Flower's (1980) for skilled writing, and Bereiter and Scardamalia's (1987) for writing development, did not assign a prominent role to the mechanics associated with writing: spelling and handwriting. In the case of Hayes and Flower's model, these skills were assumed to be automatized and, consequently, not part of the problem-space that writing posed to adult, skilled writers. The process of putting ideas generated during planning into written form, the translation process, was the only writing process without further subdivisions or subcomponents in the original formulation of the model (Berninger & Swanson, 1994). In the case of Bereiter and Scardamalia's (1987) model, writing development was more focused on successfully achieving communicative goals in a new medium—i.e., written language. In this sense, writing development was seen as a move from mere notation (of speech) towards the appropriation of the new medium, which becomes a privileged terrain for thought.

In contrast, revisions to these early models of writing, particularly those that were concerned with the development of writing skills (e.g., Berninger & Swanson, 1994; Hayes, 2011, 2012; Juel, et al., 1986), elevated the importance of basic cognitive skills, such as working memory (e.g., Kellogg, 1996; McCutchen, 1986) and, fundamentally, spelling and handwriting. These last two skills were reconfigured as the driving forces in most developmental accounts of writing (e.g., Berninger et al., 1992; 1994; Berninger & Swanson, 1994; Bourdin & Fayol, 1994), as well as in some studies with adult writers (Bourdin & Fayol, 2002; Chenoweth & Hayes, 2001; Hayes, 1996). The influential revision of the original Hayes and Flower's (1980) by Berninger and Swanson (1994)—based on the findings from a series of studies with novice writers (e.g., Berninger et al., 1991; 1992; 1994)—called attention to the lack of specification of subprocesses within the translation component of the model, which were considered of particular relevance to accommodate the development of writing. Thus, they proposed the translation process to be composed of two subprocesses: idea generation, concerned with the translation of prelinguistic content into linguistic ideas; and transcription, concerned with the mechanics of writing (Berninger et al., 1992). In turn, transcription was conceived as being composed of two more basic skills: handwriting and spelling. From their viewpoint, the mechanics of writing constituted the biggest challenge for developing writers, given that they were assumed to use up the best part of available cognitive resources, preventing enough allocation of attention to other (higher level) writing processes. It follows from this account that the most important goal of the first stages in writing development should be to automatize as much and as quickly as possible these basic abilities.

This view of transcription skills as the bottleneck of writing development has had a vast influence on developmental models of writing. Notably, however, the vast majority of studies on the role of transcription in writing development have been carried out with learners of the English orthography, while more consistent orthographies have been largely ignored, with a few exceptions (e.g., Babayiğit & Stainthorp, 2010; Borzone de Manrique & Diuk, 1994; Lerkkanen et al., 2004; Mäki et al., 2001), and no large-scale cross-linguistic studies have been published to date. Given the important differences in the rate of development of spelling as a function of orthographic consistency (e.g., Study 2; Caravolas & Bruck, 1993), it is necessary to test the extent to which English-based notions about the central role of spelling may be extended to more consistent orthographies. The present study thus aimed to determine the relationship between transcription skills and text composition in two cohorts of English- and Spanish-speaking monolinguals over the first year and a half of formal literacy instruction. Specifically, this study explored the degree of dependency of non-transcription microstructural features of writing (e.g., text structure, syntax, text length, etc.) on children's level of spelling skills.

5.1.1 The Transcription-writing Relationship

Studies exploring the role of transcription in writing development have typically compared variations in the quantity and quality of text produced as a function of whether text was produced in a spoken as opposed to a written condition. The spoken condition is typically one where the child is required to *dictate* to an adult the text that he would want to write, thus getting rid of the burden of writing mechanics (e.g., De la Paz & Graham, 1995; Graham, 1990; Hayes

& Berninger, 2009; McArthur & Graham, 1987; Scardamalia et al., 1982). Scardamalia et al. (1982) conducted one of the first studies of that kind. The rationale for these studies was that there are two main sources of interference of writing mechanics on higher-order processes during text construction: One is the lower rate of delivery of writing as compared to speech, which would make it difficult for immature writers to keep up with their thoughts while on the transcription phase. A second main source of interference is the attention that beginner writers devote to mechanical aspects of writing, which would interrupt or disturb higher-level processes. The authors devised an experiment in which these two sources of interference—rate of text delivery and mechanical interference—were manipulated, so that their individual effect on both the quantity and quality of texts produced by students could be assessed. The experiment consisted of three conditions of text production: *written production*, where children were asked to produce a written text; *normal dictation*, where children had to tell (and not write) a similar story to an adult who would write it for him; and *slow dictation*, where the rate at which the adult transcribed the text dictated by the child was adjusted to mimic the child's *written* text delivery rate. The design allowed comparing differences in quality and quantity of text production as a function of rate of delivery (slow vs. normal dictation rate) or as a function of writing mechanics (written vs. normal dictation). Results showed that the mechanics of writing and the lower rate of text delivery affected text length, but not quality: Students produced shorter texts in the written condition than in the slow dictation condition, while texts produced in the normal dictation condition were the longest. It should be noted, though, that the quality of the texts was not tested statistically. These experiments were later replicated by Graham (1990) with learning-disabled students, for whom he had hypothesized that the effect of writing mechanics on text length and quality would be maximized. His results indicated that children produce shorter, as well as lower quality texts in the written production condition. However, the learning-disabled children in Graham's (1990) study produced twice as much text (of similar quality) during the slow dictation condition—which had been hypothesized to prevent efficient flow of ideas and information.

In a more recent study, Hayes and Berninger (2009) looked at the relationships between idea generation and transcription demands. They counted the number of ideas produced by primary-school, English-speaking children in two different task conditions: orally (dictation) and in writing. The tasks were limited in time (5 minutes for the dictation condition and 10 minutes for the handwritten condition). Their results showed that children expressed significantly more ideas in the dictation condition than in the handwriting condition, a finding that was understood as being provoked by the demands of transcription. Older children expressed more ideas than younger ones, a finding that was taken as further support of the effects of automatization of transcription skills; namely, with age and experience with writing, writing mechanics become automatized and exert a smaller demand for cognitive resources. Interestingly, Hayes and Berninger (2009) results contradict those of a previous study (Boscolo, 1990) of similar objectives and comparable methodology, carried out with Italian-speaking children. The Italian children produced more ideas in the written, rather than in the oral text production condition. Hayes and Berninger (2009) argued that the transparent orthography of Italian may be partly responsible for the conflicting results.

In short, these earlier findings pointed to a key role of writing mechanics in the amount of content being generated during the text production process, especially for novice writers. The impact that these aspects had on text quality, however, was not clear-cut. Finally, these studies did not distinguish between the effect of each of the two component processes of transcription; that is, spelling and handwriting. Rather, the differential impact that each of these skills may have on writing fluency and text quality was confounded.

5.1.2 Handwriting and Text Composition

Handwriting has been claimed to be an important predictor of text composition, particularly in early developmental stages (Berninger et al., 1991; Berninger, Yates, et al., 1992). Moderate to strong relationships between handwriting and same-year text composition skills have been reported by a number of studies (e.g., Berninger et al., 1992; Graham, et al. 1997; Jones & Christensen, 1999). Wagner et al. (2011) found evidence of a latent construct of handwriting skills having a significant correlation with a latent construct of macrostructural aspects of text production (ordering of ideas, inclusion of a thesis statement, etc.) in a cross-sectional study with first and fourth graders. Interestingly, the size of the correlations between the handwriting and the text macrostructure constructs was considerably higher in the fourth grade group (r = .81), thus emphasizing the importance of handwriting in early, as well as intermediate writing development. It should be noted, however, that the contribution of handwriting to text composition outcomes appears to be much stronger if measured concurrently (e.g., Berninger, Yates, et al., 1992; Wagner et al., 2011) than longitudinally (Abbott et al., 2010).

A number of studies by Berninger and her colleagues have explored the cognitive underpinnings of handwriting skills in the primary grades (e.g., Berninger & Abbott, 1993; Berninger et al., 1992; 1994), as well as their relationships with text composition skills (Abbott et al., 2010; Berninger, Yates, et al., 1992; Graham et al., 1997). Berninger, Yates, et al. (1992) investigated the low-level developmental abilities that underlie fundamental writing skills: handwriting, spelling, and text composition. Handwriting was evaluated by counting the number of words copied correctly in 1.5 minutes from a paragraph shown to the child. Considerable variance (R^2 = .66) in handwriting ability during primary school was explained by a series of low-level developmental predictors: the alphabet task, where children are asked to write all letters of the alphabet as quickly as possible; whole word (which was only marginally significant, p = .056, p. 270) and letter cluster coding, a task where children are asked to decide whether a printed word is the same as another word (whole word coding) and whether a cluster of letters has appeared in the word shown before in that exact order (letter cluster coding); and finger succession, where children are required to touch their thumb with all other fingers in sequential order as fast as they can until told to stop. Spelling skill, on the other hand, was quite robustly explained by a different set of predictors (R^2 = .60), of which two (whole-word coding and the alphabet task) were shared with handwriting. These predictors were (ordered according to the amount of variance explained, from max. to min.), word attack; nonword reading; visual-motor integration, a task where children need to copy geometric shapes; phonological segmentation; the alphabet task; and verbal IQ. Moreover, spelling and handwriting had a small significant correlation, r = .26, in this large-scale study (N=300). Abbott and Berninger (1993) run structural equation modelling to determine the component skills of handwriting and spelling across the primary and intermediate school years. They found that, in the primary school years, handwriting was explained by motor skills and orthographic coding, while spelling was explained by orthographic coding alone. Collectively, these results seem to point at (1) the fact that spelling and handwriting are slightly related though distinct constructs; and (2) that handwriting is essentially tapping low-level motor skills, in contrast to spelling, which also draws on higher-level cognitive skills, such as verbal IQ.

In short, low-level, motor abilities are at the basis of the distinction between handwriting and spelling skills during the early stages of writing development. In this sense, handwriting seems to be an irrelevant component from a cross-linguistic viewpoint, particularly in contrast with single-word spelling. Consequently, the present study focuses solely on spelling skill and its role in accounting for text-based measures in a consistent as opposed to an inconsistent orthography.

5.1.3 Spelling and Text Composition

The spelling-writing relationship has been assumed to be strongest at the earlier phases of literacy development. While a number of studies have explored this association, the different methodological approaches—particularly in how the text-level has been defined—limit the scope of the conclusions that may be derived of such research efforts. Juel et al.'s (1986) seminal study of early literacy development found small to moderate-strong contributions of spelling to text composition. Notably, the influence of spelling over writing tended to decrease with time.

Graham et al. (1997) found that a latent construct of word spelling predicted writing fluency (number of words in a time-limited writing task) significantly in the primary grades. However, spelling did not contribute to explaining a significant proportion of variance in text quality directly, but indirectly, through a handwriting latent construct.

A more recent study by Abbott et al. (2010) looked at the longitudinal relationships between spelling and text composition in a structural cross-lagged longitudinal model. They tested a model in which spelling predicted text composition in the next adjacent school year. The time span covered the first seven years of formal literacy instruction in the United States (grade1-grade 7), with participants belonging to two cohorts, each comprising 5 years of instruction (i.e., there was a 3-year overlap). The longitudinal cross-lagged models revealed significant paths from spelling to text composition from grades 1 to 7 of small or medium size (range: .25 - .67, p. 290). These findings were taken as support for the key, though modest, role that spelling has in early written text production.

The importance of spelling to text writing fluency and quality was also explored in intervention studies. Berninger et al. (1998) conducted a study that trained either reading and/or spelling, dividing participants into different treatment conditions. Compared to controls, only one group of children who had received spelling intervention showed an increase in text writing fluency, but training spelling did not transfer to text quality scores. Graham, Harris, & Chorzempa (2002) replicated Berninger et al.'s (1998) study and incorporated a number of methodological improvements: They doubled the duration of each intervention session; they included not just text, but also word and sentence production tasks; and they added a maintenance evaluation, six months after post-test. Assessment of the written products consisted of the total number of words (to determine writing fluency) and of a holistic text quality measure. Texts had been typewritten and corrected for handwriting quality, spelling, and punctuation, to avoid possible biases in text quality ratings. Results showed that, on post-test, children in the spelling condition showed improvement in word spelling and sentence construction, but not in compositional length or quality, when compared to controls. In the maintenance assessment, only the effect on spelling had been maintained, but children who had received spelling instruction did not outperform controls in either the sentence construction or text writing tasks.

A few studies have investigated the spelling-writing relationship in more consistent orthographies. Mäki et al. (2001) ran a longitudinal (preschool to Year 3) study looking at the relationship between spelling and text coherence in the very consistent Finnish orthography. Spelling was assessed as the number of errors in single-word and sentence dictation task in Year 1; while the Year 2 and Year 3 assessments consisted of a composite measure comprising errors in spelling, grammar, punctuation, word segmentation, and verb inflection obtained from a narrative text writing task. Year 1 spelling explained a significant but small proportion of variance (R^2 = .17) of Year 2 text coherence. The same path was nonsignificant in subsequent time points. The authors interpreted these results as evidence of the fact that, once the mechanics of writing become relatively automatized, they cease to tax severely on available cognitive resources. Nevertheless, the assessment of "the mechanics of writing" corresponding to Years 2 and 3 may have arguably confounded writing mechanics with aspects pertaining to the text generation component; that is, to non-transcriptional aspects of text construction (Abbott & Berninger, 1993). Lerkkanen et al., (2004) run a 1-year longitudinal study (four time points) to investigate the development of the writing-reading relationship throughout the first year of primary education in Finland. Their measures of single-word spelling accuracy and of spelling accuracy in a text writing task (both consisting of counting the number of correctly spelled words) were consistently related to each other over the four Year 1 measurements, and showed concurrent correlations ranging from .39 to .50, while the longitudinal correlations ranged from .28 to .52. The stronger relationship between the word and text levels in the Lerkkanen et al. (2004) than in the Mäki et al. (2001) study are probably attributable to the fact that text level performance was defined in two very different ways: While the latter related word spelling to a higher-level writing feature of text coherence, the former was mainly relating spelling to spelling in tasks involving different language levels (word and text).

Babayiğit and Stainthorp (2010) conducted a longitudinal study with Turkish-speaking children, looking at the early development of literacy skills. Specifically, they looked at Year 1 predictors of Year 2 writing skills, which included word and sentence spelling accuracy, text spelling error rate, writing fluency, and text structure and content. Word spelling showed moderate to strong associations with sentence spelling and text spelling error rate, but no significant correlations were found between word spelling and writing fluency, text content, or text structure. The same pattern was found for sentence spelling, which did not have a significant association with text spelling error rate either. Their results, thus, point to a reduced or negligible value of spelling skills to explain text-level performance, beyond spelling accuracy. The fact that Babayiğit and Stainthorp (2010) distinguished between different types of writing performance—dissociating spelling accuracy from text length from text quality (content and structure)—provides a strong empirical basis to the claim that consistent orthographies, such as Turkish or Finnish, may not exhibit a consistent spelling-writing relationship if the latter construct is assessed beyond the word level of performance; that is, spelling accuracy.

One study explored the spelling-writing relationship in a Spanish-speaking population (Borzone de Manrique & Diuk, 2003). Children were low-SES, Argentinean 3rd and 4th graders. Children wrote short narrative texts, from which measures of spelling accuracy and writing fluency were obtained. Spelling and writing fluency were moderately correlated in 3rd grade and a strong correlation was found for the 4th graders (r = .52). The authors claim that poor spellers tended to produce lower quality texts, but made no systematic assessment of text quality or of its relationship to word spelling or text fluency.

In short, previous studies on the spelling-writing relationship seem to yield a number of preliminary conclusions: (1) investigations with English speaking children point to a consistently small to medium relationship between spelling and writing, although it seems to be more reliable for writing fluency than for writing quality; (2) studies in more consistent orthographies showed significant relationships among spelling measures across levels of language (word, sentence, or text), but a reduced to negligible role of spelling skill to explain text-level performance. Moreover, the few significant spelling-writing associations in consistent orthographies tended to take place in very early measurements (e.g., Year 1), unless specific populations are considered (e.g., low SES, as in Borzone de Manrique & Diuk, 2003); (3) the type of text-level assessment varies greatly from one study to another, making both within- and cross-linguistic comparisons unfeasible. This study will thus contribute to existing knowledge about the role of spelling in early writing development by providing a direct comparison of two cohorts learning to write orthographies with highly contrastive levels of internal consistency, using an identical methodology for word and text-level assessment.

5.1.4 The Nature of the Spelling-writing Relationship

Even if some studies have failed to find a significant relationship between spelling and writing, a cross-linguistic systematic comparison should help elucidate the true effect of orthographic consistency in early writing development. This is particularly relevant given the unexpected findings of Study 2 (Section 4), where children learning a consistent orthography did not show any clear advantage over their English-speaking peers beyond word-level performance. In this sense, even if children are showing from relatively similar (e.g., lexical density, punctuation by mid-Year 2) to virtually identical average performance (e.g., text length, percentage of subordinated clauses), it may well be the case that the two groups of children are attaining such results following different developmental routes. In a similar vein, those aspects where they do differ (e.g., word-level writing) may be rooted in similar underlying processes. This last scenario has been shown to be the case in two recent studies (Caravolas et al., 2012; 2013). These studies found that the same subset of cognitive skills measured at preschool predicted Year 1 gains in word reading and spelling (Caravolas et al., 2012), as well as initial growth in silent word reading (Caravolas et al., 2013) across languages with inconsistent (English) and consistent (Czech, Spanish, Slovak) orthographies—despite the large differences in the developmental rate of these skills, particularly from the end of Year 1, onwards. Therefore, at least for literacy development, it is possible to observe quantitative differences (i.e., observable behaviour, such as reading scores) that surface from qualitative similarities (i.e., same underlying skills, such as phoneme awareness and letter knowledge). By the same token, the reverse situation, namely, that there may be quantitative similarities (e.g., average text length) in the presence of qualitative different underlying patterns (role of spelling), is also worth testing. In other words, similar developmental rates do not necessarily mean that the underlying developmental patterns are equal; conversely, different developmental rates do not necessarily imply that the underlying patterns are essentially different.

The issue, then, is to establish the nature of the spelling-writing relationship. In this sense, a widespread assumption is that text writing depends on word spelling and not the opposite. There are a series of reasons why this is the case. Firstly, beginner writers may need to interrupt any ongoing writing process to sound a word out or figure out the correct spelling for a particular word. This "...internal dialogue on how to spell a specific word may interfere with higher-order skills (...) [and] may cause the writer to forget plans and ideas already developed or may disrupt the associative process involved in content generation" (Graham 1990, p. 781). Thus, the interruptions that spelling provokes during writing may have a detrimental effect on the final written product. Second, devoting most attention resources to spelling will likely be at the expense of ignoring other aspects of text construction. In this sense, spelling taxes heavily on available cognitive/attention resources, with the logical negative effect in text composition, at least until it has been fairly automatized (Bourdin & Fayol, 1994). Additionally, it has been suggested that children may choose vocabulary and syntactic constructions on the basis of their ease of transcription, rather than on communicatively driven reasons (e.g., Graham, 1990; Mackie & Dockrell, 2004). Finally, struggling spellers may be lacking motivation to write, which would affect the length and quality of their texts (Scardamalia et al., 1982). Indeed, the role of motivation in writing development has been considerably upgraded in recent models (e.g., Hayes, 2012).

The reasoning behind the argument that spelling drives writing development seems so strong, that it has been very scarcely empirically tested. One of such rare studies is Abbott et al.'s (2010) longitudinal study, which tested both directions of the relationship. They found that spelling was a much stronger and more stable predictor of writing than the reverse: Text composition showed significant longitudinal paths to spelling only from Years 3 to 6, which were of small size (range: .12 - .20, p. 289). Therefore, there is a strong (conceptual and empirical) basis supporting the view of spelling as the bottleneck of writing development. However, the inconsistency of results, the variety of assessments and definitions of what constitutes "writing", the lack of support from intervention studies, as well as the contrasting results with more consistent orthographies, call for several clarifications.

This study thus extends previous research on the role of spelling in early writing development in a number of ways. Firstly, it provides within-language tests of the spelling-writing relationship which were conducted using identical measures in two contrasting orthographies: English and Spanish. Second, it presents a longitudinal design, thus avoiding potential spurious findings, while testing the stability of the relationship. Thirdly, it distinguishes between different writing outcomes, including separate measures of text spelling, text length, as well as other word- and text-level writing measures. Last, it avoids subjective quality ratings, and adopts a fully quantitative approach to analyze text-level performance.

Many studies used global/holistic scores to evaluate writing skills (for a review, see Berman & Nir, 2009). There are two main issues with this procedure: On the one hand, spelling may have been one of the aspects evaluated in the assessment of text composition. This may happen if text scores are applied by raters to unedited texts (i.e., not corrected for spelling or punctuation errors) or if the scoring criteria included spelling as one of the aspects being evaluated (as in the standardized test used by Abbott et al., 2010). The result is an unavoidable circularity by which, whatever amount of variance in text production is explained by spelling, it could be argued to be a mere autoregression effect. On the other hand, even if spelling has been carefully set aside from all evaluation criteria, the question still remains as to which (of the many) aspect(s) of text composition is being predicted by spelling skills.

5.1.5. Plan for the Present Study

I propose an alternative way to look at the spelling-writing relationship. If word spelling skill exerts an influence on text composition, this must be the result of the child being able to transfer such skill to the text writing task; that is, it will firstly affect *text* spelling. There is robust evidence that word- and text spelling are related, though distinct, constructs (Abbott & Berninger, 1993). Correlation-based studies looking at the two skills typically show moderate associations (e.g., Babayiğit & Stainthorp, 2010; Lerkkanen et al., 2004). Certainly, the nature and the relevant factors in a single-word writing task are essentially different (and fewer) from those involved in text composition (Abbott & Berninger, 1993; Alamargot & Chanquoy, 2001). Therefore, *word* spelling skill would mainly influence *text* spelling skill which, in turn, may affect other aspects of text construction. This will be referred to as the *indirect* of spelling on writing. Additionally, word spelling may impact text writing directly, for example, if a child selects vocabulary or grammar structures on the basis of how easy/difficult they are to spell. Another example of a direct influence from spelling to writing outcomes could be the negative self-perception or lack of motivation of a child resulting in poor writing performance. The term “direct” might be confusing, since I am claiming that spelling would impact on writing through, for example, vocabulary selection or motivation/self-perception. Therefore, by “direct” I mean “not mediated by text spelling”. Even if slightly misleading, I will keep the terms given that they are consistent with the statistical terminology that will be used to test the relationship. Given that I am only focusing on the mediation effect of text spelling, this is the only indirect effect relevant to the study.

To assess the direct and indirect effect of word spelling in writing outcomes, a series of simple mediation models were run for each time point, language, and text-based measure. Mediation analyses allow exploring not only the relationship between an exogenous variable X and a dependent variable Y—termed the *total effect* or *c path*—, but also the *indirect effect* of X on Y through a mediator variable M. In this sense, M is both a dependent and an independent variable at once. The *indirect effect* is thus defined as the product of the effect of X on the mediator variable M (*a path*) and, in turn, the effect of M on Y (*b path*; Baron & Kenny, 1986).

In the models tested in the present study, word spelling skill was entered as the exogenous variable, while text spelling skill, estimated by the proportion of words accurately spelled in the texts, was entered as the mediator variable. Lastly, all other writing measures were the dependent variables. Both the mediator and the DV were measures obtained in the immediately preceding time point as the independent variable (word spelling). This procedure enables testing of the

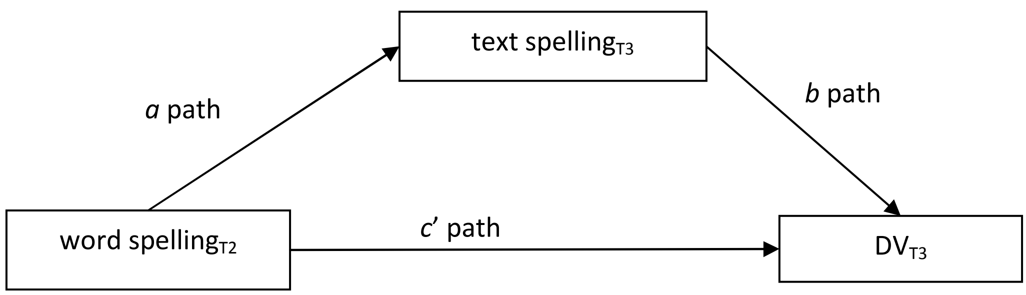


Figure 5.1. Simple-mediation model used to test the direct and indirect effect of word spelling measured at the end of Reception Year (Time 2) on text-based measures (DV) at mid-Year 1 (Time 3). The *a* path indicates the effect of word spelling at Time 2 on text spelling at Time 3. The *b* path indicates the effect of Time 3 text spelling on other text-based measures also at Time 3, controlling for the effect of word spelling. The *c*’ path indicates the direct effect of word spelling on the text-based measure, controlling for its effect on text spelling.

proposed causal model. One model was run for each dependent variable, time point, and language. Models were run using Process, a free-access macro for SPSS developed by Preacher and Hayes (Hayes, 2013; see also Hayes, 2009; Preacher & Hayes, 2008). Process allows one to calculate the direct effect of X on Y (*c'* path), after controlling for the effect of M. In addition, hypothesized covariates may also be entered into the model. Finally, significance *p* values are given for all paths by default[10], as well as estimates of model fit.

Perhaps the most important characteristic of this type of analysis is that it not only explores causal relationships between variables, but it may moreover uncover "hidden" sources of influence between variables not evident in a correlation matrix. According to Preacher and Hayes (2004), "it is quite possible to find that an indirect effect is significant even when there is no evidence for a significant total effect" (p. 719). In this way, the analyses proposed here will explore the full scope of possible influences of spelling on writing performance, both directly (c' path) and indirectly (ab path), at different moments in the early development of writing in two contrasting orthographic systems (Spanish and

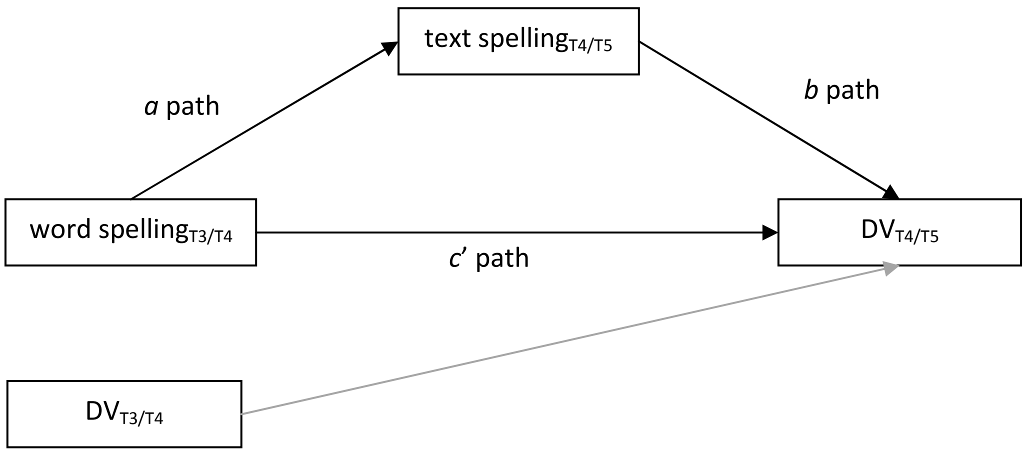


Figure 5.2. Simple-mediation model used to test the direct and indirect effect of word spelling, measured at mid- (Time 3) or end (Time 4) of Year 1, on text-based measures (DV) at Time 4 or Time 5 (mid-Year 2), respectively. In contrast to models from previous time points, these models also control for autoregressive effects.

[10] Except for the indirect effect, which is the product of paths *a* and *b*; in this case, bootstrapped 95% confidence intervals are calculated by the program.

English). For the Time 3 models, the exogenous (word spelling at Time 2), mediator (text spelling at Time 3), and dependent variables (all other writing measures at Time 3) were entered (Figure 5.1). For the Time 4 and Time 5 models (Figure 5.2), the score obtained for the dependent variable at the immediately preceding time point was entered as a covariate, in order to control for autoregressive effects.

It was hypothesized that, (1) given that they constitute tightly related constructs, word and text spelling will be significantly associated in both language groups; (2) given that spelling is, in essence, a word-level measure itself and it showed to be strongly related to word-level measures at this stages (Study 1 and Study 2 above), it will have a stronger effect on word-level writing measures (word segmentation and use of case) than on text-level writing measures (e.g., syntax, vocabulary, etc.); (3) given that children learning a consistent orthography reach near-mastery levels of word spelling by the end of the first year of instruction, it is expected that the predictive value of word spelling (whether direct or indirect) decreases with time more markedly in Spanish than in English; (4) if the proposed mediation model is valid, then the indirect effect of spelling on text writing outcome measures should be stronger and more reliable than the direct effect.

5.2 Method

5.2.1 Participants

The same subsample of English (n = 188) and Spanish (n = 190) participants from Study 2 (Section 4) were chosen for this study. Data from a previous time point (end of Reception Year; henceforth Time 2) were used to estimate children's word spelling skill (see details of the task below). Only children who were present at all relevant times points (Time 2 through to Time 5) are included in this study. The sample used in all subsequent analyses consisted of 151 English-speaking children (79 boys, mean age at Time 3 68.12 months, range: 61-74 months) and 131 Spanish speaking children (73 boys, mean age at Time 3 75.93 months, range: 70-82 months).

5.2.2. Tasks and Procedure

5.2.2.1 Word spelling task (Times 2, 3, and 4). Children were administered a spelling task of isolated words with carrier sentences. The list of words at Time 2 consisted of 9 and 7 words in English and Spanish, respectively. At Times 3 and 4 it included an additional 27 words in English and 36 words in Spanish. Words increased in their complexity, starting with high frequency monosyllabic words of high phoneme-to-grapheme consistency, progressively incorporating longer tokens with more complex syllabic structures and less consistent phonographic mappings. The administrator pronounced each target word, then used it in a carrier sentence, and repeated it one more time. For example, for item *lock* the administrator said, "'Lock.' As in *I lock the door.* Write 'lock'." Both the target items and the carrier sentences were uttered with normal speech rate and intonation, without making any particular emphasis. Children were encouraged and praised for any attempts at spelling the words and were given no feedback if they asked how to write a particular word. At Time 2, the task was carried out individually. At subsequent time points, primer words (the original set of 7-9 words in each language) were also administered individually, while the remaining items were administered in small groups of 4 to 5 children. The task was never discontinued. Each word spelled correctly received a score of 1. Percentage accuracy scores were calculated, since the number of items was different in each language. Cronbach's α for this task at Time 2 was .81 in English and .87 in Spanish.

5.2.2.2 Text writing task (Times 1, 2, and 3). See complete details of the characteristics and procedure for this task in Section 3.3.

5.2.3 Writing Measures

Except for text spelling accuracy, here the mediator variable, the same set of writing measures used in Study 2 constituted the dependent variables of the present study. The distinction between word-level/transcription measures and text-level measures was kept. Word-level variables were word segmentation and use of case. Text-level variables included two measures of text structure: punctuation and percentage of connectors; two measures of syntactic complexity: words per clause and percentage of subordination; and three measures of vocabulary: lexical density (percentage of content words), percentage of adjectives and adverbs, and average content-word length (in letters). Recall that this last measure was excluded from studies making direct cross-linguistic comparisons because of the possible bias resulting from the different average word length in each language (see Section 3.7.15). Finally, writing productivity, estimated by text length in words, constituted a separate group. The coding criteria for each of these measures can be found in Section 3.6.

5.3 Results

In this section, I report the descriptive statistics for the word spelling measure, the longitudinal correlations between word spelling and text-based measures and the concurrent correlations between text spelling and the rest of the text-based measures. Finally, results of the simple mediation analyses are reported of all measures that showed a significant association with either word or text spelling at each time point (Times 3, 4, and 5).

5.3.1 Descriptive Statistics

Table 5.1 shows the mean scores and standard deviations for all text writing measures and the word spelling task in each language and time point. Spanish and British children showed similar word spelling skills at Time 2, $t(374.16) = 0.77$, $p = .442$, $d = 0.67$; but Spanish children quickly outperformed their peers at Time 3, $t(214.48) = -4.84$, $p < .001$, $d = -0.59$; and Time 4, $t(235.50) = -17.83$, $p < .001$, $d = -2.16$. The effect sizes confirm that a very large advantage for the Spanish group emerges by the mid Year 2.

5.3.2 Correlations

Table 5.2 shows the longitudinal correlations between word spelling and all text-based measures and the concurrent correlations between text spelling and the remaining text-based measures. Only longitudinal correlations between adjacent time points were explored, since they are the ones that were used in the mediation analyses. Exploring the pattern of correlations between both spelling measures is a crucial step prior to conducting mediation analyses.

Table 5.1

Means and SDs for word spelling at Time 2, Time 3, and Time 5 and text-based measures at Time 2, Time 3, and Time 4

	Time 3		*Time 4*		*Time 5*	
	English	Spanish	English	Spanish	English	Spanish
	M (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*M* (SD)
WSPELL[1]	26.05% (11.32)	35.10% (18.54)	26.05% (11.32)	35.10% (18.54)	36.44% (14.40)	63.02% (9.38)
TSPELL[2]	47.34% (23.77)	55.30% (26.66)	64.55% (20.62)	81.06% (12.64)	70.12% (18.95)	85.16% (9.91)
SEG[2]	48.22% (36.54)	54.88% (34.22)	69.11% (33.04)	87.66% (14.24)	80.22% (25.79)	92.00% (10.62)
CASE[2]	65.31% (26.27)	64.71% (32.65)	75.40% (21.89)	91.18% (11.47)	77.27% (19.18)	94.08% (5.60)
WORD[3]	15.63 (8.52)	14.41 (8.56)	25.60 (13.67)	24.08 (8.69)	32.12 (16.20)	34.63 (13.80)
PUNCT[4]	5.82 (11.31)	1.43 (5.49)	5.95 (6.36)	4.11 (5.58)	6.75 (6.68)	5.45 (4.99)
CONN[2]	4.21% (6.91)	6.58% (8.01)	5.70% (5.96)	9.34% (6.63)	8.22% (5.82)	10.55% (5.90)
WCL	5.09 (1.32)	5.00 (2.62)	5.85 (2.03)	5.82 (2.17)	5.58 (1.24)	5.66 (1.59)
SUBORD[6]	5.89% (11.84)	6.99% (15.86)	5.98% (11.12)	9.14% (14.60)	12.61% (15.35)	12.59% (14.81)
LEXDEN[7]	46.56% (17.00)	48.78% (17.29)	43.67% (11.61)	44.22% (8.35)	45.71% (9.52)	44.21% (6.64)
ADJADV	2.31% (4.33)	1.98% (3.72)	4.82% (5.79)	3.70% (4.86)	6.28% (5.41)	4.83% (4.34)
WLEN	4.76 (1.41)	5.08 (1.33)	4.89 (1.04)	5.38 (7.17)	4.85 (0.92)	5.38 (5.23)

Note. WSPELL = word spelling; TSPELL = text spelling; SEG = segmentation; CASE = use of case; WORD = text length in words; PUNCT = punctuation; CONN = connectors; WCL = words per clause; SUBORD = subordination index; LEXDEN = lexical density; ADJADV = adjectives and adverbs; WLEN = average word length in letters. [1]Mean accuracy in word spelling task in the immediately preceding time point. [2]Mean percentage out of all words in text. [3]Mean number of words. [4]Number of punctuation marks over total number of words multiplied by 100. [5]Percentage of texts written in paragraph layout. [6]Mean percentage of subordinated clauses out of all clauses in text. [7]Mean percentage of content words.

Even though the break-down of the effect that is carried out in a mediation analysis may reveal significant relationships in the absence of a significant total effect (i.e., a simple correlation between the exogenous and the outcome variable; Preacher & Hayes, 2004), there needs to be some indication that spelling and the different writing measures are related. For this reason, the criteria applied here was such that DVs which were not significantly correlated with *either* word (longitudinally) or text spelling (concurrently) would not be subjected to mediation path analyses. Additionally, given that large sample sizes are used and even low associations are likely to render significant results, probability levels were adjusted to the $p < .01$ level.

5.3.2.1 Correlations between word-spelling and text-spelling. The longitudinal correlations between word spelling and text spelling were high in English across time points (Table 5.2); in Spanish, they were also high, with the exception of the Time 3 word spelling-Time 4 text spelling correlation, which was of small size, $r = .29$. Thus, this pattern of correlations is in line with predictions that both spelling measures are tapping into a similar skill, while constituting separate constructs.

5.3.2.2 Correlations between spelling measures and word-level writing measures. Time 2 word spelling and Time 3 text spelling showed high correlations with all Time 3 word-level measures: segmentation and use of case (Table 5.2). This pattern applied to both English and Spanish. The English groups maintained a consistent pattern of relationships between spelling (word or text) and word-level writing up until mid-Year 2 (Time 5). In Spanish, word spelling inconsistently predicted use of case: Time 2 and Time 4 word spelling showed low to medium correlations with Time 3 and Time 5 use of case, respectively, while there was a nonsignificant correlation in between Time 3 word spelling and Time 4 use of case. The concurrent correlations between text spelling and use of case were much more stable, especially at Time 3 and Time 4, and slightly higher at Time 5 ($r = .32$, at Time 3; $r = .35$, at Time 4; and $r = .53$ at Time 5). In sum, both spelling measures were strongly related to word-level writing measures throughout the study. The size of the correlations was, in general, higher for the English sample, which also showed a more consistent pattern of relationships.

5.3.2.3 Correlations between spelling measures and text length. Both spelling measures (word and text) showed consistently medium to high correlations with text length in the English sample. Correlation coefficients rangedfrom .36 to .58 between word-spelling and length, and from .42 to .60 between text-spelling and length. Note that correlation coefficients tended to decrease with time (Table 5.2). In the Spanish group, correlation coefficients at the initial time points were high, r = .53 and r = .57, for the correlation between Time 3 length and Time 2 word- and Time 3 text-spelling, respectively. However, the tendency of correlation coefficients between spelling and text length to decrease with time was more marked in Spanish. Time 1 word spelling and Time 4 text spelling were not as strongly associated with Time 4 text length in Spanish as in the corresponding correlations in the previous time point, with r values falling below .30. By Time 5, text length was no longer significantly correlated in Spanish with either spelling measure. In sum, there was a strong text length-spelling association from the start to mid-Year 1 in both language groups. However, it seems that the predictive strength of spelling over this feature of text construction tends to lose power with time. Its role became practically negligible in Spanish by Year 2, while it still remained relevant in English until the end of the study.

5.3.2.4 Correlations between spelling and text-level writing measures. The pattern of correlations between spelling (word and text) and text-level measures (syntax, vocabulary, text structure) was characterized by low correlations in both languages (Table 5.2). In English, word spelling significantly predicted punctuation and the percentage of adjectives and adverbs throughout the study. It also predicted the percentage of connectors, except at Time 5, and lexical density at Time 4 and Time 5. A significant, low correlation was also obtained for Time 3 word spelling and Time 4 percentage of subordination. The concurrent correlations between text spelling and text-level measures were remarkably similar.

Table 5.2

Product-moment correlations of exogenous and mediator variables with text-based measures

	TSPELL3	SEG3	CASE3	WORD3	PUNCT3	CONN3	WCL3	SUBORD3	LEXDEN3	ADJADV3	WLEN3
WSPELL2-Eng	.75*	.60*	.63*	.58*	.29*	.22*	-.06	.19	.12	.32*	-.21
TSPELL3-Eng	-	.71*	.77*	.60*	.16	.35*	-.05	.12	.13	.25*	-.27*
WSPELL2-Spa	.63*	.44*	.25*	.53*	.15	.27*	.22*	-.14	-.02	.15	.12
TSPELL3-Spa	-	.43*	.32*	.57*	.08	.35*	.23*	-.05	-.03	.21	-.13
	TSPELL4	SEG4	CASE4	WORD4	PUNCT4	CONN4	WCL4	SUBORD4	LEXDEN4	ADJADV4	WLEN4
WSPELL3-Eng	.69*	.50*	.47*	.38*	.21*	.23*	-.12	.25*	.21*	.23*	-.01
TSPELL4-Eng	-	.68*	.71*	.54*	.25*	.27*	-.09	.22*	.20*	.18	-.15
WSPELL3-Spa	.29*	.35*	.02	.23*	.12	.09	.11	.05	-.10	-.10	.03
TSPELL4-Spa	-	.43*	.35*	.29*	.09	-.10	.06	.10	-.03	-.22	-.18
	TSPELL5	SEG5	CASE5	WORD5	PUNCT5	CONN5	WCL5	SUBORD5	LEXDEN5	ADJADV5	WLEN5
WSPELL4-Eng	.71*	.38*	.41*	.36*	.20*	.19	.17	.06	.33*	.30*	-.01
TSPELL5-Eng	-	.65*	.65*	.42*	.15	.25*	.11	.09	.34*	.29*	-.18
WSPELL4-Spa	.50*	.46*	.31*	.13	.15	.11	-.17	.08	.18	-.11	-.20
TSPELL5-Spa	-	.61*	.53*	.03	.04	.09	-.09	.08	.28*	-.06	-.06

Note. WSPELL = word spelling; TSPELL = text spelling; SEG = segmentation; CASE = use of case; WORD = text length in words; PUNCT = punctuation; CONN = connectors; WCL = words per clause; SUBORD = subordination index; LEXDEN = lexical density; ADJADV = adjectives and adverbs; WLEN = average word length.
* = $p < .01$.

Text spelling in English correlated significantly with connectives at all three time points, with lexical density at Time 4 and Time 5, with percentage of adjectives and adverbs at Time 3 and Time 5, and with percentage of subordination at Time 4. In contrast with word spelling, text spelling correlated significantly with punctuation only at Time 4. Surprisingly, text spelling in English showed a significant, negative correlation with word length at Time 3, of small size, $r = -.27$.

The Spanish sample presented, in general, fewer significant correlations with text-level measures in comparison to English, particularly during Time 4 and Time 5. At Time 3 text spelling correlated significantly with percentage of connectors and the word-per-clause measure. No significant correlations were obtained between text spelling and text-level measures at Time 4 and only one was observed at Time 5, with lexical density, $r = .28$.

To sum up, in contrast with the strong longitudinal and concurrent associations between spelling and word-level measures, text-level measures were not consistently predicted by word spelling and did not show strong relationships with a concurrent measure of text spelling. The low value of the few significant correlations, together with the longitudinal instability of the patterns in both languages, point to a lack of reliability of the relationship between spelling and text-level writing.

5.3.3 Simple-mediation Analyses

A series of simple mediation analyses was run to test the direct and indirect influence of word spelling skill on the various measures obtained from the text writing task. In the models tested, word spelling skill, estimated by the scores in the word writing task, was entered as the exogenous variable, while text spelling skill, estimated by the percentage of words accurately spelled in the texts, was entered as the mediator variable.

Unstandardized coefficients and standard errors for the direct and indirect paths from word spelling to Time 3, Time 4 and Time 5 dependent variables are shown in Tables 5.3, 5.4, and 5.5., respectively. Unstandardized coefficients are preferred over standardized values to facilitate interpretation (Preacher & Hayes, 2004; 2008). Moreover, effect sizes for all mediation models are also provided (R^2). Note that p values can only be provided for the direct path. For the indirect path 95% confidence intervals (5,000 bootstrap samples) are reported; if zero is not a value within the confidence intervals, then the indirect effect is assumed to be significantly different from 0 at $p < .05$ (two-tailed; Preacher & Hayes, 2004, p. 722).

5.3.3.1 Word spelling and text spelling (*a* paths). The path from word spelling to later text spelling was always significant in both language groups and at all time points. In the case of Time 4 and Time 5 text spelling, the path was significant even after controlling for autoregressive effects of the dependent variable, although including the covariate resulted in slight variations in the exact *a*-path coefficients.

5.3.3.2 Text spelling and text-based measures (*b* paths). B paths are those measuring the concurrent effect between the moderator variable (text spelling) and the dependent variable, controlling for the effect of the exogenous predictor (word spelling at the immediately previous time point). The significance of *b* paths in the models conducted varied according to the type of outcome variable and the time point under scrutiny. For word-level variables (i.e., segmentation and case), a significant *b* path was always obtained even after controlling for autoregressive effects in the Time 4 and Time 5 models. Significant *b* paths were also found for text length across time points and languages (though note that the Time 5 model for text length was not run in Spanish due to the absence of a significant correlation with either spelling measure). The *b* path to Time 3 connectives was significant in Spanish and English. The same path was nonsignificant at later time points in English, the only language group for which this model was conducted (because in Spanish the bivariate correlations were nonsignificant). Finally, Time 3 punctuation and Time 3 word length showed significant *b* paths in English. The Time 3 text spelling-word length path was of negative value, although 95% confidence intervals showed that the true value of the effect was very close to zero (lower limit= -0.02; upper limit = -0.001). In short, while word-level measures were virtually always affected by text spelling, text-level measures were mostly unaffected. The main exception was constituted by the length of texts, which showed to be consistently determined by concurrent text spelling skill.

Table 5.3

Results of Time 2-Time 3 mediation models

	English					Spanish				
	indirect effect		*direct effect*			*indirect effect*		*direct effect*		
	β (SE)	95% C.I LLCI - ULCI	*β (SE)*	95% C.I LLCI - ULCI	R^2	*β (SE)*	95% C.I LLCI - ULCI	*β (SE)*	95% C.I LLCI - ULCI	R^2
SEG	2.35 (0.39)	1.58 – 3.13	0.84 (0.48)*	-0.10 – 1.78	.51*	0.99 (0.47)	0.04 – 1.89	1.78 (0.63)*	0.53 – 3.03	.48*
CASE	1.96 (0.29)	1.41 – 2.57	0.51 (0.31)	-0.11 – 1.12	.60*	1.15 (0.47)	0.31 – 2.16	0.33 (0.65)	-0.93 – 1.58	.12*
WORD	38.95 (9.02)	21.91 – 56.95	33.94 (12.25)*	9.72 – 58.16	.41*	38.50 (8.93)	23.33 – 58.57	42.80 (13.98)*	15.14 – 70.47	.37*
PUNCT	-7.55 (14.17)	-38.45 – 17.95	43.91 (15.42)*	13.42 – 74.41	.08*	—	—	—	—	—
CONN	29.68 (8.64)	14.27 – 48.09	-10.21 (10.64)	-31.25 – 10.83	.13*	25.69 (8.80)	10.35 – 45.32	10.95 (14.37)	-17.49 – 39.39	.13*
WCL	—	—	—	—	—	3.62 (3.73)	-3.43 – 11.29	6.33 (5.04)	-3.65 – 16.30	.06
SUBORD	—	—	—	—	—	—	—	—	—	—
LEXDEN	—	—	—	—	—	—	—	—	—	—
ADJADV	2.99 (4.80)	-7.08 – 12.08	16.39 (7.29)	1.98 – 30.81	.10*	—	—	—	—	—
WLEN	-2.31 (1.29)	-4.85 – 0.24	-.052 (1.57)	-3.63 – 2.58	.08*	—	—	—	—	—

Note. SEG = segmentation; CASE = use of case; WORD = text length in words; PUNCT = punctuation; CONN = connectors; WCL = words per clause; SUBORD = subordination index; LEXDE = lexical density; ADJADV = adjectives and adverbs; ADJADV = average word length in letters. LLCI: Lower-level confidence interval; ULCI: Upper-level confidence interval.

* p < .01.

Table 5.4

Results of Time 3-Time 4 mediation models

	English					**Spanish**				
	indirect effect		*direct effect*			*indirect effect*		*direct effect*		
	β (SE)	95% C.I LLCI - ULCI	*β (SE)*	95% C.I LLCI - ULCI	R^2	*β (SE)*	95% C.I LLCI - ULCI	*β (SE)*	95% C.I LLCI - ULCI	R^2
SEG	0.64 (0.17)	0.36 – 1.02	-0.08 (0.23)	-0.53 – 0.38	.55*	0.06 (0.03)	0.02 – 0.14	0.11 (0.07)	-0.02 – 0.24	.27*
CASE	0.63 (0.15)	0.38 – 0.96	-0.26 (0.16)	-0.57 – 0.05	.54*	0.05 (0.02)	0.02 – 0.10	-0.05 (0.04)	-0.12 – 0.04	.13*
WORD	29.38 (7.91)	15.48 – 45.72	-10.17 (10.85)	-31.60 – 11.27	.35*	2.09 (1.24)	0.30 – 5.47	2.06 (4.53)	-6.91 – 11.02	.14*
PUNCT	8.32 (5.20)	-2.74 – 17.85	4.62 (6.42)	-8.08 – 17.33	.07	—	—	—	—	—
CONN	5.99 (3.76)	-1.31 – 13.77	3.01 (5.71)	-8.27 – 14.30	.12*	—	—	—	—	—
WCL	—	—	—	—	—	—	—	—	—	—
SUBORD	6.47 (6.50)	-6.38 – 19.70	16.11 (10.52)	-4.67 – 36.89	.07	—	—	—	—	—
LEXDEN	5.94 (10.07)	-14.06 – 25.27	9.91 (10.72)	-11.27 – 31.09	.09*	—	—	—	—	—
ADJADV	0.81 (3.47)	-6.14 – 7.60	9.60 (5.59)	-1.45 – 20.66	.08*	—	—	—	—	—
WLEN	—	—	—	—	—	—	—	—	—	—

Note. SEG = segmentation; CASE = use of case; WORD = text length in words; PUNCT = punctuation; CONN = connectors; WCL = words per clause; SUBORD = subordination index; LEXDEN = lexical density; ADJADV = adjectives and adverbs; AVWL = average word length in letters. LLCI: Lower-level confidence interval; ULCI: Upper-level confidence interval.

* $p < .01$.

Table 5.5

Results of Time 4-Time 5 mediation models

	English					Spanish				
	indirect effect		*direct effect*			*indirect effect*		*direct effect*		
	β (SE)	95% C.I LLCI - ULCI	*β (SE)*	95% C.I LLCI - ULCI	R^2	*β (SE)*	95% C.I LLCI - ULCI	*β (SE)*	95% C.I LLCI - ULCI	R^2
SEG	0.49 (0.12)	0.29 – 0.74	-0.29 (0.15)	-0.58 – 0.01	.55*	0.26 (0.07)	0.14 – 0.43	0.18 (0.09)	0.003 – 0.35	.41*
CASE	0.25 (0.08)	0.09 – 0.43	-0.06 (0.10)	-0.26 – 0.15	.51*	0.14 (0.04)	0.07 – 0.22	0.04 (0.05)	-0.06 – 0.14	.30*
WORD	6.65 (6.55)	-6.77 – 18.85	9.19 (9.92)	-10.42 – 28.80	.40*	-0.66 (6.27)	-11.76 – 13.12	8.13 (12.71)	-17.03 – 33.29	.22*
PUNCT	-0.57 (3.79)	-8.49 – 6.45	6.97 (4.83)	-2.57 – 16.52	.10*	—	—	—	—	—
CONN	5.22 (3.33)	-1.64 – 11.53	0.88 (4.38)	-7.77 – 9.53	.10*	—	—	—	—	—
WCL	—	—	—	—	—	—	—	—	—	—
SUBORD	—	—	—	—	—	—	—	—	—	—
LEXDEN	3.31 (5.26)	-6.70 – 13.74	10.97 (5.59)	-0.07 – 22.02	.23*	7.63 (4.16)	-0.19 – 16.20	6.70 (6.84)	-6.84 – 20.25	.10*
ADJADV	2.34 (2.37)	-2.17 – 7.20	7.43 (3.96)	-0.40 – 15.26	.12*	—	—	—	—	—
WLEN	—	—	—	—	—	—	—	—	—	—

Note. SEG = segmentation; CASE = use of case; WORD = text length in words; PUNCT = punctuation; CONN = connectors; WCL = words per clause; SUBORD = subordination index; LEXDEN = lexical density; ADJADV = adjectives and adverbs; AVWL = average word length in letters. LLCI: Lower-level confidence interval; ULCI: Upper-level confidence interval.

* $p < .01$.

5.3.3.3 Direct and indirect effects of word spelling to text-based measures. Tables 5.3, 5.4, and 5.5 show the results of the indirect paths (i.e., the product of *a* and *b* paths) from word spelling to text-based measures at Time 3, Time 4, and Time 5. The indirect effect of word spelling on subsequent text-writing performance indicates whether the effect of word spelling on writing takes place via text spelling, since children may only be able to transfer to some extent their spelling skills to a text writing task, where multiple different demands and constraints are at play. In line with predictions, the indirect effect of word spelling on subsequent (next time point) word-level writing measures was robust in both language groups across time points. The indirect effect of word spelling on subsequent (next time point) text length was also significantly different from zero across time points and languages, except for Spanish at Time 5, when this model was not run due to the absence of a significant total with either spelling measure. Finally, the indirect effect of word spelling on subsequent text-level performance was unreliable across the board.

Given the virtual lack of an indirect effect of word spelling on subsequent text-writing performance, it was possible that word spelling had a direct effect (that is, not mediated by text spelling) on later text-writing performance. However, very few direct effects were significant and/or their 95% confidence intervals showed the coefficient to be unreliable, probably because it was not sufficiently different from zero (Preacher & Hayes, 2004). Significant and reliable (i.e., confidence intervals not including zero) direct effects (*c'* paths) were obtained for text length at Time 3, in both language groups, segmentation in Spanish at Time 3, and both punctuation and proportion of adjectives and adverbs at Time 3 in English (Table 5.3).

5.3.3.4 Autoregressive effects. The effect of the same skill at the immediately preceding time point was included as a covariate in Time 4 and Time 5 models to control for autoregressive effects. In the English sample autoregressive effects were always significant, with a few exceptions constituted by the syntactic complexity measures and percentage of adjectives and adverbs in the Time 4 models; in the Time 5 models, syntactic complexity, punctuation, and percentage of adjectives and adverbs did not show a significant auoregressive effect either. In Spanish, many of the paths from the covariate to the dependent variable were nonsignificant. This is not surprising given the inconsistency among text-level measures, already noted in Study 2 (Section 4).

5.4 Discussion

This study set out to explore the idea that transcription skills are the bottleneck of writing development. The study is unique in that it presents cross-linguistic data, looking at the text construction skills of children learning two contrasting orthographies, Spanish and English, which differ greatly in the consistency of their phonographic mappings. Several previous studies have explored the role of transcription skills in writing development in English (e.g., Berninger, Yates, et al., 1992; Graham, 1990; Scardamalia et al., 1982) and in other more consistent orthographies (e.g., Babayiğit & Stainthorp, 2010; Borzone de Manrique & Diuk, 2003; Lerkkanen et al., 2004), but the inconsistency of results, the variety of methods employed, and the different definitions of what constitutes "writing" prevented drawing any conclusions. In order to test the various ways in which word spelling might exert an influence in text writing, a simple-mediation model was proposed to help elucidate the nature of the relationship between spelling skills and text composition. This model allowed an examination of whether word spelling (the exogenous variable) influenced text writing directly and/or indirectly, via its effect on text spelling, the mediator variable. The model proposed assumed, therefore, that word and text spelling constitute related, though distinct constructs. This assumption was corroborated by the data: Word and text spelling showed high correlations across time points and language groups, and the path in the mediation models from word to text spelling was always significant.

In Studies 1 and 2, above (Sections 3 and 4, respectively), a split was observed among writing measures, by which spelling, segmentation, and use of case seemed to constitute a natural group, given that they (1) showed high correlations in both languages and at all time points; (2) showed to be greatly affected by orthographic consistency; and (3) showed a similar developmental rate, with remarkably strong improvements during Year 1. In contrast, text-level variables (except for text length) were characterized by a slow developmental rate and a small to null effect of orthographic consistency. It was therefore hypothesized that word spelling would impact greatly on word-level variables, rather than on text-level variables. This hypothesis was also corroborated by the data. Word spelling was an excellent predictor of word-level writing in both consistent and inconsistent orthographies, while it failed to predict text-level writing in a robust, consistent way. Instances of a significant relationship between word spelling and text-level writing were of rare occurrence and small in size, with low proportions of variance explained. In contrast, the models predicting word-level writing tended to show large R^2 values, especially in English. The finding that, at the word-level of writing performance, both language groups are showing very similar patterns is not trivial, since the mean differences between the English and the Spanish groups were large, especially at Time 4 and Time 5, both for text spelling (Section 4) and word spelling (see Section 5.3.1). Thus, in spite of behavioural differences in word spelling performance, children learning consistent and inconsistent orthographies present the same underlying patterns.

It was hypothesized that the effect of spelling in text writing would tend to decrease in Spanish after the first year of formal literacy instruction had elapsed. This prediction was partially supported by the data. The effect of word spelling on word-level writing (segmentation and case) fluctuated in Spanish, but by mid-Year 2 the effect was still significant. However, the effect on text length was in line with the departing hypothesis. Already in Study 2, text length had revealed a different pattern of behaviour, by which it showed a fast, almost linear, developmental rate, thus aligning itself with word-level measures. However, it was unaffected by orthographic consistency, with both language groups revealing nearly identical mean counts at each time point. Nevertheless, word spelling was a good predictor of text length in English across time points, accounting from 35-40% of the variance. In Spanish, a very similar result was obtained for text length at mid-Year 1 (Time 3), but the predictive power of spelling decreased with time. By mid-Year 2, there were no significant correlations between either word or text spelling and text length in this language group. This suggests that transcription skills,—specifically, spelling—may have a slightly more protracted effect in inconsistent than in consistent orthographies.

In this study a novel way of accounting for the influence that spelling exerts on writing was proposed, where word spelling was hypothesized to impact text writing indirectly through its effect on text spelling. The mediated relationship was supported by the data, since word spelling exerted its effect primarily via its influence on text spelling which, in turn, influenced text-based measures. Only rarely did word spelling directly impact on text writing and, if so, it was an additional effect over and above the indirect path. This finding is not surprising, since children may only be able to transfer their spelling skills to a text writing task to some extent, considering the several additional demands of the writing process. The degree to which children will effectively take full advantage of their spelling skills during text composition may be dependent on a number of factors, both cognitive and psychological, such as self-regulation strategies, attitudinal factors, and so on. Interestingly, this account of the role of spelling in writing is very much in line with the *simple view* triangle, which proposes that higher-level writing (the top vertex) is supported by *both* transcription and executive function skills (Berninger & Amtmann, 2003; Berninger & Winn, 2006). I would suggest that the present findings point to the interrelationships among these two constructs at the base of the triangle.

Previous literature has insisted on the complexity of the writing process and the high cognitive demands that spelling imposes on beginner writers (e.g., Berninger & Swanson, 1994; Bourdin & Fayol, 1994). Indeed, novice writers have not yet automatized spelling and must therefore devote considerable attention resources to spelling, at the expense of reducing or neglecting attention to other aspects of text construction. Several explanations have been proposed about the likely reasons why a focus on transcription has a detrimental effect on the written product. However, what is not usually considered is the detrimental effect that writing may have on spelling skills during composition. Certainly, even the simplest writing task would presumably interfere with the child's word-spelling processes and strategies. Even if children reduce attention and ignore certain component processes of writing—such as it has been suggested that children do with planning or revising (Bereiter & Scardamalia, 1987; Glynn et al., 1982; McCutchen, 1996)—, content generation (as in the writing task used in this study) or content retrieval (as in a retelling task) cannot be avoided. Sentence parsing cannot be ruled out even in a simple text copying task in the same way that an oral sentence repetition task taps into syntactic (and not just phonological) processing (e.g., Fattal, Friedmann, & Fattal-Valevski, 2011; Lust, Flynn, & Foley, 1996). It may be argued, nonetheless, that children in both language groups showed higher mean accuracy scores in the text, rather than the word-spelling task. However, while tokens in the word spelling task are determined by the researcher (and chosen specifically because of the types of inconsistencies they presented), children are at freedom to choose the vocabulary of their texts, thus resulting in superior spelling scores in the writing task.

To conclude, this study has provided support to the idea that single-word spelling supports certain aspects of the writing process: word-level writing and text length. It has strengthened the validity of the multiple-feature approach, given that it allowed distinguishing the differential effect of transcription at different writing levels and domains. Cross-linguistic observations revealed that children in both language groups show similar performance levels and underlying patterns—crucially, in the role of spelling in writing—during a foundational stage, but start to diverge around Year 2. Investigations involving a longer time span are needed in order to determine whether such differences are qualitative or quantitative in nature. The tendencies obtained for Spanish, such as the loss of predictive power of word spelling on text length, could also be operative in English, though only evident later in development.

5.5 Limitations

The lack of consistent relationships within text-level measures across time points is major concern. It could be that such a fine grained analysis of certain domains (e.g., syntax or vocabulary) is not entirely adequate for these developmental stages, especially in combination with the type of text that was required of the children. Taking into account the short average length of the texts produced by children at these stages, having such a detailed breakdown of counts and measures may not give sufficient statistical power for true effects to be detected. For this reason, the next study will combine the quantitative approach with an analytical strategy, aggregating text-level measures. Note, however, that the more consistent results for the text length measure are reassuring, in the sense that they may be considered as a (imperfect) proxy for other aspects of higher-level writing competence (Berninger, Yates, et al., 1992).

Chapter 6

Study 4: Cognitive predictors of early writing skills in English and Spanish

Identification of the cognitive skills underlying the development of literacy is crucial to obtaining a better understanding of the nature of the target abilities, as well as to improving the means for early detection and remediation of poor achievers. The present study followed the predictors of mid-Year 1 to mid-Year 2 text writing skills in two alphabetic orthographies differing in terms of orthographic consistency: English, a prototypically inconsistent orthography, and Spanish, a very consistent one. Multigroup Structural Equation Models (SEM) were carried out, including predictor variables that assessed a series of cognitive skills. The dependent variables were obtained from a semi-structured text-writing task and included spelling accuracy, text length, scores from a written vocabulary factor, and an analytic score of text-level writing which comprised children's scores on different microstructural writing features. Results showed that (1) a small set of cognitive predictors (word spelling, phoneme awareness, and RAN) accounted for considerable proportions of variance in text spelling and text length; (2) the contribution of these predictors was stable until the end of Year 1, when slight cross-linguistic divergences, as well as strong autoregressive effects, were attested; and (3) both languages drew on the same set of predictors during the foundational stage of literacy development. In addition, poor explanatory models were obtained for text-level performance scores. Theoretical implications are discussed.

6.1 Introduction

Being able to express ideas through writing is one of the most important achievements of elementary education. Early on in children's academic life writing becomes the main means of interaction at school. Students are expected to obtain information from written material and to show their knowledge about different topics through composing written texts. Children with writing difficulties are thus highly likely to experience academic failure throughout the curriculum. In this sense, early detection and remediation of writing difficulties is key, since children identified as poor writers in the elementary school years only seem to worsen their situation relative to other peers in later years (Hooper, Roberts, Nelson, Zeisel, & Fannin, 2010; Juel, 1988). However, very few studies have attempted to identify the cognitive precursors of writing skills. A first aim of this study was thus to examine the cognitive underpinnings of early writing development.

It is also important to investigate the longitudinal predictors of early text composition from a cross-linguistic perspective. This is to test the putative differential effects of the constraints of inconsistent orthographies, as opposed to those of more consistent ones. In turn, the varying demands of a major component of writing, namely spelling, may arguably also affect the type and relative weight of the cognitive skills supporting early writing development. A second aim of this study was to test whether the predictors of writing during the foundational stages of this skill are the same in English, an inconsistent orthography, and Spanish, a consistent one.

6.1.1 Previous Research on the Cognitive Predictors of Writing

6.1.1.1 Spelling. The vast majority of longitudinal studies of early writing development were concerned with investigating the predictive power of transcription skills (handwriting and spelling). As seen in Study 3 above (Section 5), there is evidence that spelling explains variance in writing, although its explanatory power varies as a function of (1) the type of writing outcome being explored: Spelling explained more variance in text spelling and other word-level writing features (e.g., segmentation and use of case) than in text length, while it had a null or unreliable role in predicting various text-level measures; (2) the point in development: Its effect was more robust during Year 1 than Year 2; (3)

orthographic consistency: Spelling explained more variance, in general, in English than in Spanish. These results seem consistent with theories that postulate that transcription skills constrain the early development of writing (e.g., Berninger & Swanson, 1994; Berninger et al., 1991; Berninger, 1999) as does orthographic consistency, since learners of more consistent orthographies are expected to be influenced by such constraints for a shorter period of time, given that they automatize transcription processes earlier (e.g., Babayiğit & Stainthorp, 2010; Mäki et al., 2001). In this study, a measure of single-word spelling skill was used as a predictor of writing development.

6.1.1.2 Reading. Reading and writing influence each other throughout development and are tightly related constructs (e.g., Mehta et al., 2005; Shanahan, 1984; Shanahan & Lomax, 1986). It is not entirely clear what the exact nature of their relationship across the school years is, in the sense of whether reading influences writing (e.g., Abbott & Berninger, 1993; Juel, 1988) or whether the reverse is true. In addition, it could be that the relationship is reciprocal. Finally, it is also possible that the way that reading and writing relate to each other varies over time (Lerkkanen et al., 2004). Abbott et al., (2010) ran a cross-lagged model to explore the longitudinal relationships between reading, reading comprehension, spelling, and written composition, covering Grades 1st to 7th. Grade 2 word reading significantly predicted Grade 3 writing, and reading comprehension significantly predicted writing in the next adjacent year from 3rd to 6th grade. In contrast, text writing never supported word reading and only predicted reading comprehension in the later primary grades. Their results support the view that, at least in the initial phases of literacy development, reading skills set the base for later writing skills rather than vice versa, although the authors note a tendency to reciprocity towards the end of the study period. Lerkkanen et al. (2004) found that reading at the outset of Year 1 predicted early writing performance in Finnish, measured as the number of correctly spelled words, only a few months after the initial reading assessment. However, later on in the same school year, it appeared that writing supported reading achievements more than the reverse, although a significant path was also found from initial reading to a later assessment of writing (carried out in March of first grade). These results were taken as evidence of the fluctuating and interactive relationship between reading and writing. The fluctuating pattern of influences between reading comprehension and writing echoes previous elaborations for the single-word domain (e.g., Frith, 1985). It should be noted, however, that Lerkkanen et al.'s (2004) study focused on reading accuracy for the initial assessment, and more on comprehension for the later assessments of reading. This is due to the fact that accuracy levels in reading reach ceiling very early on in the highly consistent orthography of Finnish. Therefore, the changes in the writing-reading association may be influenced by variations in the specific skills involved in each assessment.

In sum, reading has been found to contribute to explaining variance in writing achievements over the primary school years, in both consistent and inconsistent orthographies. The exact nature of the relationship is still a topic open for discussion, which is far from the scope of the present study. However, given the tight relationship between the two skills and the different ways in which reading may support gains in writing performance, measures of reading were included as predictors of writing in this study.

6.1.1.3 Oral language. In literate communities, the language learning taking place after the age of 5 is difficult to dissociate from written language (Berman, 2008; Ravid & Tolchinsky, 2002; Tolchinsky, 2004). Oral language abilities and written expression are tightly linked. Several aspects of written language involve acquiring knowledge of the way in which writing represents oral language and, in turn, exposure to print shapes and reshapes linguistic knowledge (e.g., Olson, 1994). It has also been proposed that oral language competence may indirectly influence text generation by assisting working memory processes (McCutchen, Covill, Hoyne, & Mildes, 1994). Not surprisingly, oral language competence has been found to be related to written composition (e.g., Cragg & Nation, 2006; Griffin, Hemphill, Camp, & Wolf, 2004). Oral language skills made a significant contribution in explaining third to fifth grade writing quality, as well as rate of writing development in a study of preschool predictors of writing (Hooper et al., 2010). Berninger and Abbott (2010) conducted a large-scale longitudinal study from grades 1-7, where an oral expression factor predicted scores of a written expression factor in 3rd and 7th grade, but not in grade 1.

Writing research with language impaired (LI) populations provides an additional source of evidence in favour of the central role of language competence for literacy and, particularly, writing development. LI children experience sustained written expression difficulties as a result of their morphosyntactic deficits (Bishop & Clarkson, 2003; Bishop & Snowling, 2004; Mackie et al., 2012). Their problems with writing are evident during the primary grades and may last up to adolescence (Dockrell et al., 2007). Compared to age-matched peers, the texts produced by LI children tend to be shorter, contain a higher proportion of spelling mistakes (usually involving function words), show lower levels of lexical diversity, and include a reduced range of grammatical structures, which tend to be simple (e.g., Mackie & Dockrell, 2004). Also, their overall text quality is lower (Connelly et al., 2012; Fey et al., 2004; Puranik & Lonigan, 2012). Importantly, their writing difficulties are not restricted to morphosyntax, but they also involve phonological competence (Bishop & Snowling, 2004; Mackie et al., 2012;).To sum up, research on LI populations provides robust evidence that writing development is at least partly supported by oral language skills.

6.1.1.3.1 Phonological awareness. The cognitive ability to manipulate single phonemes—that is, phoneme awareness—, is at the heart of basic spelling and reading skills, since it enables the acquisition of the alphabetic principle (Byrne, 1998). Recent findings have indicated that it constitutes a foundational skill of spelling and reading development across alphabetic scripts (e.g., Caravolas et al., 2001; 2012; Vaessen, et al., 2010; Ziegler, et al, 2010). However, phonological awareness was the only oral language measure which did not explain unique variance in a study measuring several oral language indicators (Hooper et al., 2010). This was explained by the fact that there were multicollinearity issues with letter- and word- identification measures. Phonological awareness has been claimed to be associated with spelling also in consistent orthographies. Borzone and Signorini (1998) compared the spelling skills of Spanish-speaking kindergarten children during a text production task. Children with higher levels of phonological awareness and knowledge of letter names had an analytic spelling strategy (i.e., sounding out letters; Adams, 1990), which usually allowed them to produce conventional spellings of words in their texts, since Spanish is very consistent in its phonographic mappings. In contrast, children with lower levels of phonological awareness typically produced pseudo-letter strings and very rarely produced accurate spellings. The robust empirical evidence indicating the central role of phonological awareness in both consistent and inconsistent orthographies motivated the inclusion of this type of measure as predictors in this study. In addition, a measure of letter knowledge was also included in the models since—although it is very closely related to phoneme awareness—it has been found to contribute to explaining unique variance in later gains in spelling and reading skill (e.g., Caravolas et al., 2012) and to be of particular importance in languages with consistent orthographies (Caravolas et al., 2013).

6.1.1.3.2 Morphosyntactic skills. Knowledge of phonographic correspondences is often not sufficient for accurate word spelling, particularly in inconsistent orthographies like English. In such cases, morphological and syntactic awareness become necessary as well (Carslile, 1996). A good number of studies have indicated that children resort to morphological knowledge when spelling in inconsistent orthographies (e.g., see Pacton & Deacon , 2008 for a review), but also in very consistent ones, like Spanish (Defior et al., 2008). Use of morphological information seems to be moderated, however, by morphological typology, with morphologically richer languages more often resorting to this type of knowledge for aid in spelling inconsistent words. Gillis and Ravid (2000, 2006) compared Hebrew- and Dutch-speaking children's use of morphological and morphophonological information to resolve phonographic ambiguities. The Hebrew children showed higher levels of morphological awareness evident in their more frequent use of morphological information to decide between alternative spellings (where the correct form involved access to morpho-graphemic rules). Crucially, Hebrew is a richly inflected language in contrast with the rather poor morphology of Dutch. Notably, though, the study did not make reference to the degrees of orthographic consistency in each language, so it is hard to anticipate in what ways orthographic consistency and morphological typology relate to each other at the early stages of literacy development.

Morphological and, more generally speaking, morphosyntactic skills, can be argued to be essential for text composing over and above their usefulness for spelling irregular words. Appropriate management of co-reference within the text, and knowledge of the available structures for expressing the same propositional content, as well as knowledge of derived words all contribute to enhancing precision in writing and ensure effective communication. Very few studies, however, have studied the relationship between levels of morphosyntactic awareness and writing outcomes in the early stages of literacy development. There is evidence that grammatical competence predicts text-quality ratings (Olinghouse, 2008). The author found that a grammaticality judgement task administered to American 3rd graders predicted ratings of text quality. However, Berninger et al. (1992) found that a measure of sentence syntax (repeating sentences of increasing syntactic complexity) did not contribute to explaining variance in the number of words or clauses produced in two writing tasks, elicited by a narrative and an expository frame. Similarly, some studies with LI children, who are characterized by poor morphosyntactic skills (Leonard, 1998), also failed to find a relationship between measures of grammar and children's written outcomes (e.g., Bishop & Clarkson, 2003; Dockrell et al., 2007; Mackie & Dockrell, 2004). Inconclusive results about the role of morphology and syntax in writing development are not restricted to English. For example, Babayiğit and Stainthorp (2010) found a measure of Year 1 morphological awareness to be a reliable predictor of Year 2 written *content* scores in Turkish, an agglutinative language (and, therefore, very rich in inflectional morphology). Surprisingly, morphological awareness was unrelated to ratings of text *structure*, which was estimated by the use of connectors and complex syntax. In Spanish, Silva, Sánchez, and Borzone (2010) found no differences as a function of grade level in the syntactic complexity—measured as the number of subordinated clauses per T-unit —of written narrations by 1st and 2nd graders. However, syntactic complexity showed medium-low but significant correlations with a word spelling task (r = .33 and .27, for 1st and 2nd grade, respectively) and with text length at the end of second grade (r = .28).

In sum, measures of morphology and syntactic competence may have an impact at both the word and text level of writing. However, the specific contribution of morphosyntactic skills may differ as a function of linguistic typology, with morphologically richer languages making more extensive use of morphological information. In addition, the influence of morphosyntactic skills on writing may not operate at all grade levels. Since research to date is unclear about the specific role of morphosyntactic skills in early writing, this study will include measures of both morphological and syntactic awareness.

6.1.1.3.3 Vocabulary. Vocabulary knowledge is decisive to ensure expressive and precise production of content and ideas. It is thought to support efficient retrieval of vocabulary during text production (Mackie et al., 2012) and, therefore, may help with the text generation process (Berninger et al., 1992). Previous studies have found support for the predictive power of vocabulary in explaining variance in early writing (e.g., Berninger et al., 1992; Berninger & Fuller, 1992; Connelly et al., 2012; Dockrell et al., 2007; Dunsmuir & Blatchford, 2004; Olinghouse & Leaird, 2009). The longitudinal study by Babayiğit and Stainthorp (2010) provides an excellent test of the extent to which vocabulary knowledge is related to writing outcomes. In assessing writing, the authors distinguished between *content* writing quality (which resulted from separate ratings for "general content" and "choice of vocabulary") and *structure* writing quality (which resulted from separate ratings for the use of connectors and subordinated clauses). They found that Year 1 vocabulary significantly predicted content ratings in texts produced by Turkish children in Year 2. Because part of the rubric for assessing content quality included ratings of vocabulary use, they removed the weight of vocabulary in the content quality assessment, but the significant association with vocabulary knowledge remained. Similarly, Dockrell et al., (2007) found support for a *semantic* factor (comprising measures of ideas, organization and coherence, and vocabulary) in LI children's written productions (as separate from a *rule* factor, including measures of sentence structure, grammar, and capitalization, p. 156), of which vocabulary knowledge was a significant predictor. In contrast, in Puranik and AlOtaiba's (2012) study with children at the end of kindergarten year, vocabulary was not found to explain unique variance in writing outcomes. Two aspects of this study must be noted, though. First, the study was cross-sectional, while the Dockrell et al.'s (2007) and Babayiğit and Stainthorp's (2010) studies were longitudinal. Second, the only dependent measure in Puranik and AlOtaiba's (2012) study was number of words produced (which showed an almost perfect correlation, r = .94, with number of ideas), whereas the study in Turkish included several writing measures, thus allowing the identification of those aspects of text composition that are supported by vocabulary.

6.1.1.4 Memory. The high cognitive demands of text composition led models of writing to include working memory and short-term memory components (e.g., Berninger & Swanson, 1994; 1996; Berninger & Winn, 2006; Kellogg, 1996; McCutchen, 1996). Efficient working- and short-term memory capacity in novice writers should be useful to compensate for the interruptions to text generation produced by transcription processes (Graham, 1990; Scardamalia, et al., 1982); for example, to sound out a difficult word without forgetting the plan for expressing content. Indeed, short-term memory has been reported to be related to composition in English, although its effect is stronger on transcription, rather than on text-level writing features (Puranik, 2006). On the other hand, working memory constraints have been found not to be operational until the intermediate grades (e.g., Berninger & Swanson, 1994; Berninger & Winn, 2006). For this reason, only a measure of verbal short-term memory was included in this study.

6.1.1.5 Naming speed for visual-verbal associations. Even if working-memory constraints are not relevant at these early stages of writing development, it could be argued that fast retrieval of verbal and orthographic information about words from both short- and long-term memory may lead to improvements of text spelling and in the quality and quantity of text produced. The task of Rapid Automatized Naming (RAN)—a task that measures the speed at which children name a limited set of items (colours, objects, letters, or digits) displayed on a page— warrants consideration as a cognitive component skill of writing. It has been argued that RAN taps onto general processing speed mechanisms (Kail, Hall, & Caskey, 1999), while others argue that it is sensitive specifically to speed of phonological retrieval (e.g., Wagner & Torgesen, 1987; see also Lervåg & Hulme, 2009). There is a wealth of empirical evidence of the validity of RAN as a predictor, in consistent and inconsistent orthographies, of later reading skills (e.g., Caravolas et al., 2012; 2013; Vaessen et al., 2009), as well as of spelling skills (Caravolas et al., 2012). The types of processing mechanisms that RAN may be argued to be tapping onto, particularly its emphasis on speed, make it a predictor of interest not only of text spelling skill, but also of writing fluency. In contrast to the argument with regards to short-term and working memory, RAN speed may be an index of individual variation in verbal-visual mapping fluency that estimates the degree of automatization of the transcription process. For these reasons, measures of non-alphanumeric RAN (considered to be less biased towards predicting literacy outcomes; Lervåg & Hulme, 2009) were included in this study.

6.1.2 The Present Study

This study is unique in several ways. Firstly, it examines the role of cognitive predictors, which are followed over the first two years of formal literacy instruction, in the development of early literacy outcomes across languages. Secondly, it tests whether orthographic consistency shapes the underlying cognitive skills that support writing development. Finally, it explores the predictors of the development of distinct writing outcomes. These outcomes were *text spelling* accuracy; *text length* or writing productivity; a *vocabulary* regression-based factor score, obtained from an exploratory factor analysis (Chapter 4); and an analytic *text-level* score comprising punctuation, connectors, syntactic complexity, and vocabulary choices.

Investigating the cognitive profiles underlying different writing outcomes should help determine the specific needs of children who are in their first steps of learning to write and, particularly, to establish research-based intervention programs for children at risk of literacy difficulties. In this sense, it has been shown that children may experience difficulties with some aspects of the writing process (e.g., transcription) but not others (e.g., idea generation; Berninger et al., 1992).

Text spelling was chosen because it is a fundamental indicator of children's word-level/transcription skills (Puranik et al., 2008). Text length or productivity in number of words has been used in writing development research as an indicator of children's text generation capacities (e.g., Puranik & AlOtaiba, 2012) and it has been shown that it correlates with global assessments of micro- and macro-structural quality (Berman & Nir, 2009; Lee, Gentile, & Kantor, 2010). Vocabulary is often included in studies assessing the content of children's written texts (e.g., Babayiğit & Stainthorp, 2010) and it is usually part of the rubric used to assess writing quality. Finally, analytic scales of writing outcomes are used frequently in writing research, as a means to ensure that—contrary to holistic assessments of writing quality—key aspects of the written product are weighed in the resulting scores (e.g., Lee et al., 2010). Analytic scales typically evaluate different aspects or dimensions of the texts and these are summed up to obtain a final, aggregated score of written quality. Here, children's scores for individual markers of text-level performance on text structure, syntactic complexity, and vocabulary were aggregated and constituted the text-level performance outcome measure at each time point and language group.

In order to fulfil the main objective of the study, analyses were designed to answer the following research questions: (1) Which are the most relevant predictors of each writing outcome?; (2) Does the relative weight of the different predictors change over time?; and (3) Is early writing development driven by the same set of cognitive predictors in English and in Spanish? The hypotheses regarding these questions are detailed for each dimension of writing separately.

With respect to the first research question, it was expected that *text spelling* skills would be predicted by word spelling, phonological awareness, RAN, and knowledge of letters, in line with a previous study on word spelling development across languages (Caravolas et al., 2012). Notwithstanding the differences between word and text spelling mentioned elsewhere (e.g., Chapter 5; Abbott & Berninger, 1993; Alamargot & Chanquoy, 2001), a high proportion of the variance in text spelling should be explained by the same subset of predictors that account for word-spelling performance.

The number of words children are able to produce during a limited amount of time, that is, the *text length* measure, was hypothesized to be predicted by word spelling skills, in line with findings from Study 3 (Section 5), as well as by other studies in English (e.g., Berninger, et al., 1992) and in other more consistent orthographies (e.g., Babayiğit & Stainthorp, 2010; Lerkkanen et al., 2004). In addition, the naming speed for visually presented objects that RAN taps into may be argued to help with retrieval of the content being generated. Similarly, efficient verbal short-term memory may aid in retaining writing plans long enough so that they are not disrupted by the interruptions of transcription processes. Moreover, vocabulary knowledge may also influence text generation by facilitating the translation of ideas into words. Finally, familiarity with letter names and sounds may contribute, over and above spelling, to transposing linguistic formulation into written form. Thus, RAN, verbal span, vocabulary, and letter knowledge were all included as predictors of text length.

It was hypothesized that preschool levels of vocabulary would contribute to explaining variance in *vocabulary* factor scores, together with RAN and verbal short-term memory span, which may arguably be essential for fast retrieval of vocabulary tokens. Morphosyntactic awareness should also be a relevant predictor, especially because the morphology measure assessed derivational morphology, which may indirectly tap into vocabulary knowledge. Better readers are likely to have more exposure to print (Juel, 1988), which in turn works as another means of vocabulary learning (Cain & Oakhill, 2011). For this reason, reading skills were included as predictors of vocabulary scores. Finally, and although spelling skills did not appear to have an impact on single text-level indicators (Chapter 5), it is possible that the combination of measures—as is the case with the vocabulary factor—increase statistical power to detect an effect. Certainly, besides the central role that current models of writing assign to spelling skills (e.g., Berninger & Swanson, 1994; Juel et al., 1986), it has moreover been argued that children's spelling abilities are a facilitator in vocabulary learning (Ehri & Rosenthal, 2010). Therefore, vocabulary knowledge, word spelling, RAN, verbal short-term memory, and morphosyntactic awareness were the tested predictors of children's vocabulary factor scores in English and Spanish. Note that only the cognitive profile of Time 5 (mid-Year 2) levels of vocabulary scores was examined, and not scores at previous time points, given that the vocabulary factor emerged at Time 4, for English, and Time 5, for Spanish. Therefore, it was only at Time 5 that a cross-linguistic comparison could be performed.

As for *text-level* performance, spelling skills were hypothesized to contribute to explaining variance given that—in spite of the inconclusive results of Study 3—it was expected that, similarly to the vocabulary factor scores, pooling various text-level scores together would result in the gain of statistical power, so that weak associations might be boosted. In addition, knowledge of vocabulary should be associated to later achievements in text-level performance by providing the child with more tools to show precision and richness in expressing events. Also, morphosyntactic skills should bring about more complex syntactic structures during text composition, as well as facilitating the text generation process. Finally, better reading skills usually indicate more experience with and exposure to written language, and they may also impact indirectly on text generation by boosting vocabulary knowledge (Juel, 1988).

As for the second research question regarding the change in the relative weight of predictors over time, it was hypothesized that the relative weight of the different predictors should be stable until the end of Year 1. However, as each type of writing dimension stabilizes, autoregressive effects should become stronger and the unique contribution of the different predictors should decline. This hypothesis is consistent with recent findings for the growth of word reading skills (Caravolas et al., 2013), where after the end of Year 1 the powerful autoregressive effects overshadowed the relative predictive role of other skills.

The last research question is concerned with the expected differences as a function of language or orthographic consistency. In this sense, it was predicted that the highly consistent Spanish orthography would lead to a shorter period of dependency upon word spelling skills relative to English speaking children, as the findings from Study 3 (Chapter 5) seemed to suggest. Based on the same findings, however, the effect of orthographic consistency on the other dimensions of writing (text length, vocabulary scores, and text-level skills) was expected to be limited.

6.2 Method

6.2.1 Participants

Data from the entire samples were included in this study because the type of statistical analyses used permitted multiple imputation procedures to manage missing data. Preschool predictors were obtained during mid-Reception Year (Time 1). The English sample thus consisted of 188 participants (98 boys; mean age at Time 1 = 60.27 months) and the Spanish sample consisted of 190 participants (104 boys; mean age at Time 1 = 66.72 months). Full demographic information and details of recruitment procedures can be found in Section 3.2.

6.2.2 Tasks and Materials

6.2.2.1 Letter spelling (Time 1, Time 3). This test assessed children's early spelling ability by evaluating their accuracy in writing letters to dictation.

Fifteen letters of the alphabet were selected such that five of them represented vowel sounds, another five represented consonants with consistent phoneme-to-grapheme mappings, and another five represented consonants with less consistent mappings. The administrator pronounced the sound (not the name) of a letter and asked the child to write the letter that s/he thought represented that sound in writing. The test was discontinued only if a child refused to write six consecutive letters. Correct letters were awarded 2 points, reversals (e.g., <d> for <b>) were awarded 1 point. The task was not administered after Time 3 because children in both language groups had reached mastery level (>80% accuracy). Chronbach's α was .87 and .90 in English and Spanish, respectively.

6.2.2.2 Word spelling (Time 1, Time 3, Time 4). This task assessed children's early spelling ability by evaluating their accuracy to write words to dictation. Children wrote high-frequency words that were orally presented by the administrator. Words (except *mom, sun,* or *bunny*) were administered together with a carrier sentence to facilitate word identification on the part of the child. The administrator said, for example, "write *in*, as in *sleeping* in *bed".* At Time 1, children were asked to write their own name and then another 9 words (7 for Spanish) were administered. At Time 3 and Time 4, 26 items, graded in difficulty, were added to the English instrument (35 additional items were added in Spanish). Administration at Time 1 was individual, and in small groups of 4 to 5 children in subsequent time points. Children were praised for their efforts and encouraged to write as they could. No feedback was given during the entire task, which was never discontinued. Each correctly spelled word received a score of 1. Chronbach's α was .81 and .83 in English and Spanish, respectively.

6.2.2.3 Block design (Time 1). The Block-design subtest of the WPPSI-III-UK (Wechsler, 2003) and the equivalent subtest of its Spanish version (Wechsler, 2001), were administered to the British and Spanish participants, respectively, in order to estimate non-verbal IQ. Children had to use blocks of different colours to imitate a model produced by the experimenter or one displayed in a picture. Standardized scoring procedures were used. The English (UK) version of this subtest has a reported reliability of .84, and the Spanish instrument reports .85.

6.2.2.4 Phoneme isolation (Time 1, Time 3, Time 4). This was one of two tasks estimating children's phonemic awareness, by asking children to manipulate single sound segments of pseudowords. The task (based on Caravolas & Bruck, 1993; Caravolas et al., 2001) consisted of isolating one phoneme from either the beginning or end of a pseudoword. In the *initial phoneme isolation* block, there were 16 items divided into two sub-blocks: 8 CVC items and 8 CCVC items. In the *final phoneme isolation* block there were another 16 items: 8 CVC and 8 CVCC[11] items. The experimenter explained that they were going to play a game with "pretend words", where the child will have to say which the first sound in a word was. There were two demonstration items, the first of which was a real word (e.g., *cat*) and the other one was a pretend word (e.g., *sot*). The experimenter said, "Here's the funny word /sot/.The first sound of /sot/ is /s/; /s/ is the sound it makes at the very beginning. Do you hear that? /s:ot/, /s:/. Can you say /s:/?" Next, four practice items were administered—two of each syllable type. Children were required to repeat the whole word first and to pronounce the isolated phoneme immediately afterwards, in order to be able to correctly assess whether the isolation had been accurate. The experimenter gave corrective feedback during the practice trials, which could be repeated if the child failed to understand the procedure of the task. After the practice trials, the CVC and CCVC/CVCC items were administered in that order, without giving feedback. A sub-block was discontinued if the child produced four consecutive wrong answers; this did not affect the administration of the following block. Chronbach's α was .94 and .97 for English and Spanish, respectively.

6.2.2.5 Phoneme blending (Time 1, Time 3, Time 4). The second task used for estimating children's phonemic awareness consisted in identifying a word by putting together a series of isolated phonemes. It was important to assess phoneme awareness using different types of tasks—phonological analysis (phoneme isolation) and phonological synthesis (blending)—given that it has been argued that they relate differently to literacy development (Wagner, Torgesen, & Rashotte, 1994). Children were required to listen to a series of recorded sound segments and to pronounce the resulting word by blending the sounds (e.g., /p/, /əʊ/, /n/, /i/ →*pony*). At Time 1 there was a total of 10 items and an additional 14

[11] Since Spanish does not allow complex word endings in native word stock, the "complex syllable structure" subset contained CCVC pseudowords.

items were included in subsequent time points. Items increased in length and in syllabic complexity. The task was presented as a game in which the child had to "guess a secret word" by putting together the sounds in that word. There were two demonstration items with picture support: the child was shown the picture of, for example, a sheep, and was asked "What's this?" The experimenter used the child's response (*sheep*) to illustrate how the word is made up of three sound segments, /ʃ/ /i:/ /p/.Next, there were two practice trials without picture support, which resembled the actual nature of the task. The experimenter said, for example, "I'll say the sounds of a secret word and you tell me what you think the word is. Listen carefully, /b/ /u:/ /t/ (*boot*). What do you think the word is?" Corrective feedback was given during the practice trials, which could be repeated if the child did not understand the procedure. To avoid different rates and delivery styles of the tests items, they were recorded and presented over loud speakers, so that participants were exposed to the same set of stimuli. The approximate duration of each sound segment ranged between 300-500 ms and a 500 ms pause was inserted to separate the sounds. The test was discontinued after six responses showing no overlap with any of the sounds in the target word. Correct answers received a score of 1. Chronbach's α was .76 and .87 for English and Spanish, respectively.

6.2.2.6 Morphological awareness (Time 1, Time 3). This task was designed to estimate children's level of inflectional and derivational morphological awareness skills. Children were asked to either complete a sentence or answer a question which required the morphological transformation of a previously presented word. Sentence-completion tasks have along-standing tradition in assessing knowledge of morphology (e.g., Berko, 1958). Items were constructed so that in each language relevant and age-appropriate structures were tapped into, but in keeping with the general design of the task (after Muter, Hulme, Snowling, & Stevenson, 2004). Items (*N*=28) were grouped into blocks, according to whether they tapped into noun-phrase, verb-phrase, or derivational morphology. For noun-phrase morphology, items were delivered with picture support. Prior to beginning administering the test items of each block (e.g., noun phrase morphology), two practice trials were administered. The goal of these trials was to familiarize children with the nature of the task, not to give corrective feedback in terms of the appropriate morpheme that should be used. Test trials for a particular block were administered without giving any kind of feedback. In the English instrument, the noun-phrase block tested children's awareness of plural formation (block 1.a) and genitive case formation (block 1.b). In Spanish, this block tested noun-adjective gender agreement. The verb-phrase morphology block tested children's awareness of present-tense, 3rd person singular formation, and past tense (regular and irregular) formation in English; in Spanish, several verbal inflections were tested. In the derivational morphology block, English children had to form agent derivatives out of nouns and verbs, while Spanish children had to form a variety of derivatives. Each correct answer received a score of 1, which were summed up, giving a range of 0-28 in both languages. Chronbach's α was .80 and .74 for English and Spanish, respectively.

6.2.2.7 Syntactic awareness (Time 1). At Time 1 an adaptation of the Test for Reception of Grammar, version 2[12] (*TROG-2*; Bishop, 2003) was administered to obtain an estimate of children's syntactic awareness skills. Children listened to a sentence containing a target grammatical structure and had to decide which of four pictures best fitted the sentence they heard. Items in the TROG-2 are grouped into blocks according to the morphosyntactic structure they test. Four blocks of four items each (total = 16 items) were selected because they tap into syntactic structures which are comparable in nature and difficulty across the two languages languages: (1) subject relatives; (2) passive voice constructions; (3) object relatives; and (4) centre-embedded clauses. Children looked at four pictures while listening to an orally presented sentence, e.g., *The man that is eating looks at the cat.* Two practice trials were administered prior to the test trials. Children then pointed to the picture which they considered to be the best fit to the sentence. Each correctly matched picture is awarded 1 point. The final score ranged between 0-16. Chronbach's α was low: .58 and .61 for English and Spanish, respectively.

6.2.2.8 Sentence repetition (Time 3). At Time 3 some of the sentence items of the TROG-2 and others of similar structure were used to build a sentence repetition test. Sentence-repetition tasks are a simple and useful means to assess children's receptive syntax, given that children need to parse the syntactic structure upon repeating it (Lust et al., 1996). To avoid biases due to differences in

[12] The TROG test was adapted to Spanish syntax, but the same target constructions were tested.

the administration of the items, sentences were recorded. The Spanish instrument was composed of 16 items, while the English instrument was composed of 20 items. Sentences increased in syntactic complexity, but the task was never discontinued. Exact repetitions were afforded a score of 1, while substitutions and omissions resulted in a score of 0 for any given item. In English, some differences between the target sentence and children's repetitions that were due to dialectal variation were not penalized. These mostly involved substituting the subordinating conjunction *that* with *what*. For example, for the item *The duck that is running looks at the girl* children often said *The duck* <u>*what*</u> *is running looks at the girl.* These types of "errors" were considered to be direct evidence of the children parsing the targeted syntactic structures. Chronbach's α for this task was .93 and .82, for English and Spanish, respectively.

6.2.2.9 Vocabulary (Time 1). The Vocabulary subtest of the WPPSI-III-UK (Wechsler, 2003) and the equivalent subtest of its Spanish version (Wechsler, 2001) were administered to the British and Spanish participants, respectively, in order to estimate vocabulary knowledge. Children had to provide definitions of a series of nouns, verbs, and adjectives. The test requires not just the identification of a word with its reference (as would be the case with a receptive vocabulary task), but it demands the child to show how much s/he knows about any given word. Questions had the structure *What is X?* or *What does X mean?,* and a score of 0-2 is afforded to each item, depending on the grasp, on the part of the child, of the concept in question. Therefore, the test might be argued to tap into both breath and depth of vocabulary. Standardized scores were used. The English (UK) version of this subtest has a reported reliability of 95, and the Spanish instrument reports .81. To ensure that administrators were scoring using the same criteria, answers to the task were transcribed and re-scored by two raters in each language group. Agreement between raters exceeded 80% and cases of discrepancy were resolved through discussion until 100% agreement was reached.

6.2.2.10 One-minute reading (Time 1, Time 3, Time 4). This task was used to obtain a timed measure of word decoding skills. Children were asked to read as many words as possible in one minute. Words increased in complexity starting with short, high-frequency items with simple syllabic structure and high phoneme-grapheme consistency. The English instrument consisted of 144 words, and the Spanish instrument consisted of 140 words. Children were instructed to read as quickly as possible each word in order, following with their finger. They were encouraged to try all words, but to go the next one if they got stuck. Each correctly identified word received a score of 1. Test-retest reliability was .98 for English, and .87, for Spanish.

6.2.2.11 Picture-word matching (Time 1, Time 3, Time 4). This timed task measured silent word-identification skills. Silent reading has been argued to constitute a related, though distinct, form of reading skill (e.g., Kim, Wagner, & Foster, 2011); thus, it was assessed in order to have a more complete picture of children's reading skills. Children had to match a picture to one of four printed words next to it. The target words were cognates matched cross-linguistically for number of letters and syllables, syllable complexity, and frequency of use (see Caravolas et al., 2012, for a sample of the task materials). The three distracters for every target item were created so that there always were (1) a phonologically close word (e.g., *rose-roads*; *lemon-linen*); (2) a semantically related word (e.g., *rose-daisy*; *chocolate-candy*); and (3) an unrelated word (e.g., *rose-cage*; *robot-winter*). Children were first administered three practice trials to familiarize them with the characteristics of the task. They were told to choose the word that best matched the picture and that they should not omit any items. The test phase lasted three minutes and children were required to stop once the time had elapsed. Each correct answer received a score of 1; incorrect and unanswered items received a score of 0. Test-retest reliability was .93 and .81, for English and Spanish, respectively.

6.2.2.12 RAN pictures (Time 1, Time 3, Time 4). A task of rapid automatized naming (RAN) with picture items was administered. Children had to name five pictures of familiar objects (dog, eye, key, lion, table), which appeared a total of eight times each and were displayed in a fixed random order in a blank page. Children were first presented the five objects and asked to name each. If the child failed to accurately name one of the items, the administrator provided him with corrective feedback. The objective of these practice trials (max. 3) was to get the child to name all five items fluently and without assistance from the administrator. Next, there were two test trials, which were timed from the moment in which the child named the first object until naming the last one. Children were told to name the pictures as quickly as they could, without making errors of omitting items. The scoring of the task was the average time (in seconds) of trials 1 and 2; errors were also tallied. Task reliability was evaluated by computing the correlation between trials 1 and 2, which was .84 and .79, for English and Spanish, respectively.

6.2.2.13 RAN colours (Time 1, Time 3, Time 4). This task was identical in nature and procedure to the RAN objects task, except that the items were five different colours (red, brown, green, black, blue and their direct translations to Spanish). Task reliability was evaluated by computing the correlation between trials 1 and 2, which was .84 and .82, for English and Spanish, respectively.

6.2.2.14 Letter knowledge (Time 1, Time 3, Time 4). The 26 letters in the English alphabet, and the 27 letters (the same 26 as English, plus letter *ñ*) as well as two digraphs (*ch* and *ll*) in the Spanish alphabet, were presented individually on 7 x 7 cm cards (Arial, 72 pt.). The task aimed to estimate children's knowledge of letters names and sounds, assessing both lower- and uppercase letters. Children were asked first to say the sound and then the name of each letter. The child's initials were presented first, followed by the rest of the letters, which were administered in fixed random order, according to the order in which they are normally learned and taught in each language. The entire alphabet was administered in this manner, first with lower case letters and then with upper case letters, given that children are taught lowercase letters first (in Spanish the order was reversed, for the same reason). Print letters were used throughout. The task was only discontinued if a child failed to give *both* the correct sound and name of 5 consecutive letters. If the discontinue criteria applied for the first set of letters (e.g., lowercase in English) the second set was nevertheless administered. One point was given for each correct response. Chronbach's α for this task was .98 and .99 in English and Spanish, respectively.

6.2.2.15 Word span (Time 1, Time 4). This task measured verbal short-term memory skills (Hulme, Maughan, & Brown, 1991; McDougal, Hulme, Ellis, & Monk, 1994). Children were asked to repeat lists of monosyllabic words accurately and in the same order as they were presented. The task started with two-word lists and progressed up to eight-word lists. There were four lists for any given list length, making a total of 56 trials. Children were instructed to repeat the words they heard in the exact same order and without omitting any words. No practice trials were administered. Within each trial, words were delivered at a rate of, approximately, one word per second. The task was discontinued if the child incorrectly repeated (omission or wrong order of words) three lists within a given list length. Each correct trial received a score of 0.25 points (i.e., all four lists of one list length = 1 point), making the range of scores 0-8 (i.e., the same scoring procedure followed by Hulme et al., 1991). Split-half reliability was lower than desirable: .73 and .55, for English and Spanish, respectively.

6.2.2.16 Text writing (Time 3, Time 4, Time 5). The same experimental writing task as previously described was used (see details in Section 3.3). As mentioned, four writing outcomes were studied: text spelling accuracy (percentage of words spelled conventionally), text length (number of words), vocabulary factor scores (only at Time 5), and text-level performance. The latter consisted in an aggregated score of some of the text-level measures used in previous studies in the present thesis. In light of the shortcomings of using perhaps extremely fine-grained measures in Study 3 above (Section 5), an aggregated score of text-level writing performance was created in order to pool performance scores for each dimension of writing. The variables of percentage of connectors, use of punctuation, words-per-clause, percentage of subordination, lexical density, and percentage of adjectives and adverbs were standardized (z-scores) for each language and time point. The individual *z*-scores were then summed up to create the aggregate score. The resulting variables showed excellent distribution properties in both language groups.

6.2.3 Procedure

Table 6.1 shows the list of measures obtained at each time point. At Time 1 the tasks were distributed into three sessions, all of them individual. In subsequent time points, a few tasks were administered in group sessions, including the text writing task. Collection, transcription, scoring, and coding of the data were carried out by the author as well as by another graduate student (based in Spain) and trained research assistants. The writing task was administered at mid-Year 1 (Time 3), end of Year 1 (Time 4), and mid-Year 2 (Time 5) to groups of about 5 children.

Table 6.1

Summary of measures obtained at each time point

Time point	Time 1	Time 3	Time 4	Time 5
	Reception (Feb-March)	Year 1 (Nov-Dec)	Year 1 (June)	Year 2 (Nov-Dec)
Measures	Block design	Phoneme isolation	Phoneme isolation	Text spelling
	Vocabulary	Phoneme blending	Phoneme blending	Text length
	Phoneme isolation	Letter knowledge	Letter knowledge	Vocabulary factor score
	Phoneme blending	Letter spelling	Word spelling	Text-level score
	Letter knowledge	Word spelling	RAN pictures	
	Letter spelling	RAN pictures	RAN colours	
	Word spelling	RAN colours	One-min. reading	
	RAN pictures	One-min. reading	Picture-word match.	
	RAN colours	Picture-word match.	Word span	
	One-min. reading	Morphological awar.	Text spelling	
	Picture-word match.	Sentence repetition	Text length	
	Morphological awar.	Text spelling	Text-level score	
	Syntactic awareness	Text length		
	Word span	Text-level score		

6.3 Results

6.3.1 Descriptive Statistics

It is apparent from Table 6.2 that children in both language groups made considerable progress in text spelling, text length, and text-level performance. A mixed-design ANOVA (of the same type as those conducted in Study 2, Section 4 above) indicated that the new text-level aggregated score increased significantly with time, $F(1.93, 540.19) = 45.20$, $p < .001$. Planned contrasts revealed that the effect of time was considerably larger in the Time 4-Time 5 interval, $F(1, 280) = 50.36$, $p < .001$, $r = .39$, than in the Time 3-Time 4 interval, $F(1, 280) = 7.65$, $p = .006$, $r = .16$, but showed no effect of language or orthographic consistency, $F(1, 280) = 0.39$, $p = .532$, and no significant interactions, $F(1.93, 540.19) = 1.61$, $p = .203$.

Table 6.2

Descriptive statistics for Time 3, Time 4, and Time 5 writing measures

	Time 3		Time 4		Time 5	
	Mean (S.D.)	Range	Mean (S.D.)	Range	Mean (S.D.)	Range
			English			
Text spelling[a]	46.92 (23.88)	0-100	62.85 (21.43)	0-96.30	69.09 (19.07)	0-100
Text length[b]	15.48 (8.62)	0-41	24.62 (12.67)	0-65	31.62 (15.67)	0-79
Text-level[c]	69.01 (25.52)	3.80-129.60	71.49 (22.52)	10.59-118.91	84.15 (21.57)	23.26-142.49
			Spanish			
Text spelling[a]	54.60 (26.62)	0-100	81.07 (12.38)	43.09-100	85.21 (9.92)	54.73-100
Text length[b]	13.95 (8.12)	0-39	23.81 (8.16)	5-48	33.95 (12.81)	4-73
Text-level[c]	71.22 (22.39)	0-146.94	75.30 (17.83)	36.32-128.60	82.68 (18.33)	46.47-126.93

Note. [a] Percentage correct; [b] number of words; [c] aggregated score.

This is consistent with the overall profile described for text-level measures (Study 2, Section 4), which tended to show small or null variations as a function of orthographic consistency. However, text-level measures typically showed small effect sizes for time, while the aggregate score—probably as a result of the increase in statistical power—showed a moderate growth from the end of Year 1 to mid-Year 2. In this sense, text-level measures show the reverse pattern to word-level measures in that the latter had a faster rate of development during the Year 1 time points, while growth in text-level seems to accelerate in Year 2. At Time 5, the vocabulary regression-based factor scores had a mean of 0 and a standard deviation of 1 in both language groups. Recall that this score was obtained from the estimation of a factor of vocabulary, whose individual indicators were lexical density and percentage of adjectives and adverbs.

Table 6.3 shows the descriptive statistics for all predictor variables in English and Spanish. On average, verbal and non-verbal IQ estimates indicated that

Table 6.3

Descriptive statistics for Time 1, Time 3, and Time 4 predictor measures

	Time 1		Time 3		Time 4		
	Mean	S.D.	Mean	S.D.	Mean	S.D.	α/r
			English				
BD[a]	10.14	3.14	—		—		.84
VOCAB[a]	9.55	2.67	—		—		.95
LTSPELL[b]	75.98	22.33	87.39	12.66	—		.87[g]
WSPELL[f]	4.07	2.56	25.94	11.35	35.50	14.63	.81[g]
SPELL_C							.90[g]
PHONIS[b]	54.84	26.62	79.45	22.30	88.46	17.09	.94[g]
PHONBL[b]	56.06	23.79	55.55	20.17	64.94	17.20	.76[g]
PHONAW_C							.94[g]
MORPHAW[b]	46.97	16.02	61.62	15.32	—		.80[g]
SYNAW[b]	47.64	16.86	—		—		.58[g]
SENTREP[b]	—		52.65	30.91	—		.93[g]
MS_C (Time 1/Time 3)							.80[g]/.95[g]
RANpic[c]	58.10	17.63	47.17	12.50	44.84	12.63	.84[i]
RANcol[c]	67.50	25.96	53.41	17.71	45.44	16.36	.84[i]
RAN_C							.91[i]
WSPAN[d]	2.66	0.63	—		2.91	0.78	.73[j]
ONEMIN[e]	8.71	13.25	27.94	23.83	51.02	27.06	.98[h]
PWM[f]	7.36	4.06	13.42	6.84	18.07	8.11	.93[h]
READ_C							.87[h]
LTKNOW[b]	52.84	22.54	71.31	19.57	82.79	16.29	.98[g]
			Spanish				
BD[a]	11.71	3.42	—		—		.85
VOCAB[a]	9.16	3.43	—		—		.81
LTSPELL[b]	60.58	27.58	76.27	22.73	—		.90[g]
WSPELL[f]	3.49	2.31	34.89	19.01	64.11	11.04	.83[g]
SPELL_C							.94[g]
PHONIS[b]	34.52	33.57	57.87	33.48	81.66	19.72	.97[g]
PHONBL[b]	23.00	28.30	38.45	25.91	60.16	21.34	.87[g]
PHONAW_C							.97[g]
MORPHAW[b]	65.49	13.93	66.41	18.57	—		.74[g]
SYNAW[b]	46.69	15.40	—		—		.61[g]
SENTREP[b]	—		49.42	20.37	—		.82[g]
MS_C (Time 1/Time 3)							.75[g]/.89[g]
RANpic[c]	54.58	12.62	45.74	10.15	39.88	8.82	.79[i]

RANcol[c]	65.15	24.11	53.08	16.24	50.29	15.06	.82[i]
RAN_C							.86[i]
WSPAN[d]	1.83	0.53	—		1.85	0.83	.55[j]
ONEMIN[e]	7.53	12.51	27.37	18.80	53.88	15.52	.87[h]
PWM[f]	7.77	6.07	14.41	7.27	23.11	6.42	.81[h]
READ_C							.90[h]
LTKNOW[b]	37.50	26.30	61.83	27.31	82.88	14.70	.99[g]

Note. [a] Standard score; [b] percentage correct; [c] seconds; [d] items correct; [e] correct number of words; [f] correct number of words at Time 1; percentage correct at Times 3 and 4; [g] Cronbach's alpha; [h] correlation between Time 1 and Time 2 test scores; [i] correlation between trials 1 and 2; [j] split-half reliability; BD = Block Design; VOC = Vocabulary; LTSPELL = Letter Spelling; WSPELL = Word Spelling; PHONIS = Phoneme Isolation; PHONBL = Phoneme Blending; MORPHAW = Morphological Awareness; SYNAW = Syntactic Awareness; SENTREP = Snentence Repetition; RANpic = Rapid Automatized Naming with picture items; RANcol = Rapid Automatized Naming with colour items; WSPAN = Word Span; ONEMIN = One-minute Reading; PWM = Picture-Word Matching; LTKNOW = Letter Knowledge; SPELL_C = Spelling Composite; PHONAW_C = Phoneme Awareness Composite; MS_C = Morpho-syntactic Awareness Composite; RAN_C = RAN Composite; READ_C = Reading Composite.

children in both language groups were within the normal range. In February/March of Reception Year (Time 1), both groups of children showed very similar achievement levels in spelling, reading, RAN, and syntactic awareness. The Spanish children were slightly behind in phoneme awareness measures, $t(329) = 6.29$, $p < .001$ (phoneme isolation); $t(329) = 12.54$, $p < .001$ (phoneme blending); and in letter knowledge, $t(329) = 6.16$, $p < .001$, as well as in verbal short-term memory, $t(329) = 13.54$, $p < .001$. Finally, Spanish children performed better than their English peers in the morphological awareness task, $t(329) = 11.05$, $p < .001$. About 10 months later, at Time 3, Spanish children outperformed their English peers on word spelling, $t(329) = 4.69$, $p < .001$, but no significant differences were observed for word reading skills, morphological awareness, sentence repetition, and RAN. English children obtained higher scores in letter knowledge, $t(329) = 3.96$, $p < .001$, letter spelling, $t(329) = 5.93$, $p < .001$, and phoneme awareness tasks, $t(329) = 7.32$, $p < .001$ (phoneme isolation), $t(329) = 7.30$, $p < .001$ (phoneme blending).

By the end of Year 1 (Time 4) the advantage of the Spanish over the English group was more evident in word spelling, $t(329) = 19.71$, $p < .001$, and in the picture-word matching task, $t(329) = 5.61$, $p < .001$. The Spanish group performed significantly below English children in phoneme awareness, $t(329) = 4.01$, $p < .001$ (phoneme isolation), $t(329) = 2.72$, $p < .001$ (phoneme blending), but caught up with their English peers in letter knowledge, where they showed very similar average scores. Spanish children were slower than the English group for RAN pictures, $t(329) = 3.58$, $p < .001$, but English children were slower in the RAN colours task, $t(329) = 2.59$, $p < .001$. The differences in the word span task, $t(329) = 12.31$, $p < .001$ (Spanish < English), might again be attributed to the lower reliability estimates, especially in Spanish. Results where this measure is involved should thus be taken with caution.

All predictor variables were standardized within each language group and time point before running the analyses. Composite scores were computed for RAN (RAN pictures and RAN colours), phoneme awareness (phoneme isolation and phoneme blending), morphosyntactic awareness (morphological and syntactic awareness at Time 1, and morphological awareness and sentence repetition at Time 3), reading (one-minute reading and picture-word matching), and spelling (letter and word spelling). Composite scores were calculated by adding up the *z*-scores of the relevant variables and the resulting variables were re-standardized, so that all predictor variables had a mean of 0 and a standard deviation of 1. Reliability estimates for the composite measures generally improved those of individual measures and were in the .75 to .97 range (Table 6.3). Except for the word span task, all individual as well as composite measures, showed acceptable to excellent reliability estimates, making them suitable measures to be included in the regression models.

6.3.2 Longitudinal Correlations

6.3.2.1 Correlations with text spelling. Tables 6.4, 6.5, and 6.6 show the longitudinal correlations between all predictor variables and writing outcomes at Time 3, Time 4, and Time 5, respectively. Text spelling in English showed consistently high correlations with word spelling, phoneme awareness, reading, and letter knowledge at previous time points. It also showed strong autocorrelations (see Tables 6.5 and 6.5). Its longitudinal correlations with morphosyntactic awareness and RAN were of medium-large size, while its correlations with verbal short-term memory (word span) were in the medium-low range. Finally, text spelling in English was weakly associated with verbal and non-verbal IQ (see Table 6.4).

Text spelling in Spanish showed significant associations with virtually all predictor variables but, in contrast to the English group, the size of these correlations was less stable over time. Time 3 text spelling in Spanish showed large or medium-large correlations with Time 1 spelling, letter knowledge, and phoneme and morphosyntactic awareness (see Table 6.4), but the same associations were weaker in later time points. RAN and reading showed more stable associations with later text spelling performance in Spanish, which were in the medium to medium-low range. The correlations with verbal short-term memory were lower for Spanish than for English and were always of small size. Non-verbal IQ had an almost identical association with Time 3 text spelling in Spanish as in English,

Table 6.4

Product-moment correlations between Time 1 predictors and Time 3 writing measures

	Time 3 writing measures					
	English[1]			Spanish[2]		
Time 1 predictors	TSPELL	Length	TLEV	TSPELL	Length	TLEV
BD	.28**	.15*	.06	.29**	.31**	.24**
VOC	.19*	.14	.12	.33**	.30**	.18*
SPELL_C	.77**	.50**	.43**	.61**	.53**	.38**
PHONAW_C	.64**	.49**	.38**	.50**	.45**	.36**
MS_C	.44**	.27**	.23**	.46**	.45**	.33**
RAN_C	-.40**	-.38**	-.27**	-.33**	-.38**	-.18*
WSPAN	.42**	.27**	.23**	.26**	.32**	.22**
READ_C	.54**	.40**	.35**	.43**	.47**	.27**
LETKNOW	.66**	.39**	.40**	.49**	.46**	.35**

Note. BD = Block Design; VOC = Vocabulary; SPELL_C = Spelling composite; PHONAW = Phoneme Awareness composite; MS_C = Morphosyntax composite; RAN_C = RAN composite; WSPAN = Word Span; READ_C = Reading composite; LETKNOW = Letter Knowledge; TSPELL: Text Spelling correct; LENGTH = Text Length in words; TLEV = Text-level aggregated score.
* = p < .05; ** = < .01.
[1] n = 173; [2] n = 167

while the correlation with verbal IQ was larger in Spanish (see Table 6.4). Finally, autocorrelations in Spanish were in the medium-low range, in sharp contrast with English.

In short, text spelling skills showed a similar pattern of correlations with predictor variables in both languages. Notably, however, the associations in the Spanish group were less stable than in English.

6.3.2.2 Correlations with text length. In English, spelling and phoneme awareness were both strong predictors of text length, though the strength of their association decreased towards the end of the study period. Reading, RAN, and letter knowledge maintained more stable associations, which were in the medium-high range for reading, and medium for RAN and letter knowledge. Morphosyntactic awareness showed a low but steady association with later text length performance.

Table 6.5

Product-moment correlations between Time 3 predictors and Time 4 writing measures

	Time 4 writing measures					
	English[1]			Spanish[2]		
Time 3 predictors	TSPELL	Length	TLEV	TSPELL	Length	TLEV
TSPELL	.72**			.30**		
Length		.49**			.29**	
TLEV			.35**			.03
SPELL_C	.77**	.51**	.53**	.27**	.27**	.04
PHONAW_C	.64**	.31**	.36**	.15	.22**	.12
MS_C	.51**	.26**	.37**	.29**	.17*	-.02
RAN_C	-.45**	-.31**	-.31**	-.17*	-.25**	.04
READ_C	.62**	.40**	.40**	.35**	.30**	.10
LETKNOW	.68**	.32**	.36**	.17*	.24**	.07

Note. SPELL_C = Spelling composite; PHONAW = Phoneme Awareness composite; MS_C = Morphosyntax composite; RAN_C = RAN composite; READ_C = Reading composite; LETKNOW = Letter Knowledge; TSPELL: Text Spelling correct; LENGTH = Text Length in words; TLEV = Text-level aggregated score.

* = $p < .05$; ** = $< .01$.

[1] $n = 167$; [2] $n = 161$

Time 1 verbal short-term memory was only weakly correlated with text length at Time 3, and the correlation was nonsignificant in a subsequent time point (see Table 6.6). Verbal IQ was not significantly correlated with text length in English, while non-verbal IQ maintained only a weak association. Finally, autocorrelations were high for English throughout the study period (Time 3-Time 4: r = .49; Time 4-Time 5: r = .61, both significant at p < .01 level). This pattern of correlations suggests that text length in English was supported by a similar set of skills as text spelling, with word spelling and oral language skills (particularly, phoneme awareness), as well as some literacy-related skills (e.g., letter knowledge), being strongly associated throughout the period under study. In contrast with text spelling, however, some slight decrease over time in the strength of the associations was attested.

Table 6.6

Product-moment correlations between Time 4 predictors and Time 5 writing measures

	Time 5 writing measures					
	English[1]			Spanish[2]		
Time 4 predictors	TSPELL	Length	TLEV	TSPELL	Length	TLEV
TSPELL	.80**			.35**		
Length		.61**			.41**	
TLEV			.45**			.08
WSPELL	.71**	.35**	.46**	.44**	.21*	.23**
PHONAW_C	.56**	.34**	.42**	.31**	.28**	.28**
RAN_C	-.36**	-.30**	-.23**	-.32**	-.26**	-.10
WSPAN	.23**	.03	.10	.16*	.06	.20*
READ_C	.68**	.46**	.48**	.41**	.30**	.21*
LETKNOW	.67**	.32**	.38**	.26**	.22*	.26**

Note. WSPELL = Word spelling; PHONAW = Phoneme Awareness composite; RAN_C = RAN composite; WSPAN = Word Span; READ_C = Reading composite; LETKNOW = Letter Knowledge; TSPELL: Text Spelling correct; LENGTH = Text length in words; TLEV = Text-level aggregated score.
* = p < .05; ** = < .01.
[1] n = 173; [2] n = 166

In Spanish, Time 1 levels of spelling, phoneme and morphosyntactic awareness, reading, and letter knowledge were robust predictors of text length at Time 3. Similarly to the English group, the strength of the associations decreased with time. However, while in English the decline was more evident in the latter interval (i.e., the correlations between Time 4 predictors and Time 5 text length), in Spanish this was clear already in the previous time point (i.e., the correlations between predictors at Time 3 with text length at Time 4). Reading maintained a more stable, medium-size correlation with text length in Spanish throughout the study period. The correlation between Time 1 word span and Time 3 text length in Spanish was medium-low, r = .32, and was nonsignificant afterwards (see Table 6.6). In contrast to English, both verbal and non-verbal IQ showed medium-size correlations with text length in Spanish. Finally, autocorrelations were of medium size in Spanish and, thus, less strong than in English, while similarly stable.

In sum, text length in both language groups tended to be correlated with the same set of cognitive predictors as text spelling. In Spanish the pattern of longitudinal associations was more stable than for text spelling, but also showed a downward tendency in the size of the correlations, which anticipated what would be found in English later on. In this sense, the two language groups showed remarkably similar underlying correlational patterns.

Table 6.7

Product-moment correlations between predictors and Time 5 vocabulary factor scores

Predictors	English[1]	Spanish[2]
BD (Time 1)	.12	.10
VOC (Time 1)	.14	.19
TSPELL (Time 4)	.37*	-.01
READ_C (Time 4)	.35*	.12
MS_C (Time 3)	.32*	.15
RAN_C (Time 4)	-.20	-.06
WSPAN (Time 4)	.07	.25*

Note. BD = Block Design; VOC = Vocabulary; WSPELL = Word spelling; READ_C = Reading composite; MS_C = Morphosyntax composite; RAN_C = RAN composite; WSPAN = Word Span.
* = $p < .01$
[1] n = 173; [2] n = 166

6.3.2.3 Correlations with text-level performance. Spelling and reading showed the highest and most stable correlations with text-level performance in English throughout the study period. Phoneme awareness and letter knowledge were also consistently correlated with text-level scores in English, with correlation sizes in the medium-high range. Morphosyntactic awareness and RAN maintained medium-low correlations with text-level performance, while verbal short-term memory showed weak or null associations. Verbal and non-verbal IQ were not significantly correlated with text-level scores in English. In sum, a remarkably similar pattern of correlations underlay levels of text spelling, text length, and text-level performance in the English group. The main difference from one writing outcome to another may be observed in the size of the correlations, which are strongest with text spelling, slightly weaker with text length, and weakest with text-level performance.

In Spanish, the longitudinal correlation patterns with text-level scores were distinct both from those shown by the English group, as well as in comparison to other writing outcomes. No single predictor was very strongly correlated with later achievements in text level in Spanish. Correlations between Time 1 predictors and Time 3 text-level scores were all of medium or small size. The highest correlations with Time 3 text-level scores in Spanish were with Time 1 spelling, phoneme and morphosyntactic awareness, and with letter knowledge (see Table 6.4). The remaining correlations at that time point were significant, but low. All of the correlations between Time 3 predictors and Time 4 text-level scores in Spanish were nonsignificant, and all correlations between Time 4 predictors and Time 5 text-level scores were low or nonsignificant. Of special concern is the lack of significant autocorrelations among text-level scores in Spanish. This is surprising, and may be taken to suggest that the reliability of the text-level construct is questionable in this language group.

6.3.2.4 Correlations with vocabulary factor scores. Very few significant correlations were observed between Time 3 and Time 4 cognitive predictors and vocabulary factor scores (Table 6.7). In Spanish, only word span was a significant predictor, and the correlation was, moreover, of small size. In English end-of-Year 1 levels of word spelling and reading were significantly correlated with mid-Year 2 vocabulary factor scores. Also, and as hypothesized, the morphosyntactic awareness composite was also significantly correlated with vocabulary. No other significant correlations were attested.

6.3.3 Structural Equation Models

Analyses were run in Mplus 6.1 (Muthén & Muthén, 1998-2011) as a series of structural equation models where language was the grouping variable. Missing data were imputed using maximum likelihood estimation with robust standard errors. For all dependent variables, an initial unconstrained model in which no parameters were fixed to be equal was run, including the Time 1 predictor variables and the Time 3 writing outcome in question (text spelling, text length, or the text-level performance aggregated score). This yielded the baseline model, which was used to identify those regression paths in each language group that were significant. Variables that made a significant contribution to explaining variance in the outcome variable in either language were retained, while nonsignificant paths in both languages were dropped from the model. A preliminary nested model was then run, fixing one of the regression parameters to be equal across languages. If no significant differences between the baseline model and the nested model were found, then another regression path was fixed to be equal across the two language groups. This more constrained, second nested model was then tested against the first nested model, and so on. This procedure was followed until all regression paths had been tested. When fixing regression paths to be equal in both languages resulted in the most constrained model being significantly different from the preceding, less constrained one, then that path was freely estimated over the two language groups. Once all regression paths had been tested, a similar iterative procedure was followed for constraining correlation paths to be equal, and the resulting model was again tested against the previous, less constrained one. Lastly, the process of testing constraints was repeated for residual variances of all variables; these were incrementally fixed to be equal and tested against the previous, less constrained model.

With this final version of the Time 1-Time 3 model, the regression paths of each of the Time 4 writing outcome variables were examined with Time 3 predictor variables, and another wave of testing nested versus less constrained models ensued. Finally, the Time 4 predictors and the corresponding Time 5 writing outcome variable were entered following the same iterative model fitting procedures. In all cases, the Satorra-Bentler scaled difference Chi-square test (S-B $\chi^2\Delta$; Satorra & Bentler, 2001) was used to determine whether two models were significantly different from each other.

6.3.3.1 Predicting text spelling. Only Time 1 spelling and phoneme awareness, in English, and Time 1 spelling and RAN, in Spanish, were significant predictors of text spelling at Time 3. The high correlations between some of the predictor variables (in particular between spelling and letter knowledge) resulted in suppressor effects; therefore, letter knowledge was excluded from subsequent versions of the model, and only spelling, phoneme awareness, and RAN were retained as Time 3 predictors. Through iterative testing, all regression paths and correlations of the model with Time 1 predictors were constrained to be equal across the two languages. All residual variances were also fixed, except for the residual variance in the outcome variable, text spelling at Time 3. This model, depicted in Figures 6.1 (for English) and 6.2 (for Spanish), fitted the data extremely well, $\chi^2(9, N = 378) = 6.55$, $p = .684$, Root Mean Square Error of Approximation (RMSEA) = 0.000 (90% CI = 0.000-0.064), Comparative Fit Index (CFI) = 1.000, Tucker-Lewis Index (TLI) = 1.003, Standardized Root Mean Residual (SRMR) = .047. A model in which all residual variances were fixed to be equal was significantly different from one where text spelling at Time 3 was free to vary, S-B $\chi^2\Delta(1) = 29.35$, $p < .001$. Up until mid-Year 1, then, children's text spelling performance in both language groups seems to rely on the same set of cognitive skills, in spite of the fact that the Spanish group obtained significantly higher scores (see Study 2, Section 4). Moreover, these results are also consistent with a larger study on cognitive predictors of word spelling across four different languages, with varying degrees of orthographic consistency (Caravolas et al., 2012).

In modelling the predictors of text spelling at Time 4, a model where the path from Time 3 phoneme awareness to Time 4 text spelling was left unconstrained, while all other regression paths and correlations were fixed to be equal over languages, fitted the data moderately well[13], $\chi^2(33, N = 378) = 52.46$, $p = .001$, RMSEA = 0.056 (90% CI = 0.024-0.083), CFI = .994, TLI = .991, SRMR = .107, whereas one where phoneme awareness was fixed to be equal across languages was significantly different, S-B $\chi^2\Delta(1) = 158.29$, $p < .001$. The negative standardized regression coefficient of Time 4 text spelling on Time 3 phoneme awareness in the Spanish model clearly indicated suppressor effects between word spelling and phoneme awareness. Indeed, the bivariate correlation between phoneme awareness and word spelling in Spanish at Time 3 was very high, $r = .80$, whereas it was lower in English, $r = .66$. The Time 4 model where only regression and correlation paths had been fixed (except for phoneme awareness) was significantly different from another model where also residual variances were constrained to be equal, S-B $\chi^2\Delta(5) = 16.78$, $p < .01$. Presumably, this deterioration in fit reflected the marked reduction in variance in text spelling of the Spanish group at Time 4 (as mastery level in text spelling accuracy was reached)—both relative to itself at the earlier time point, and crucially relative to the English group at Time 4 (see Table 6.2). In short, both language groups showed remarkably similar patterns in the cognitive predictors that support text spelling performance from Reception Year and up until the end of Year 1. Note, however, that the few unconstrained parameters impacted the amount of variance explained in each language group at this phase of development.

The models that were tested to predict text spelling performance at Time 5 (mid-Year 2) showed differences between the two groups. The best fitting model was one where the Time 4 autoregressor and Time 4 word spelling were freely estimated over groups, while the regression paths from phoneme awareness and RAN were fixed to be equal, $\chi^2(87, N = 378) = 166.76$, $p = .001$, RMSEA = 0.070 (90% CI = 0.054-0.086), CFI = .972, TLI = .960, SRMR = .091. A model where all Time 4 regression paths were fixed to be equal in both languages fitted the data less well, S-B $\chi^2\Delta(2) = 32.00$, $p < .001$, as did another model where Time 4 RAN and

[13] The χ^2 statistic is very sensitive to sample size and tends to reject models that use large sample sizes (Hooper, Coughlan, & Mullen, 2008). Therefore, to interpret significant χ^2, Wheaton, Muthén, Alwin, & Summers' (1977) relative χ^2 was used (χ^2/df). Recommended ratios range between 5 and 2 (Hooper, et al., 2008).

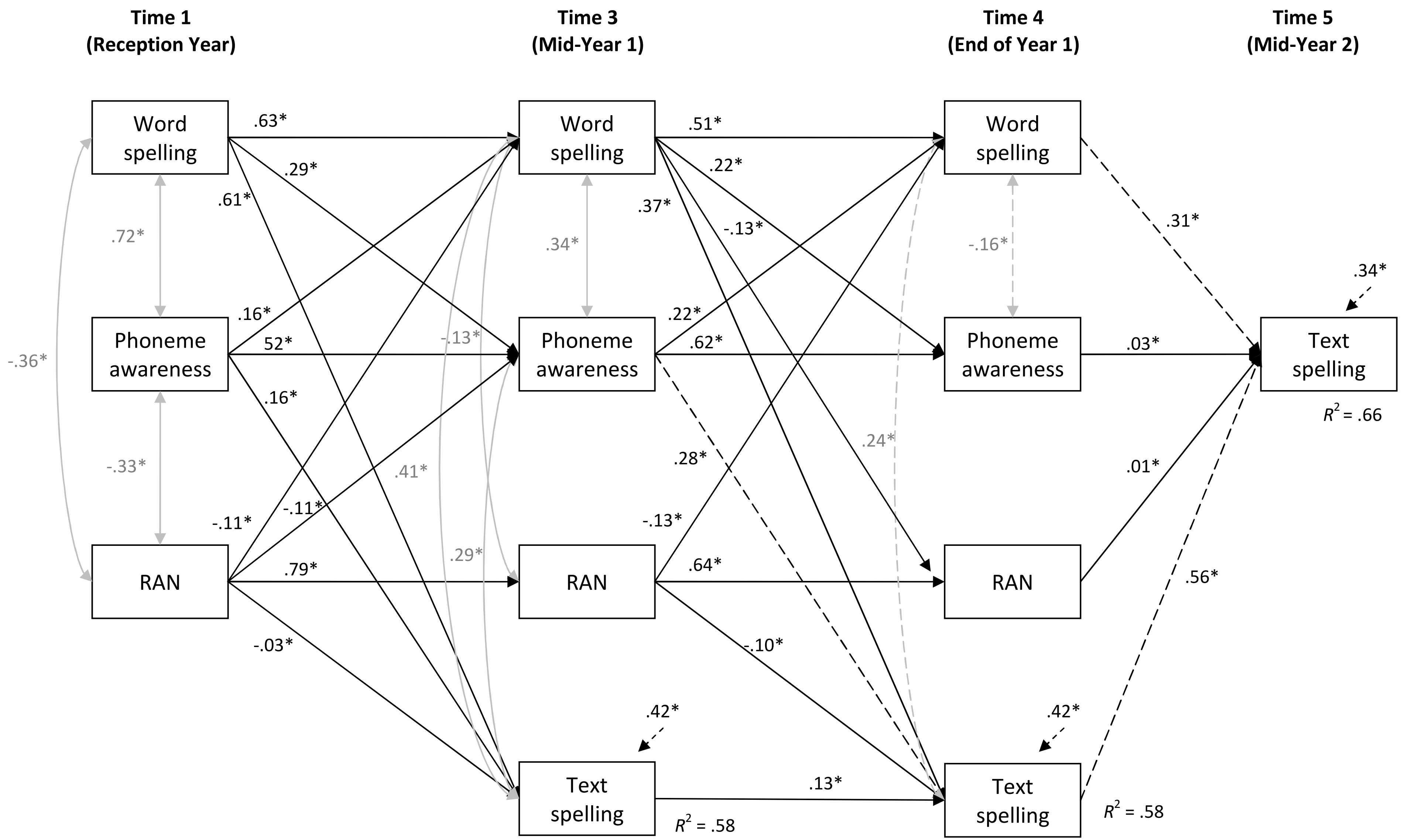


Figure 6.1. Path models of growth in text spelling skill in English from mid-Year 1 through to mid-Year 2. Solid black arrows indicate that the regression path was constrained to be equal across languages. Dotted black arrows indicate that the regression path was left unconstrained. Grey arrows and figures indicate correlations. Nonsignificant paths were omitted, for clarity. Standardized path weights are shown and asterisks indicate significant effects (p < .05)

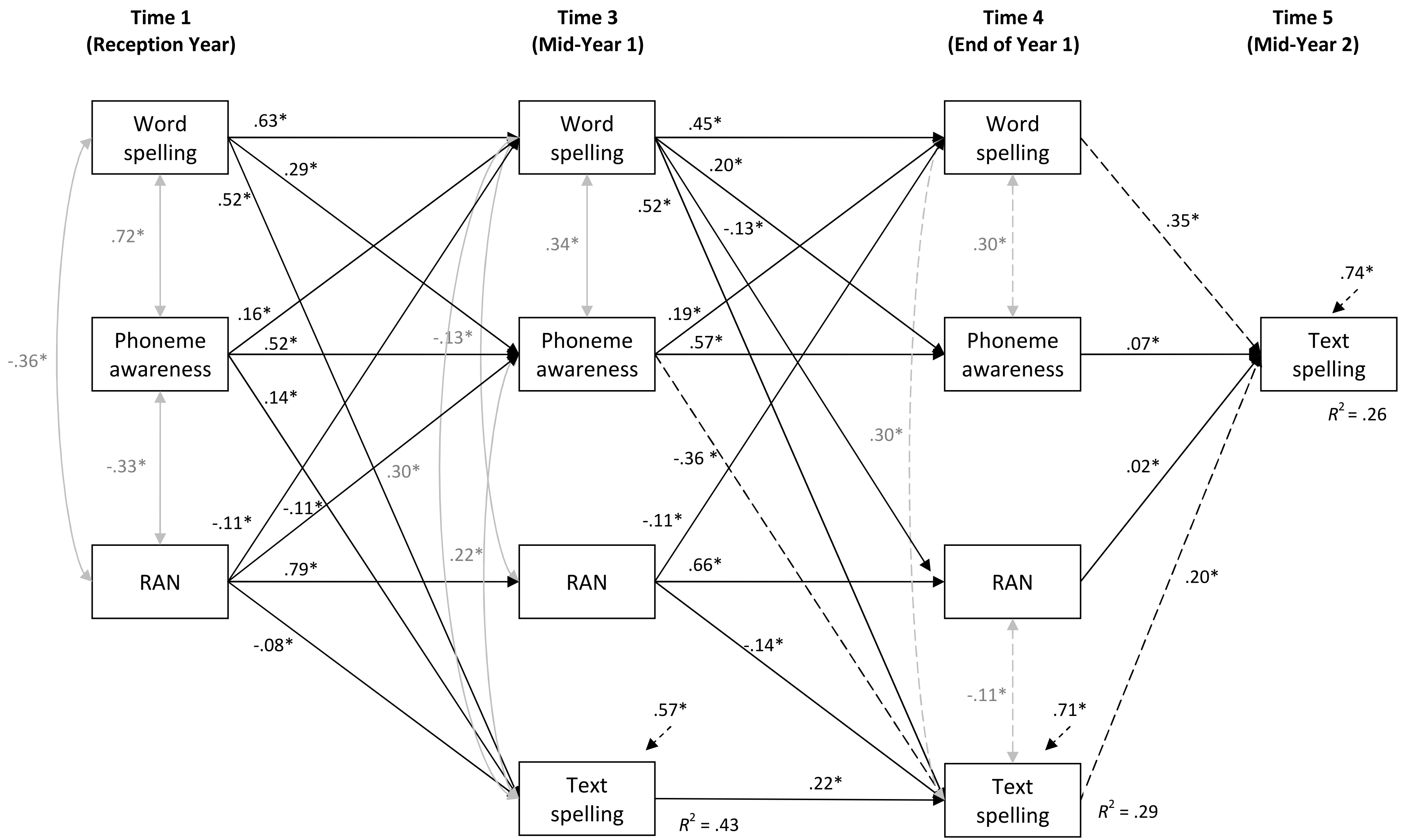


Figure 6.2. Path models of growth in text spelling skill in Spanish from mid-Year 1 through to mid-Year 2. Solid black arrows indicate that the regression path was constrained to be equal across languages. Dotted black arrows indicate that the regression path was left unconstrained. Grey arrows and figures indicate correlations. Nonsignificant paths were omitted, for clarity. Standardized path weights are shown and asterisks indicate significant effects (p < .05)

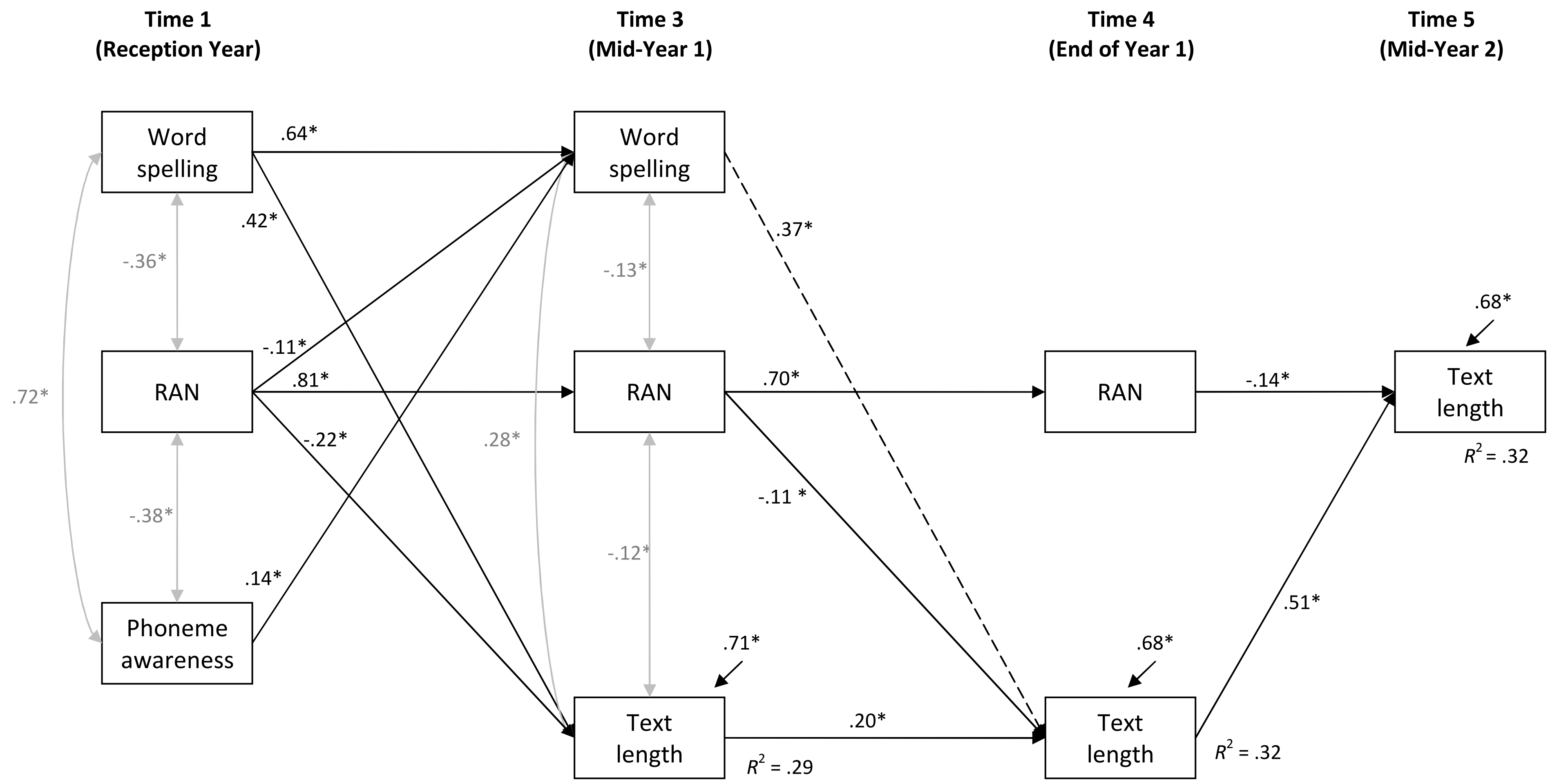


Figure 6.3. Path models of growth in text length (number of words) in English from mid-Year 1 through to mid-Year 2. Solid arrows indicate that the path was constrained to be equal across languages. Dotted arrows indicate that the path was left unconstrained. Grey arrows and figures indicate correlations. Nonsignificant regression paths were omitted, for clarity. Standardized path weights are shown and asterisks indicate significant effects ($p < .05$)

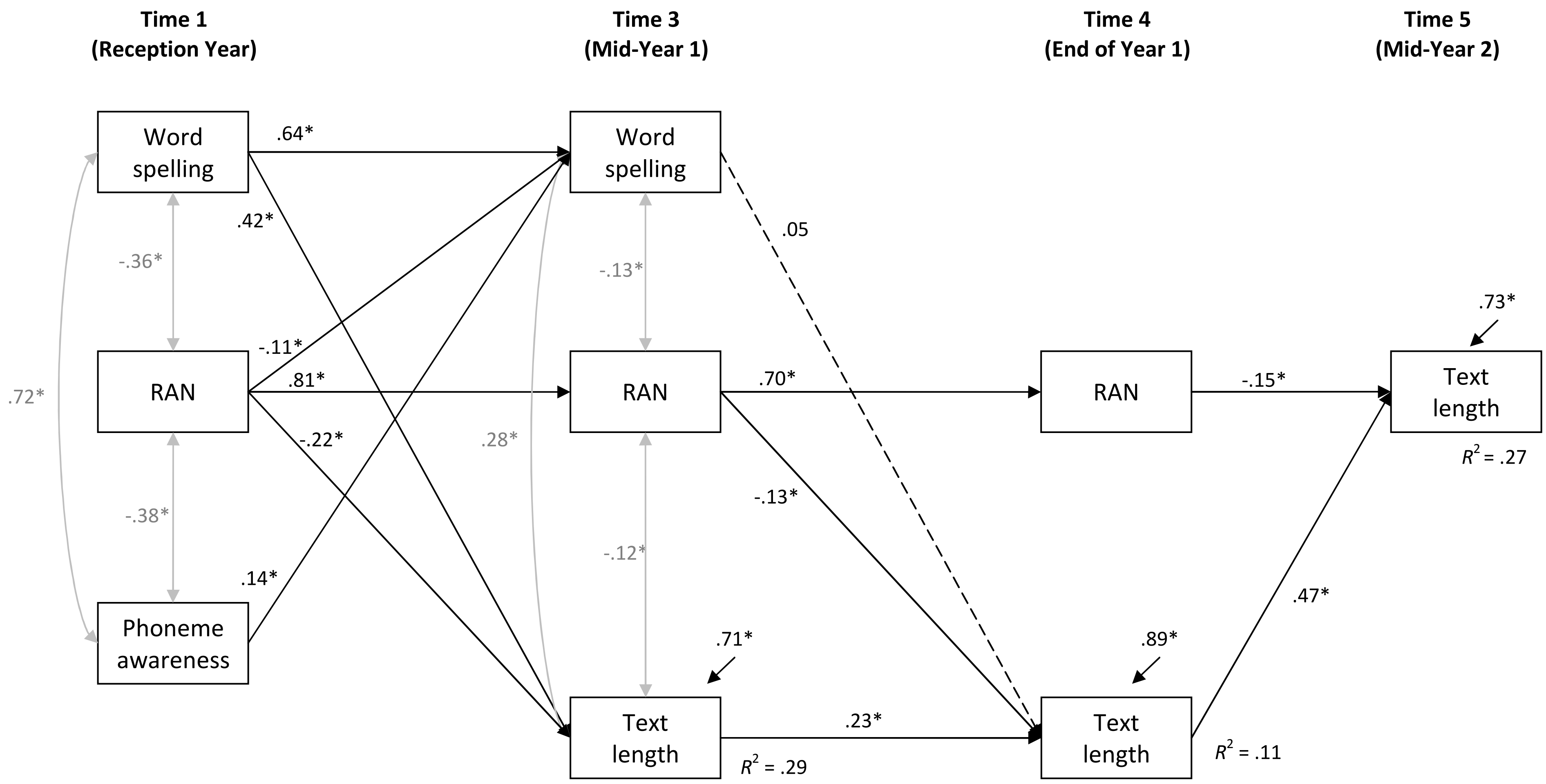


Figure 6.4. Path models of growth in text length (number of words) in Spanish from mid-Year 1 through to mid-Year 2. Solid arrows indicate that the path was constrained to be equal across languages. Dotted arrows indicate that the path was left unconstrained. Grey arrows and figures indicate correlations. Nonsignificant regression paths were omitted, for clarity. Standardized path weights are shown and asterisks indicate significant effects (*p* < .05)

Time 4 phoneme awareness regression paths, as well as Time 4 correlations, were fixed (with the Time 4 word and text spelling regression paths left free to vary), S-B $\chi^2\Delta(6) = 74.23$, $p < .001$.

It would appear from the model (Figures 6.1 and 6.2) that differences as a function of language or orthographic consistency emerge between the end of Year 1 and mid-Year 2. However, this divergence may reflect some artefacts of the text writing task as suggested by the considerably lower proportion of variance explained by text writing at Time 4 in Spanish as compared to English. That is, notably, the autoregressor effects are much larger in English, while word spelling is the most important predictor in Spanish. However, this is a product of the fact that Spanish children were showing near-ceiling effects in text spelling performance in the previous time point, but not in the experimental word spelling task, which was graded and designed to be a sensitive measure of spelling ability over the primary school years. These divergent patterns suggest that the text spelling and isolated word spelling become differentially sensitive estimators of spelling ability by end-Year 1, while both task types remain sensitive estimators of spelling in English. Despite these differences, however, both language groups showed that text spelling at Time 5 was mainly accounted for by earlier spelling ability, while the component skills (phoneme awareness and RAN) showed a sharp decrease in predictive power at this point. The positive sign in the regression path from Time 4 RAN to Time 5 text spelling is, thus, negligible.

6.3.3.2 Predicting text length. Only Time 1 spelling and Time 1 RAN made significant contributions to explaining variance in Time 3 text length in both English and Spanish. Again, suppressor effects determined that other highly correlated predictors, such as Time 1 letter knowledge or reading, were excluded from subsequent analyses. A model where all parameters were fixed to be equal (regression paths, correlations, and residual variances) was a good fit to the data, $\chi^2(6, N = 378) = 4.37$, $p = .627$, RMSEA = 0.000 (90% CI = 0.000-0.079), CFI = 1.000, TLI = 1.009, SRMR = .038. A similar model where residual variances and correlations were unconstrained was equally good, S-B $\chi^2\Delta(4) = 3.25$, $p > .05$. The final models predicting text length from Reception Year until mid-Year 2 are shown in figures 6.2a (English) and 6.2b (Spanish). As predicted, RAN explained unique variance in the number of words produced by children during text composition in mid-Year 1, over and above the strong contribution of the spelling composite. These results indicate that the striking similarities in the average number of words that children produced at this developmental stage (Study 2) are further substantiated by the same set of cognitive predictors.

When Time 3 predictors and Time 4 text length were introduced in the model, a model where Time 3 spelling was unconstrained was a better fit than one where this parameter was fixed, S-B $\chi^2\Delta(1) = 16.69$, $p < .001$, and was similar to one where correlations were constrained to be equal over groups, S-B $\chi^2\Delta(1) = 1.80$, $p > .05$. This last model, where Time 3 spelling was free to vary, while all other regression paths and correlations were fixed to be equal over groups, fitted the data well, $\chi^2(26, N = 378) = 25.86$, $p = .471$, RMSEA = 0.000 (90% CI = 0.000-0.057), CFI = 1.000, TLI = 1.000, SRMR = .042. Another model where residual variances where also fixed was a significantly worse fit to the data, S-B $\chi^2\Delta(5) = 18.39$, $p < .01$. In line with findings in Study 3 (Section 5), the power of spelling to predict number of words during text production was lost for Spanish, while it still explained variance in English. This resulted in a major drop in the proportion of variance explained in Spanish. Interestingly, the role of RAN was remarkably stable, while both language groups showed equal autoregressor effects.

Time 4 spelling did not make a significant contribution to explaining variance in text length in English or Spanish. For this reason, it was removed from subsequent analyses. Thus, the last step of the model included only RAN and text length at Time 4 as the only predictor variables. A model where all regression paths, correlations, and residual variances were constrained to be equal fitted the data moderately well, $\chi^2(59, N = 378) = 97.75$, $p = .001$, RMSEA = 0.059 (90% CI = 0.037-0.079), CFI = .976, TLI = .973, SRMR = .104. This model was then preferred over another one where correlations and residual variances were freely estimated over languages, which fitted the data similarly well, S-B $\chi^2\Delta(4) = 7.37$, $p > .05$. Removing Time 4 spelling as a predictor allowed the powerful autoregressor effects to emerge, while it did not affect the unique contribution of RAN, which remained stable from Reception Year and until mid-Year 2. Also, the proportion of variance explained in each language group was again very similar ($R^2 = .32$, for English; $R^2 = .27$, for Spanish) and remarkably stable if the entire period under scrutiny is considered.

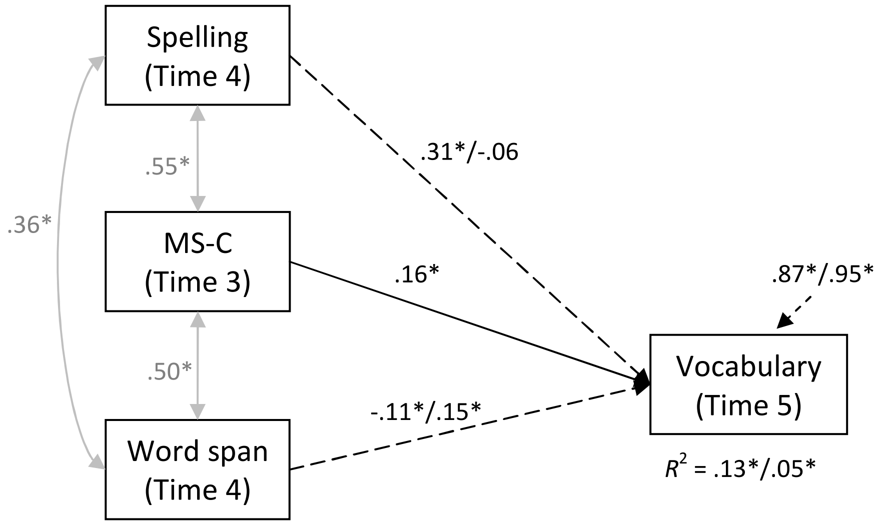


Figure 6.5 Path models of the Time 5 vocabulary factor scores with Time 3 and Time 4 predictor measures. Path and correlation coefficients to the left correspond to English, and rightmost to Spanish. Solid arrows indicate that the path was constrained to be equal across languages and only one path/correlation coefficient is reported. Dotted arrows indicate that the regression path was left unconstrained. Grey arrows and figures indicate correlations. Nonsignificant regression paths were omitted, for clarity. Standardized path weights are shown and asterisks indicate significant effects (p < .05)

6.3.3.3 Predicting vocabulary factor scores. The predictors of Time 5 written vocabulary factor scores provided a very poor explanatory model in both language groups. Block design, vocabulary, RAN, and reading did not make any unique contributions to explaining variance in the outcome variable and were dropped from the model. Thus, only spelling, word span, and morphosyntax were retained. In the best fitting model (Figure 6.5) only the regression path from the morphosyntax composite and all regression coefficients were constrained to be equal across language groups. The model fitted the data well, $\chi^2(4, N = 378) = 1.94$, $p = .748$, RMSEA = 0.000 (90% CI = 0.000-0.077), CFI = 1.000, TLI = 1.042, SRMR = .025, but it only explained a meagre 13% and 5% of the variance in English and Spanish, respectively. A model where correlations were unconstrained was an equally good fit, S-B $\chi^2\Delta(3) = 1.82$, $p > .05$, while a model where word spelling and word span were also constrained to be equal was a significantly poorer fit, S-B $\chi^2\Delta(2) = 27.82$, $p < .001$.

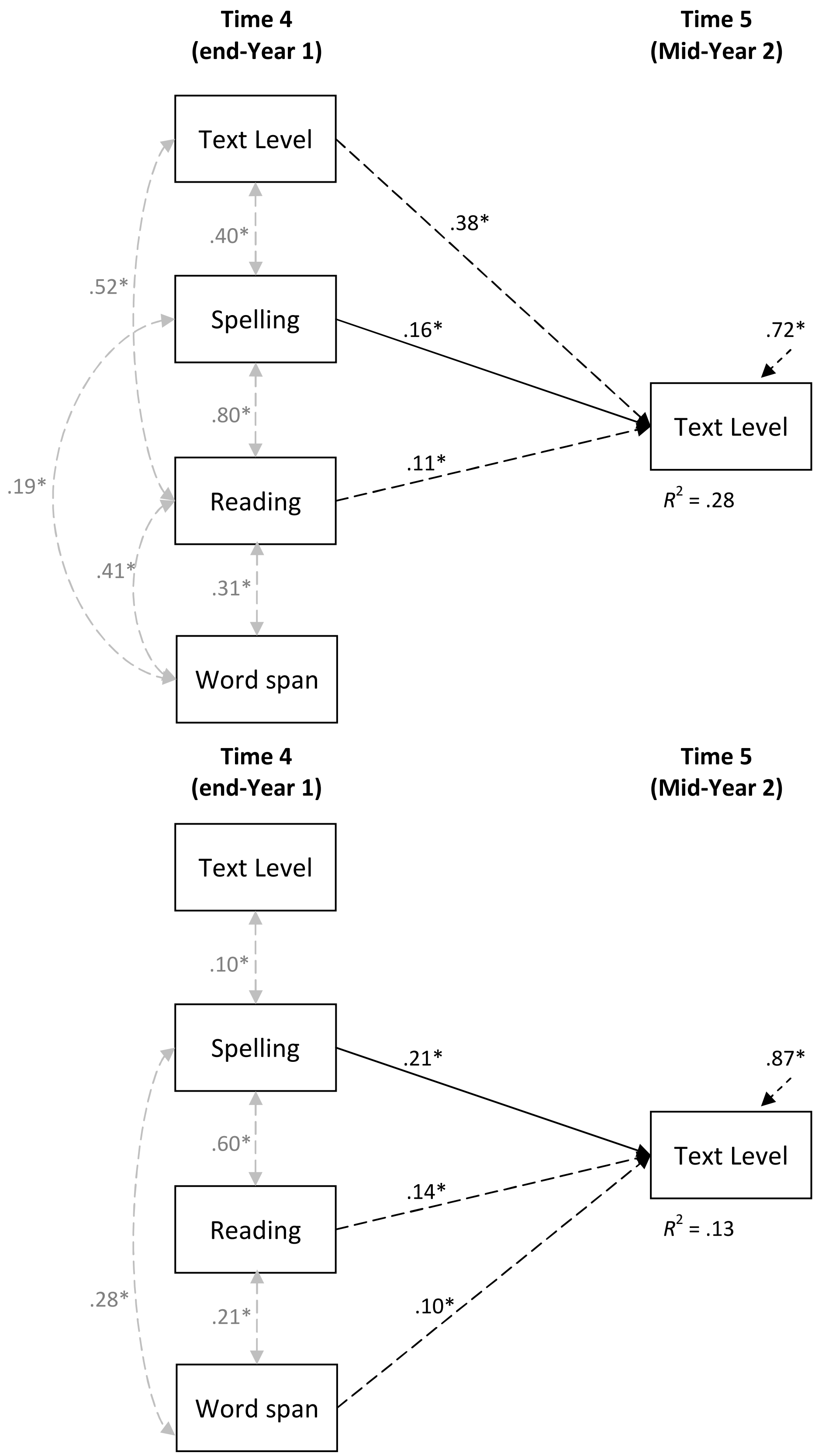


Figure 6.6 Path models of text-level performance with Time 4 predictor measures in English (on top) and Spanish (below). Solid arrows indicate that the path was constrained to be equal across languages. Dotted arrows indicate that the regression path was left unconstrained. Grey arrows and figures indicate correlations. Nonsignificant regression paths were omitted, for clarity. Standardized path weights are shown and asterisks indicate significant effects (p < .05)

6.3.34 Predicting text-level performance. Models predicting text-level performance at Time 3 and Time 4 yielded uninterpretable results. For this reason, only models predicting Time 5 text-level scores from Time 4 predictor variables are reported. In English, Time 4 text-level, word spelling, and reading scores were all significant predictors of text-level performance at Time 5. In Spanish, the effect of the autoregressor was nonsignificant, indicating problems with the stability of this construct in this language. Similarly to English, however, both spelling and reading were significantly associated with text-level scores at Time 5. Moreover, word span was significantly associated to Time 5 text-level performance in Spanish only.

In the best fitting model predicting Time 5 text-level performance (Figure 6.6), only the regression path from Time 4 spelling was fixed to be equal over languages. This model fitted the data well, $\chi^2(2, N = 378) = 4.12$, $p = .127$, RMSEA = 0.077 (90% CI = 0.000-0.183), CFI = .991, TLI = .953, SRMR = .022. A more constrained model where also the regression paths from the text-level, reading, and word span at Time 4 were fixed fitted the data significantly worse, S-B $\chi^2\Delta(2) = 39.72$, $p < .001$, and so did one where the regression path from Time 4 spelling and all correlation paths were fixed, S-B $\chi^2\Delta(6) = 49.81$, $p < .001$. Aside from the contribution of previous text-level performance at Time 4, which was the strongest predictor in English (though made no contribution in Spanish), the levels of word spelling and word reading that children showed at the end of Year 1 predicted their text-level scores some 6 months later in both language groups. In Spanish, children's verbal short-term memory scores also contributed to explaining variance in text level performance. However, the lack of autoregressive effects in Spanish clearly indicates that the text-level construct (both as individual variables, as in the previous Studies 2 and 3 above, and as an aggregated score, as in the present study) did not provide a reliable measure of children's microstructural text-level abilities during text composition. This issue is further elaborated in the Discussion section below.

6.4 Discussion

The main objective of this study was to determine the cognitive predictors of mid-Year 1 to mid-Year 2 writing skills from a cross-linguistic perspective. A series of cognitive measures hypothesized to be instrumental to the early development of text writing skills was traced from the time when children were attending Reception Year classes and until mid-Year 2, thus comprising 2 years from the outset of literacy instruction. Importantly, the study period begins when children had only had a few months of formal literacy education. Therefore, this study explores the foundational stages of writing skills in a formal context.

The first research question concerned the cognitive skills that predicted later achievements in the expression of microstructural aspects of writing: text spelling, text productivity (text length), and text-level writing. In line with the hypothesis, text spelling was predicted by a similar set of cognitive predictors to those reported to support word spelling across alphabetic orthograhies: word spelling, phoneme awareness, and RAN (Caravolas et al., 2012). Caravolas et al. (2012) also found letter knowledge to be a strong predictor of spelling skills across a range of orthographies, but in the present study the high correlations among the predictor variables resulted in suppressor effects, in particular with single word spelling, and this variable was dropped from subsequent analyses. Notably, the effect of letter knowledge is subsumed primarily by the word spelling measure. This rather reduced set of cognitive predictors, along with the contribution of the autoregressors, accounted for a large proportion of variance in text spelling in English (range in R^2: .58 - .66), though considerably less variance was explained in Spanish (range in R^2: .26 - .43). This is probably because the Spanish group showed near-ceiling levels of text spelling as early as Time 4 (end of Year 1).

Only two skills reliably predicted gains in text length: prior spelling ability and RAN. Again, the high correlations of other literacy-related measures with spelling prevented their having a substantial independent contribution to the models. About a third of the variance in text length in both language groups was explained on the basis of these two predictors until the end of Year 1. After that, strong autoregressor effects made the single word spelling measure redundant. Of interest is the remarkably stable role of RAN throughout the period under study, which maintained a modest but steady association with later levels of text productivity in both languages. To my knowledge, this is the first time that RAN has been considered as a skill associated with writing development beyond the word level. It is likely that this reliable and stable role of RAN reflects the need for fast retrieval of content as well as of phonographic mappings. It also seems likely that, due to its characteristics, it is uninfluenced by variations in orthographic consistency or by training, thus ensuring longitudinal stability and constituting an independent predictor of literacy skills. Future studies should address in what ways non-alphanumeric RAN facilitates production of content or boosts other mechanisms of relevance during text composition.

Altogether, the models offered a poor account of the cognitive foundations of text-level scores at mid-Year 2, both in the case of the vocabulary factor as well as for the analytic text-level scores. A possible interpretation of the poor explanatory value of text-level models may be related to the maturity and reliability of the constructs *per se*. Very little support was found for the stability and reliability of the text-level constructs beyond productivity. In this sense, the text-level analytic score was also an unreliable construct. This finding is, however, not unprecendented. Babayiğit and Stainthorp (2011) found no significant regression paths in a model where IQ, vocabulary, working memory, spelling accuracy, and age were entered as exogenous predictors. Only autoregressive effects were attested, between two consecutive measures of writing (assessed via an analytic score).

It was surprising to find that a measure assessing oral vocabulary was not significantly associated with written vocabulary scores in either language group. One reason for the lack of association between these two variables might be that there was a rather long interval between the time when oral vocabulary was assessed (mid-Reception year) and the collection of texts from which vocabulary scores were obtained (mid-Year 2). Children's preschool levels of vocabulary may have been substantially altered during this period of time, hence the loss of a connection between the two measures. Certainly, multiple new opportunities for vocabulary gains are opened up when children begin primary school—crucially, the role of reading and spelling in vocabulary learning.

The second research question addressed the issue of the stability in the predictors over the period under study. It was hypothesized that, at the initial stages, the underlying cognitive predictors would make more substantial and consistent contributions; however, as constructs stabilized, autoregressive effects would make other predictors redundant, in line with findings for other facets of literacy development (Caravolas et al., 2013). This prediction was corroborated by the data. Indeed, in the text length models in English and in Spanish, as well as in the text spelling model in English, powerful autoregressive effects were attested and other predictors had to be dropped (as was the case of Time 4 spelling in the text length models) or they made significant, but effectively negligible, contributions (as was the case of Time 4 RAN in the text spelling models). A minor exception was constituted by text spelling in Spanish at Time 5, where the contribution of word spelling was larger than that of the autoregressor. However, as mentioned, this may be the result of the near-ceiling effects attested for text spelling by the Spanish children already at Time 4.

The third research question was related to the effect of the differences in orthographic consistency between English and Spanish. It was predicted that, since Spanish children showed an earlier command in word-level writing (as shown in Study 2, Section 4), their dependence upon spelling skills for higher-level writing features would last a shorter period of time in comparison to English. This hypothesis was only partially supported by the data. The two language groups started to diverge as early as in the Time 3 predictors of text length: The contribution of spelling in the Spanish model was negligible, while in English it still made the most substantial contribution. Similarly, at Time 4 in the text spelling models both the Time 4 autoregressors and Time 4 spelling made different contributions to the models in each language group, which were larger in English. In other words, English showed a protracted influence of spelling skills on writing.

In sum, this study has provided robust evidence that some microstructural aspects of writing—specifically, text spelling and written productivity—rely on the same set of cognitive predictors: spelling, phoneme awareness, and RAN. Crucially, this finding seems to hold across alphabetic writing systems, since the two languages under scrutiny are good representatives of the two ends of the orthographic consistency continuum. Whenever differences were observed, they were quantitative, rather than qualitative, in nature. For example, they were apparent in the size of correlations or in the proportions of explained variance, more so than constituting essentially different underlying patterns. Finally, support was found for a foundational stage of writing development characterized by strong cross-linguistic similarities, while divergence would start to take shape around the end of Year 1.

6.5 Limitations

Measurement of text-level outcomes has been shown to be unreliable in this study. The most plausible explanation for this is that writing skill beyond the word level is simply not yet sufficiently developed or stable to warrant measurement within the first year or two of literacy learning. Additionally, it may be that using proportion scores, to control for the effect of text length on the measures, limited the scope for variation at this level of performance. In addition, the type analytic scale used in this study may not be suitable to assess text-level writing development, although we are not aware of an alternative objective measurement scale/method. While the measure adopted here aimed to avoid the manual, qualitative rating of different aspects of text construction and used text-based quantities instead, it might not be suitable for the study of text features which are at their starting point, from a developmental perspective. On the other hand, one should make room for the possibility that key predictor variables were not included in the present study. Future research should investigate in depth the cognitive underpinnings of vocabulary and other text-level writing features in early text composition.

A final note is concerned with the absence of a handwriting measure in the text-length models. Handwriting is a key predictor of productivity in text production in languages with consistent and inconsistent orthographies (e.g., Babayiğit & Stainthorp, 2010; Berninger et al., 1992; Graham et al., 1997; Wagner, et al., 2011). It seems safe to assume that adding a handwriting component to the models presented here would have considerably increased the proportion of variance explained in the dependent variable. At the same time, including a handwriting component would likely lessen the absolute contribution of the spelling component—of more consequence here, because of the cross-linguistic nature of the study—, given the high correlations reported in previous studies (e.g., Graham, et al., 1997)[14]. However, the predictive power of RAN should be comparable, since handwriting and RAN do not appear to draw on closely related constructs. In other words, a large portion of the variance explained in text length by the models proposed in this study should not overlap with a similar model including handwriting skills. Future work should strive to disentangle the overlap between RAN and handwriting and, especially, between handwriting and spelling.

---

[14] See Section 5.1.2 for a rationale for focusing on spelling rather than on handwriting in cross-linguistic research

Chapter 7

General Discussion

The main goal of this thesis was to contribute to the field of writing development with a systematic comparison of the early writing skills of English- and Spanish-speaking children. The studies that compose this work are unique in that they constitute a first attempt to provide a large-scale, cross-linguistic, longitudinal exploration of the writing skills of two groups of children who were learning orthographies with highly contrasting degrees of consistency. Three main issues were explored: (1) the role of orthographic consistency in the attainment levels of early text composition; (2) the dependency of writing features on spelling skill; and (3) the cognitive underpinnings of early written composition. This chapter then discusses the main cross-linguistic findings, as well as the implications for current models of writing development. In addition, methodological issues, limitations of the studies, and suggestions for future research are considered.

7.1 Cross-linguistic Findings

A chief outcome of the present thesis is the finding that differences in orthographic consistency do not determine the development of text writing beyond the word level. Spanish children, who were learning a consistent orthography, showed spelling skills superior to those shown by their English peers, but both groups performed similarly in most text-level features, including the amount of text produced under time constraints. That English children were outscored by the Spanish children in word-level skills, such as spelling, is line with a large body of research showing the effects of orthographic consistency on word-level literacy (e.g., Caravolas & Bruck, 1993; Seymour et al., 2003; Share, 2008). These results thus extend our understanding of the effects of orthographic consistency to *text* spelling, as well as to other word-level writing aspects, such as word segmentation and capitalization. At the same time, they suggest that early writing development at higher levels of language—sentence, text—is relatively unaffected by variations in orthographic consistency.

The assessment of the dimensionality of early writing development was also characterized by cross-linguistic similarities. Support for virtually the same constructs was obtained in both language groups at all time points. Word-level/transcription and productivity were the clearest and most robust dimensions in both languages, and they were composed of the same set of individual variables. This finding is in line with recent research looking at the dimensionality of writing in other orthographies and writing systems (Guan, Ye, Wagner, & Meng, 2012).

This thesis has also demonstrated that the cognitive underpinnings of text spelling and written productivity are shared by languages whose orthographies differ greatly in terms of the consistency of their phonographic mappings, at least during the first year and a half of formal literacy instruction. This result further substantiates claims in favour of common (perhaps, universal) patterns in the foundational stage of literacy development across alphabetic orthographies. (e.g., Caravolas et al., 2012; 2013; Ziegler et al., 2010), by extending such claims to the text writing domain.

A possible explanation for the lack of cross-linguistic differences in writing beyond word-level performance could be that, at these initial stages, children are not much concerned with representing words conventionally, while they are satisfied with representing their phonological structure (Chomsky, 1970; 1976). Thus, as long as they have a general grasp of the alphabetic principle and know enough (though not all) phoneme-letter pairings, they may carry on producing connected text using their unconventional spellings. As Chomsky (1976) pointed out,

> The children spell independently, making their own decisions. They have no preconceptions of how the word ought to be spelled, not any expectation that there is a right or a wrong way to do it. They spell creatively, according to some combination of what they perceive and what they consider worthy of representation [...]. Once the children get started, they can go on to write any message at all. For it is not that they know the spelling of certain words. Rather they possess the means to write any words. (p. 6-8)

Share (2008) claimed that, until children have learned all the basic phonographic mappings of their corresponding spelling systems, all orthographies are “functionally opaque” (p. 597). This means that, if a child ignores the ways in which the phonological structure of single words is represented in writing, s/he cannot read words, even if s/he already understands that letters stand for sounds. Of course, whether the orthography is more or less consistent is irrelevant at this point. I would like to suggest that, within the framework of spelling/writing development, all orthographies are “functionally transparent” (or consistent) to early writers. This is because the writer, in contrast to the reader, has control over the semantics that s/he wants to convey (Chomsky, 1976). Therefore, as long as they can roughly represent the phonological structure of words, they can proceed with the writing process.

Differences across groups were not only rooted in disparities in orthographic consistency. Some cross-linguistic differences were better explained on the basis of oral language contrasts. For example, English children’s higher average percentage of adjectives and adverbs, or Spanish children’s larger mean content-word length or lexical diversity. On the contrary, striking similarities were observed in children’s early written syntactic complexity. These results suggest that children very early on reflect language-specific trends in the composition of their texts. However, more refined measures may be needed to detect language-specific interplays of morphosyntax and discourse pragmatics (e.g., Aparici, Perera, Rosado, & Tolchinsky, 2000; Berman, 2004; Jisa, Reilly, Rosado, & Verhoeven, 2002; Ravid, van Hell, Rosado, & Zamora, 2002; Salas, 2010; Verhoeven et al., 2002).

7.2 Implications for Developmental Models of Writing

Studies in this thesis provided longitudinal and cross-linguistic evidence of the validity of key components of the writing process—word-level/transcription and text generation. The word-level dimension shares many features with an *accuracy* construct identified in recent studies concerned with the microstructure of early text writing. This accuracy/word-level factor also included spelling skill and capitalization components across studies (Puranik et al., 2008; Wagner et al., 2011). A *productivity* dimension was reported by these studies, all of which included a number-of-words component, which has moreover been regarded as a prime indicator of text generation skills (Puranik & AlOtaiba, 2012; Puranik et al., 2008; Wagner et al., 2011). In this sense, the thesis also extends models of writing development based on writing by English-speaking participants—i.e., the simple view (Juel et al., 1986; Juel, 1988), the developmental constraints theory (Berninger et al., 1991; 1992), and the "triangle" models (Berninger & Amtmann, 2003; Berninger & Winn, 2006)—to more consistent orthographies, in terms of the dimensionality of early writing.

Additionally, the early maturity of a productivity or text generation factor concurs with Bereiter and Scardamalia's (1987) *knowledge transforming* model of writing. Indeed, around six months after the outset of formal literacy instruction most children are able to produce short pieces of discourse, without the support of an interlocutor, which has been argued to be one of the primary challenges to beginner writers (Bereiter & Scardamalia, 1987). Children were only given a simple prompt (to put in writing what they had done the day before) and the vast majority produced intelligible, connected text, without any further interventions on the part of the administrator. The fact that the task was carried out in small groups (and not individually) further substantiates the claim about children's awareness that writing is a task that does not involve an immediate interlocutor. This is a non-trivial finding, in that it implies the appropriation of a fundamental aspect of the written modality: the ability to produce sustained discourse without the assistance of a direct audience. It also contradicts claims about early writing "style", by which children are assumed to transfer task schemas from oral language (Bereiter & Scardamalia, 1987). Several aspects of spoken language interactions that do not apply to writing (e.g., turn-taking, hesitations, interjections, etc.), which school-aged children have certainly mastered do not surface in their early written productions (Perera, 1984). Importantly, the absence of an "oral language style" was attested in a writing task (report of recent past events) that does not belong to a particular discourse genre from which children could have been drawing.

The solid findings for the transcription and text-generation components contrast sharply with the longitudinal instability of the text-level measures throughout the study. However, such instability is consistent with the view that some writing components develop prior to others (e.g., Berninger & Swanson, 1994; McCutchen, 2000).

This thesis also provided support for the levels of language framework (Abbott, et al., 2010; Whitaker, et al., 1994). First, because children across language groups showed intraindividual variation at the word and text levels; and second, within-level relationships (e.g., word-word) were stronger than relationships across levels of language (word-text). However, the pattern of relationships between the various text measures seems to fit best within the framework of *linguistic literacy* (Ravid & Tolchinsky, 2002). From this perspective, the disconnection between word- and text-level performance reflects the distinction between writing as notation, that is, writing understood as "the representational system that is used in the written modality" (p.18); and writing as discourse style, which "involves the variety of genres appropriate for 'language in writing' (...), each with its typical thematic content, global structure, and linguistic features" (p. 428). Early writing development may be characterized, in this sense, by rapid gains in writing as notation, especially from the outset of formal schooling. In contrast, writing as discourse style entails command over a large repertoire of (linguistic) forms and structures, as well as conventions, together with their respective contexts of use (Berman, 2008). Therefore, its developmental route is considerably longer and is possibly affected by a larger range of factors. This account is a closer match to the pattern of results reported in these studies, by which notational aspects of writing constituted reliable constructs and progress took place at a fast and steady rate. Conversely, text-level aspects of writing, arguably more sensitive to discourse knowledge and rhetorical style, showed little or no variation over time and were, in general, immature and still unreliable. What is more, rooted, as it is, in social practices and conventions, writing as a discourse style should be unaffected by variations in orthographic consistency, thus accommodating the lack of cross-linguistic findings at this level of performance.

Some cross-linguistic (i.e., cross-orthographic) findings in this thesis cannot be accommodated by models of writing development claiming a strong dependency of high-level writing components upon low-level components (e.g., Berninger & Amtmann, 2003; Berninger et al., 1991; Juel et al., 1986). The lesser demand in the transcription component for Spanish children should have led to their outperforming their English peers in writing achievements beyond the word level. This conclusion is further substantiated by the finding that spelling exerted a small and short-lived influence in the expression of text-level features, particularly on the amount of text produced. One possibility for the short-termed supporting role of spelling to text length or productivity may be the stabilization of productivity as a construct and, thus, the emergence of powerful autoregressor effects, which relegate spelling to a secondary or redundant role. This is consistent with studies showing that spelling does not directly impact on text length beyond the elementary grades (e.g., Graham et al., 1997; Lerkkanen, et al., 2004), as well as with research in the reading domain showing that exogenous predictors only account for variance in the very early developmental stages: From as early as the end of Year 1, powerful autoregressive effects preclude determining the influence of other individual variables, while suggesting a high degree of stability of the outcome measure (Caravolas et al., 2013).

The finding that spelling did not support the expression of text level performance moreover conflicts with studies that have reported a prolonged influence of spelling over text generation skills (e.g., Berninger & Swanson, 1994; Berninger et al., 2002; Connelly et al., 2012; Wagner et al., 2011). One explanation for the discrepancies across studies is related to the possible confound between the predictor (spelling) and the outcome variable, in those studies where the text-level assessment did not completely remove the direct or indirect effect that spelling might have in the final score. A direct effect may be found in holistic assessments of writing performance in which part of the score is derived from the evaluation of writing mechanics. An indirect effect of spelling in the assessment of text-level writing may occur when raters assess texts that have not been corrected for spelling and handwriting and may therefore bias their evaluations (MacArthur & Graham 1987). A second explanation is related to the possibility that spelling may “return” as an explanatory factor of text-level writing once children realize that their invented spellings are not sufficient to communicate effectively. A third explanation alludes to affective/motivational factors: Children’s growing concerns for spelling may result in negative self-perceptions, thus lowering their motivation to write and to produce longer and high-quality texts (Hayes, 2011; 2012).

A dominant claim in current thinking on text composition states that writing processes compete for the available cognitive resources, hence the need to automatize some of them. Particularly in the realm of writing development, transcription processes have long been suggested as the ideal candidate for automatization (Bourdin & Fayol, 1994, 2002; Juel et al., 1986). Nevertheless, other elements of the writing process may also need to achieve a relative degree of automatization to lead to successful writing. An attractive candidate for relative automatization is writing or task schemas, whose role in the writing process has been recently upgraded (Hayes, 2012). Being able to retrieve from long-term memory a ready-made writing “template” may arguably make the writing process more efficient by lessening the demand for online planning, as well as ensuring that task and genre-specific demands are met. There is some evidence that knowledge and familiarity with macrostructural aspects of discourse construction predicts later writing quality (Griffin et al., 2004; Juel et al., 1986). Finally, given that they may be obtained by instruction, practice, or a combination of both, task schemas constitute an interesting basis for intervention studies.

7.3 Methodological Issues in the Assessment of Early Writing Development

The studies presented in this thesis revealed the critical importance of making distinctions among writing outcomes. Only by establishing clear boundaries between different text features and skills was it possible to determine which constructs were stable and mature and to distinguish them from those that were in their infancy and, thus, unreliable. Also, this methodological approach allowed detecting the selective impact of low-level writing skills on other aspects of text composition. In this sense, the present studies highlight the need for carefully avoiding confounds between dependent and independent measures. On the one hand, it is essential to rule out the influence of handwriting and spelling skill on the scoring of other text features to disentangle the nature of the relationship between low-level and high-level writing skills. On the other hand, proportion scores must be used to control for the effect that text length could have on other measures and ratings.

In sum, many of the reported findings require that quantitative, multifeature approaches be taken to investigate and assess early writing development. Nevertheless, the difficulties in all the studies included in the thesis to draw any conclusions with regard to text level features might be indicating that a fully quantitative approach may not be suitable for all text features and skills at any developmental stage. Evidence was provided that some text-level measures were still showing near-floor effects even towards the middle and end of the study period. These measures were often those showing longitudinal instability (e.g., punctuation, complex syntax), suggesting that some aspects of text level performance may not yet be consolidated, whether they are considered individually or within an aggregated or regression-based score. The few studies which have attempted the operationalization of text-level aspects of early written composition have used qualitative ratings for their assessment (Puranik et al., 2008; Wagner et al., 2011). More research is needed to identify the most valid and sensitive measures at different stages of writing development.

7.4 Limitations

The studies included in this thesis presented a number of limitations. Some aspects of relevance to literacy development were not available for all participants.

This involved full SES information as well as information on handedness. Although the large sample sizes should minimize the effect of these population variables, the extent of their impact on current results cannot be determined with precision. Similarly, the aim to obtain an unselected sample, representative of the population in English and Spanish classrooms may have resulted in the lack of control over different sub-populations within the samples. A study detecting different writer profiles would be a fruitful endeavour.

The writing task may have been slightly inadequate to gauge more proficient children's full command of writing. The text that children were asked to produce did not have a clear correlate in the real world, in the sense that children could not refer to previous writing experience or exposure to a similar type of text. This aspect of the task may have arguably limited the extent to which drawing on such knowledge could enhance children's text-level writing skills. In addition, the time constraints for text production may have hindered the range of abilities that children's could have demonstrated without such constraints. It is thus highly likely that the present results are only applicable to this type of text. Future studies should strive to elicit a more varied range of text types, obtained both with and without time limits.

Also in relation to the nature of the writing task, certain aspects of its procedure might have had an impact on the results obtained. On the one hand, children were asked to speak about their recent past events prior to writing about them, with the objective to activate ideas and facilitate content retrieval. However, such a procedure may have had undesired effects, such as biasing results by making children focus solely or at least primarily in transcription aspects. In this sense, it has been suggested that when oral production precedes written production, the quality of the latter is improved well beyond transcription aspects. From this viewpoint, the spoken versions of the texts have been argued to constitute a pre-planning phase (Boscolo, 1990). Nevertheless, little is known about the specific chain of effects that a speaking-first procedure may have on the resulting written product.

On the other hand, children were asked to "interpret" their written productions, when told to "read" back to the administrator the text produced. This is common practice in studies with very young children who are still struggling with their spelling and handwriting skills (e.g., Berninger et al., 1992; Hayes & Berninger, 2009).

Although these “readings” were never the only—nor the most important—source of information for the transcriber, it is hard to determine whether their existence may have biased the ultimate interpretation of the texts.

A number of predictor measures that would have been of interest were not included in these studies. No task measured children’s oral language production skills. Although children’s metalinguistic and receptive oral language skills were evaluated, an oral production task would have been useful to control for children’s text generation skills without the demands of transcription (e.g., Hayes & Berninger, 2009; Graham, 1990), as well as to determine the predictive power of oral language production in text composition. A study of children’s productive language skills and their relation with writing outcomes is underway. Similarly, a measure of children’s knowledge of writing and/or print awareness may have been useful to determine the role of higher-level writing knowledge in early writing development. Finally, verbal short-term memory, but not working memory measures, were included in the thesis. Although working-memory has been argued to constrain writing development starting in the intermediate grades (Berninger & Swanson, 1994; Berninger & Winn, 2006), testing its predictive role particularly in Spanish would have been of interest, especially in light of the lack of cross-linguistic differences as a function of orthographic complexity and of the lower proportion of variance explained in models of Spanish, in contrast to those of English.

Some improvements in the strategy of data analyses could have achieved a more detailed picture of the developmental patterns both within as across language groups. Analysis of growth trajectories would have allowed expanding present results on the achievement levels of children learning to write in English as opposed to Spanish, to include an assessment of rate and speed of development. In addition, the exploratory nature of the compositionality of early writing across languages could inform theory-driven confirmatory factor analyses, to obtain more conclusive results.

Finally, in view of the subtle tendencies to an orthographic consistency/language effect towards the end of the study period, a study covering a longer time span would have helped determine whether such tendencies became more pronounced as both groups further acquired higher level writing skills. Moreover, the intervals between time points were relatively short and may have thus been best suited for capturing the fast development of word-level writing skills, rather than text-levels skill, which may arguably need longer periods for marked differences to emerge.

7.5 Concluding Remarks

This work has extended single-word spelling research by demonstrating that several trends applicable to spelling development also hold within the text writing domain. The studies here reported indicated that, when writing is considered as a notation system, there is continuity from single-word to text writing tasks, since in both of them spelling skill is equally sensitive to variations in orthographic consistency, as well as it is driven by the same set of precursor skills. Less is known about the early links between writing as notation and writing as discourse style, for which only poor explanatory models have been proposed to date. The skills that are usually considered to contribute to writing quality are still unstable constructs at these early developmental stages. Therefore, future research should strive to identify the foundational abilities that drive development in writing during this foundational period, as well as the extent to which they may be subject to variations as a function of orthography- and language-specific properties.

References

Abbott, R. D., & Berninger, V. W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary-and intermediate-grade writers. *Journal of Educational Psychology*, *85*(3), 478-508.

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.

Abercrombie, D . (1967). *Elements of phonetics*. Edinburgh: Edinburgh Univ- Press.

Adams, M. J. (1990). Beginning to read: Learning and thinking about print.*Beginning to read: Learning and thinking about print*.

Alamargot, D., & Chanquoy, L. (2001). *Through the Models of Writing: With Commentaries by Ronald T. Kellogg & John R. Hayes* (Vol. 9). Kluwer Academic Pub.

Alegría, J., Marín, J., Carrillo, S., & Mousty, P. (2003). Les premiers pas dans l'acquisition de l'orthographe en fonction du caractère profond ou superficiel du système alphabétique: comparaison entre le français et l'espagnol [First steps in the aquicistion of orthography as a function of whether the alphabetic system is deep or shallow: Comparison between French and Spanish]. *L'apprentissage de la lecture: Perspective comparative interlangues*, 51-67.

Aparici, M., Perera, J., Rosado, E., & Tolchinsky, L. (2000). Developing literacy across genres, modalities, and languages. (International Literacy Project, Working papers, 3.) Barcelona: Institute of Educational Sciences. *University of Barcelona*.

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.

Babayiğit, S., & Stainthorp, R. (2011). Modeling the relationships between cognitive–linguistic skills and literacy skills: New insights from a transparent orthography. *Journal of Educational Psychology, 103*(1), 169-189.

Baddeley, A. D. (1986). *Working-memory.* New York: Oxford University Press.

Bedore, L. M., & Leonard, L. B. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment.*Journal of Speech, Language and Hearing Research*, *44*(4), 905.

Bedore, L. M., & Leonard, L. B. (2005). Verb inflections and noun phrase morphology in the spontaneous speech of Spanish-speaking children with specific language impairment. *Applied Psycholinguistics*, *26*(02), 195-225.

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. (2011). 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 Associates.

Berko, J. (1958). *The child's learning of English morphology*. Doctoral dissertation, Radcliffe College.

Berman, R. & Nir, B. (2009). Cognitive and linguisticfactors in evaluating text quality: Global versus local? In V. Evans & S. Pourcel (eds.). *New directions in cognitive linguistics.* Amsterdam:John Benjamins, p. 421-440.

Berman, R. (1997). Preschool knowledge of language: What five-year olds know about language structure and language use. In C. Pontecorvo (ed.). *Writing development: An interdisciplinary view*, 61-76.

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

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., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study* (Vol. 1). Psychology Press.

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. (2009). Highlights of programmatic, interdisciplinary research on writing. *Learning Disabilities Research & Practice*, *24*(2), 69-80.

Berninger, V. W., & Amtmann, D. (2003). Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: Research into practice.

Berninger, V. W., & Fuller, F. (1993). Gender differences in orthographic, verbal, and compositional fluency: Implications for assessing writing disabilities in primary grade children. *Journal of School Psychology*, *30*(4), 363-382.

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. W., Abbott, R. D., Augsburger, A., & Garcia, N. (2009). Comparison of pen and keyboard transcription modes in children with and without learning disabilities. *Learning Disability Quarterly*, 123-141.

Berninger, V. W., Cartwright, A. C., Yates, C. M., Swanson, H. L., & Abbott, R. D. (1994). Developmental skills related to writing and reading acquisition in the intermediate grades. *Reading and Writing*, *6*(2), 161-196.

Berninger, V. W., Fuller, F., & Whitaker, D. (1996). A process model of writing development across the life span. *Educational Psychology Review*, *8*(3), 193-218.

Berninger, V. W., Mizokawa, D. T., & Bragg, R. (1991). Scientific practitioner: Theory-based diagnosis and remediation of writing disabilities. *Journal of School Psychology*, *29*(1), 57-79.

Berninger, V. W., Vaughan, K., Abbott, R. D., Brooks, A., Abbott, S. P., Rogan, L., ... & Graham, S. (1998). Early intervention for spelling problems: Teaching functional spelling units of varying size with a multiple-connections framework. *Journal of Educational Psychology*, *90*(4), 587-605.

Berninger, V., & Winn, W. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. *Handbook of writing research*, 96-114.

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.

Biber, D. (1995). *Dimensions of register variation: A cross-linguistic comparison*. Cambridge University Press.

Bishop, D. V. (2003). *Test for reception of grammar: TROG-2 version 2*. Pearson Assessment.

Bishop, D. V., & Clarkson, B. (2003). Written Language as a Window in to Residual Language Deficits: A Study of Children With Persistent and Residual Speech and Language Impairments. *Cortex*, *39*(2), 215-237.

Bishop, D. V., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different?. *Psychological bulletin*, *130*(6), 858.

Blom, E., Jong, J., Orgassa, A., Baker, A., & Weerman, F. (2013). Verb inflection in monolingual Dutch and sequential bilingual Turkish–Dutch children with and without SLI. *International Journal of Language & Communication Disorders*.

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.

Boscolo, P. (1990). The construction of expository text. *First Language*, *10*(30), 217-230.

Bosman, A. M., & Van Orden, G. C. (1997). Why spelling is more difficult than reading. *Learning to spell: Research, theory, and practice across languages*,*10*, 173-194.

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.

Bourdin, B., & Fayol, M. (2002). Even in adults, written production is still more costly than oral production. *International Journal of Psychology*, *37*(4), 219-227.

Britton, J., T. Burgess, N. Martin, A. McLeod & H. Rosen. (1975). *The development* of *writing abilities.* London: Macmillan.

Bruck, M., & Treiman, R. (1990). Phonological awareness and spelling in normal children and dyslexics: The case of initial consonant clusters. *Journal of Experimental Child Psychology*, *50*(1), 156-178.

Bruck, M., Genesee, F., & Caravolas, M. (1997). A cross-linguistic study of early literacy acquisition. *Foundations of reading acquisition and dyslexia: Implications for early intervention*, 145-162.

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

Byrne, B. J. (1998). *The foundation of literacy: The child's acquisition of the alphabetic principle*. Hove, England:Psychology Press.

Cain, K., & Oakhill, J. (2011). Matthew Effects in Young Readers Reading Comprehension and Reading Experience Aid Vocabulary Development. *Journal of learning disabilities, 44*(5), 431-443.

Calkins, L. M. (1986). *The art of teaching writing*. Heinemann Educational Books Inc., 70 Court St., Portsmouth, NH 03801.

Caravolas, M. (2006). Refining the psycholinguistic grain size theory: effects of phonotactics and word formation on the availability of phonemes to preliterate children. *Developmental Science*, *9*(5), 445-447.

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., & Landerl, K. (2010). The influences of syllable structure and reading ability on the development of phoneme awareness: A longitudinal, cross-linguistic study. *Scientific Studies of Reading*, *14*(5), 464-484.

Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and language*, *45*(4), 751-774.

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.*

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

Caravolas, M., Volín, J., & Hulme, C. (2005). Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: Evidence from Czech and English children. *Journal of experimental child psychology*, *92*(2), 107-139.

Cardoso-Martins, C., Corrêa, M. F., Lemos, L. S., & Napoleão, R. F. (2006). Is There a Syllabic Stage in Spelling Development? Evidence from Portuguese-Speaking Children. *Journal of Educational Psychology*, *98*(3), 628-641.

Carlisle, J. F. (1996). An exploratory study of morphological errors in children's written stories. *Reading and Writing*, *8*(1), 61-72.

Carney, E. (1994). *Survey of English Spelling*. London:Routledge.

Carrillo, M. S., Alegría, J., & Marín, J. (2012). On the acquisition of some basic word spelling mechanisms in a deep (French) and a shallow (Spanish) system. *Reading and Writing*, 1-21.

Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology*, *89*(4), 631.

Charney, D. (1984). The validity of using holistic scoring to evaluate writing: A critical overview. *Research in the Teaching of English*, *18*(1), 65-81.

Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in Writing Generating Text in L1 and L2. *Written communication*, *18*(1), 80-98.

Chomsky, C. (1970). Reading, writing, and phonology. *Harvard Educational Review*, *40*(2), 287-309.

Chomsky, C. (1976). Approaching reading through invented spelling.

Chomsky, N. (2002). *Syntactic structures*. de Gruyter Mouton.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. Addison-Wesley Educational Publishers.

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, CliC–UB.

Clark, E. V. (1995). *The lexicon in acquisition* (Vol. 65). Cambridge University Press.

Connelly, V., Dockrell, J. E., Walter, K., & Critten, S. (2012). Predicting the quality of composition and written language bursts from oral language, spelling, and handwriting skills in children with and without specific language impairment. *Written Communication*, *29*(3), 278-302.

Correa, J., & Dockrell, J. E. (2007). Unconventional word segmentation in Brazilian children's early text production. *Reading and Writing*, *20*(8), 815-831.

Cossu, G., Shankweiler, D., Liberman, I. Y., Katz, L., & Tola, G. (1988). Awareness of phonological segments and reading ability in Italian children.*Applied Psycholinguistics*, *9*(1), 1-16.

Cragg, L., & Nation, K. (2006). Exploring written narrative in children with poor reading comprehension. *Educational Psychology*, *26*(1), 55-72.

Cutler, A. (2012). Native Listening: Language Experience and the Recognition of Spoken Words. Cambridge, MA: MIT Press.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language*, *2*(3), 133-142.

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: Sunjin Printing Co.

De La Paz, S., & Graham, S. (1995). Dictation: Applications to writing for students with learning disabilities. *Advances in learning and behavioral disorders*, *9*, 227-247.

de Manrique, A. M. B., & Diuk, B. (2003). La escritura de textos en niños pequeños: conocimiento ortográfico y producción textual [Composition in young children: Spelling and text production]. *Cultura y Educación*, *15*(1), 17-27.

de Manrique, A. M. B., & Signorini, A. (1998). Emergent writing forms in Spanish. *Reading and Writing*, *10*(6), 499-517.

de Vega Rodríguez, M., Alvarez, C., & Valiña, M. F. C. (1992). Estudio estadístico de la ortografía castellana (1): La frecuencia silabica. *Cognitiva*,*4*(1), 75-106.

Deacon, S. H., & Kirby, J. R. (2004). Morphological awareness: Just "more phonological"? The roles of morphological and phonological awareness in reading development. *Applied psycholinguistics*, *25*(02), 223-238.

Defior, S., Alegría, J., Titos, R., & Martos, F. (2008). Using morphology when spelling in a shallow orthographic system: The case of Spanish. *Cognitive Development*, *23*(1), 204-215.

Diccionario de la Lengua Española [Dictionary of the Spanish Language]. (2001). Madrid:Espasa Calpe.

Diccionario panhispánico de dudas. Real Academia Española, 2005.

Dockrell, J. E., Lindsay, G., Connelly, V., & Mackie, C. (2007). Constraints in the production of written text in children with specific language impairments.*Exceptional Children*, *73*(2), 147-164.

Dressler, W. (2005). Morphological typology and first language acquisition: some mutual challenges. *Morphology and linguistic typology*, 7-20.

Dressler, W. (2010). A typological approach to first language acquisition.*Language acquisition across linguistic and cognitive systems*, 109-124.

Duncan, L. G., Colé, P., Seymour, P. H., & Magnan, A. (2006). Differing sequences of metaphonological development in French and English. *Journal of Child Language*, *33*(02), 369-399.

Dunsmuir, S., & Blatchford, P. (2004). Predictors of writing competence in 4-to 7-year-old children. *British journal of educational psychology*, *74*(3), 461-483.

Durgunoğlu, A. Y. (2006). Learning to read in Turkish. *Developmental Science*,*9*(5), 437-439.

Ehri, L. C., & Rosenthal, J. (2010). Spellings of words: A neglected facilitator of vocabulary learning. In D. Aram & O. Korat (Eds.). *Literacy Development and Enhancement Across Orthographies and Cultures,* Springer US, p. 137-152.

Fattal, I., Friedmann, N., & Fattal-Valevski, A. (2011). The crucial role of thiamine in the development of syntax and lexical retrieval: a study of infantile thiamine deficiency. *Brain*, *134*(6), 1720-1739.

Ferreiro, E. (2000). Reading and writing in a changing world. *Publishing research quarterly*, *16*(3), 53-61.

Ferreiro, E. (2009). The Transformation of Children's Knowledge of Language Units During Beginning and Initial Literacy. In *Changing Literacies for Changing Times: An Historical Perspective on the Future of Reading Research, Public Policy, and Classroom Practices*, 61.

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

Ferreiro, E., & Pontecorvo, C. (2002). Word segmentation in early written narratives. *Language and Education*, *16*(1), 1-17.

Ferreiro, E., & Teberosky, A. (1979). Los *sistemas de escritura en el desarrollo del niño.*[Literacy before schooling]. México: Siglo XXI.

Ferreiro, E., Pontecorvo, C., Ribeiro-Moreira, N., & García-Hidalgo, I. (1996). *Caperucita roja aprende a escribir: estudios psicolingüísticos comparativos en tres lenguas.* [Little Red Riding Hood learns to write: Comparative psycholinguistic studies in three languages] Barcelona:Gedisa.

Field, A. (2009). *Discovering statistics using SPSS*. Sage Publications Limited.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College composition and communication*, *32*(4), 365-387.

Frith, U. (1985). Beneath the Surface of Developmental Dyslexia. *Surface Dyslexia: Neuropsychological and Cognitive Studies of Phonological Reading*, 301.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: a multilingual comparison. *Journal of experimental psychology. Human perception and performance*, *13*(1), 104-115.

Gentry, J. R. (1982). An Analysis of Developmental Spelling in" GNYS AT WRK". *The reading teacher*, *36*(2), 192-200.

Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *Journal of Educational Psychology*, *100*(3), 566.

Gibson, E. J., & Levin, H. (1975). *The psychology of reading* (pp. 454-465). Cambridge, MA: MIT press.

Giegerich, H. J. (1992). *English phonology: An introduction*. Cambridge University Press.

Gillis, S., & Ravid, D. (2000). Effects of phonology and morphology in children’s orthographic systems: a cross-linguistic study of Hebrew and Dutch. In *The proceedings of the 30th annual child language research forum* (pp. 203-210).

Gillis, S., & Ravid, D. (2006). Typological effects on spelling development: A crosslinguistic study of Hebrew and Dutch. *Journal of child language*, *33*(03), 621-659.

Graham, S. (1990). The role of production factors in learning disabled students' compositions. *Journal of Educational Psychology*, *82*(4), 781.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, *99*(3), 445.

Graham, S., & Weintraub, N. (1996). A review of handwriting research: Progress and prospects from 1980 to 1994. *Educational Psychology Review*,*8*(1), 7-87.

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.

Graham, S., Harris, K. R., & Chorzempa, B. F. (2002). Contribution of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology*, *94*(4), 669-686.

Graves, D. H. (1975). An examination of the writing processes of seven year old children. *Research in the Teaching of English*, *9*(3), 227-241.

Griffin, T. M., Hemphill, L., Camp, L., & Wolf, D. P. (2004). Oral discourse in the preschool years and later literacy skills. *First Language*, *24*(2), 123-147.

Guan, C. Q., Ye, F., Wagner, R. K., & Meng, W. (2012). Developmental and individual differences in Chinese writing. *Reading and Writing,* 1-26.

Guirao, M., & de Manrique, A. M. B. (1975). Identification of argentine Spanish vowels. *Journal of Psycholinguistic Research*, *4*(1), 17-25.

Halliday, M.A.K. & Hassan, R. (1985), *Language: Context and Text*. Burwood,

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological review*, *111*(3), 662.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*(4), 408-420.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Randall (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1-27). Mahwah, NJ: Erlbaum.

Hayes, J. R. (2011). Kinds of knowledge-telling: Modeling early writing development. *Journal of Writing Research*, *3*(2), 73-92.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*,*29*(3), 369-388.

Hayes, J. R., & Berninger, V. W. (2010). 13 Relationships between idea generation and transcription. *Traditions of writing research*, 166.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. *Cognitive processes in writing*, 3-30.

Hayes, J. R., & Flower, L. S. (1986). Writing research and the writer. *American psychologist*, *41*(10), 1106.

Hayes, J. R., & Nash, J. G. (1996). On the nature of planning in writing. In C. M. Levy & S. Ransdell (Eds.), The science of writing: theories, methods, individual differences and applications (pp. 29–55). Mahwah, NJ: Lawrence Erlbaum Associates.

Hayes, J. R., Flower, L., Schriver, K. A., Stratman, J., & Carey, L. (1987). Cognitive processes in revision. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics:Vol. 2. Reading writing and language processing* (pp. 176-241). Cambridge:Cambridge University Press.

Henderson, E. H., & Beers, J. (Eds.). (1980). Developmental and cognitive aspects of learning to spell: A reflection of word knowledge. Newark, DE: International Reading Association.

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). Hillsdale, NJ:Lawrence Erlbaum.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: guidelines for determining model fit. *Articles*, 2.

Hooper, S. R., Roberts, J. E., Nelson, L., Zeisel, S., & Kasambira Fannin, D. (2010). Preschool predictors of narrative writing skills in elementary school children. *School Psychology Quarterly*, *25*(1), 1-12.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and writing*, *2*(2), 127-160.

Hulme, C., Maughan, S., & Brown, G. D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30*(6), 685-701.

Jiménez González, J. E., & García, C. R. H. (1995). Effects of word linguistic properties on phonological awareness in Spanish children. *Journal of Educational Psychology*, *87*(2), 193-201.

Jisa, H., Reilly, J. S., Verhoeven, L., Baruch, E., & Rosado, E. (2002). Passive voice constructions in written texts:: A cross-linguistic developmental study. *Written Language & Literacy*, *5*(2), 163-181.

Jones, D., & Christensen, C. A. (1999). Relationship between automaticity in handwriting and students' ability to generate written text. *Journal of educational psychology*, *91*(1), 44.

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.

Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in cognitive sciences*, *3*(9), 323-328.

Justicia, F., Santiago, J., Palma, A., Huertas, D., & Gutierrez, N. (1996). La frecuencia silábica del español escrito por niños: estudio estadístico [Syllable frequency: A statistical study of written productions by Spanish children]. *Cognitiva*, *8*(2), 131-168.

Kail, R., Hall, L. K., & Caskey, B. J. (1999). Processing speed, exposure to print, and naming speed. *Applied Psycholinguistics*, *20*(2), 303-314.

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition*, *23*(2), 95-147.

Karmiloff-Smith, A. (1992). *Beyond Modularity: A developmental Perspective on Cognitive Science*. The MIT Press.

Katz, L., & Frost, R. (1992). The Reading Process is Different for Different Orthographies: The Orthographic Depth Hypothesis. *Advances in Psychology*,*94*, 67-84.

Kaufer, D. S., Hayes, J. R., & Flower, L. S. (1986). Composing written sentences. *Research in the Teaching of English, 20*(2), 121-140.

Kellogg, R. T. (1996). A Model of Working Memory in Writing. *The science of writing: theories, methods, individual differences, and applications*, 57.

Kellogg, R. T. (2001). Competition for Working Memory among Writing Processes. *The American Journal of Psychology*, 175-191.

Kessler, B. (2003). Is English spelling chaotic? Misconceptions concerning its irregularity. *Reading Psychology*, *24*(3-4), 267-289.

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

Kim, Y. S., Wagner, R. K., & Foster, E. (2011). Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of first-grade readers. *Scientific Studies of Reading, 15*(4), 338-362.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. The Guilford Press.

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology*, *100*(1), 150.

Lee, Y. W., Gentile, C., & Kantor, R. (2010). Toward automated multi-trait scoring of essays: Investigating links among holistic, analytic, and text feature scores. *Applied Linguistics, 31*(3), 391-417.

Leonard, L. (1998). Children with specific language impairment. Cambridge, MA: MIT Press.

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.

Lervåg, A., & Hulme, C. (2009). Rapid automatized naming (RAN) taps a mechanism that places constraints on the development of early reading fluency. *Psychological Science*, *20*(8), 1040-1048.

Llauradó, A. & Tolchinsky, L. (in press). Growth of text-embedded lexicon in Catalan: from childhood to adolescence. *First Language.*

Lust, B., Flynn, S., & Foley, C. (1996). What children know about what they say: Elicited imitation as a research method for assessing children's syntax.*Methods for assessing children's syntax*, 55-76.

MacArthur, C. A., & Graham, S. (1987). Learning Disabled Students' Composing Under Three Methods of Text Production Handwriting, Word Processing, and Dictation. *The Journal of Special Education*, *21*(3), 22-42.

Mackie, C. J., Dockrell, J., & Lindsay, G. (2012). An evaluation of the written texts of children with SLI: the contributions of oral language, reading and phonological short-term memory. *Reading and Writing*, 1-24.

Mackie, C., & Dockrell, J. E. (2004). The nature of written language deficits in children with SLI. *Journal of Speech, Language and Hearing Research*, *47*(6), 1469.

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: Palgrave Macmillan.

Manrique, A. M. B., & Signorini, A. (1994). Phonological awareness, spelling and reading abilities in Spanish-speaking children. *British Journal of Educational Psychology*, *64*(3), 429-439.

Martin, J. R. (2009). Genre and language learning: A social semiotic perspective. *Linguistics and Education*, *20*(1), 10-21.

McCutchen, D. (1986). Domain knowledge and linguistic knowledge in the development of writing ability. *Journal of Memory and Language*, *25*(4), 431-444.

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.

McCutchen, D., Covill, A., Hoyne, S. H., & Mildes, K. (1994). Individual differences in writing: Implications of translating fluency. *Journal of Educational Psychology*, *86*(2), 256-266.

McDougall, S., Hulme, C., Ellis, A., & Monk, A. (1994). Learning to read: The role of short-term memory and phonological skills. *Journal of Experimental Child Psychology, 58*(1), 112-133.

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.

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication*, *27*(1), 57-86.

Mehta, P. D., Foorman, B. R., Branum-Martin, L., & Taylor, W. P. (2005). Literacy as a unidimensional multilevel construct: Validation, sources of influence, and implications in a longitudinal study in grades 1 to 4. *Scientific Studies of Reading*, *9*(2), 85-116.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Developmental Psychology, 40*(5), 665-681.

Muthen, L. K., & Muthen, B. O. (2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Author.

Nystrand, M. (2006). The social and historical context for writing research.*Handbook of writing research*, 11-27.

Olinghouse, N. G. (2008). Student-and instruction-level predictors of narrative writing in third-grade students. *Reading and Writing*, *21*(1-2), 3-26.

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.

Olson, D. R. (1994). *The world on paper: The conceptual and cognitive implications of reading and writing*. Cambridge:Cambridge University Press.

Olson, D. R. (1996). Towards a psychology of literacy: On the relations between speech and writing. *Cognition*, *60*(1), 83-104.

Pacton, S., & Deacon, S. H. (2008). The timing and mechanisms of children's use of morphological information in spelling: A review of evidence from English and French. *Cognitive Development*, *23*(3), 339-359.

Pajares, F., & Valiante, G. (2006). Self-efficacy beliefs and motivation in writing development. *Handbook of writing research*, 158-170.

Perera, K. (1984). Children's writing and reading: Analyzing classroom language. Oxford: Basil Blackwell.

Perfetti, C. A., & McCutchen, D. (1987). Schooled language competence: Linguistic abilities in reading and writing. *Advances in applied psycholinguistics*,*2*, 105-141.

Pollo, T. C., Kessler, B., & Treiman, R. (2005). Vowels, syllables, and letter names: Differences between young children's spelling in English and Portuguese. *Journal of Experimental Child Psychology*, *92*(2), 161-181.

Pollo, T. C., Treiman, R., & Kessler, B. (2008). Three perspectives on spelling development. *Single-word reading: Behavioral and biological perspectives*, 175-189.

Pons, F., & Bosch, L. (2010). Stress Pattern Preference in Spanish-Learning Infants: The Role of Syllable Weight. *Infancy*, *15*(3), 223-245.

Pontecorvo, C., & Zucchermaglio, C. (1988). Modes of differentiation in children's writing construction. *European Journal of Psychology of Education*,*3*(4), 371-384.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 717-731.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models.*Behavior research methods*, *40*(3), 879-891.

Puranik, C. S., & AlOtaiba, 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., & Lonigan, C. J. (2011). From scribbles to scrabble: Preschool children's developing knowledge of written language. *Reading and writing*, *24*(5), 567-589.

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.

Puranik, C.S. (2006). *Expository writing skills in elementary school children from Third through sixth grades and contributions of short-term and working memory*. Unpublished doctoral dissertation. University of Florida.

Quilis, A. (2010). *Principios de fonología y fonética españolas*. Madrid:Arco.

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.

Ravid, D., & Tolchinsky, L. (2002). Investigating literacy development and language acquisition. Reply to commentaries on 'Developing linguistic literacy: a comprehensive model'. *Journal of child language*, *29*(02), 489-494.

Ravid, D., van Hell, J. G., Rosado, E., & Zamora, A. (2002). Subject NP patterning in the development of text production: Speech and writing. *Written Language & Literacy*, *5*(1), 69-93.

Read, C. (1971). Pre-school children's knowledge of English phonology. *Harvard educational review*, *41*(1), 1-34.

Read, C. (1986). Children's creative spelling . London: Routledge & Kegan Paul.

Riedemann, H. (1996). Word-length distribution in English press texts. *Journal of Quantitative Linguistics, 3*(3), p. 265-271.

Rittle-Johnson, B., & Siegler, R. S. (1999). Learning to spell: Variability, choice, and change in children's strategy use. *Child Development*, *70*(2), 332-348.

Roach, P. (1982). On the distinction between 'stress-timed'and 'syllable-timed'languages. *Linguistic controversies*, 73-79.

Roach, P. (2004). British English: Received Pronunciation. *Journal of the International Phonetic Association*, *34*(2), 239-245.

Salas, N. (2010). *Developmental and discursive underpinnings in the expression of Spanish noun phrases.* Unpublished Masters' Dissertation. Linguistic Department, University of Barcelona, Spain.

Sánchez-Abchi, V., Borzone, A. M., & Diuk, B. (2007). La escritura de textos en niños pequeños: relación entre la transcripción y la composición [Text writing in small children: relationships between transcription and composition]. *Universitas psychologica*, (3), 559-569.

Sandoval, A. M., Toledano, D. T., Curto, N., & de la Torre, R. (2006). Inventario de frecuencias fonémicas y silábicas del castellano espontáneo y escrito. *IV Jornadas en Tecnologia del Habla, Zaragoza*.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507-514.

Scardamalia, M., Bereiter, C., & Goelman, H. (1982). The role of production factors in writing ability. *What writers know: The language, process, and structure of written discourse*, 173-210.

Scott, C. (2005). Learning to write. In H. W. Catts & A. G. Kamhi (Eds.)*, Language and reading disabilities* (2nd ed.). Needham Heights, MA: Allyn & Bacon, p. 233–273

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.

Shanahan, T. (1984). Nature of the Reading-Writing Relation: An Exploratory Multivariate Analysis. *Journal of Educational Psychology*, *76*(3), 466-77.

Shanahan, T., & Lomax, R. (1986). An Analysis and Comparison of Theoretical Models of the Reading-Writing Relationship. *Journal of Educational Psychology*,*78*(2), 116-23.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: the perils of overreliance on an" outlier" orthography. *Psychological bulletin*, *134*(4), 584.

Siegler, R. S. (1996). *Emerging minds: the process of change in children's thinking*. Oxford University Press, USA.

Silva, M.L., Sánchez-Abchi, V., & Borzone, A. (2010). Subordinated clauses usage and assessment of syntactic maturity: A comparison of oral and written retellings in beginning writers. *Journal of Writing Research, 2*(1), 47-64.

Slobin, D. I. (1996). Two ways to travel: Verbs of motion in English and Spanish. In *Grammatical constructions: Their form and meaning*, 195-219.

Slobin, D. I. (2004). The many ways to search for a frog. *Relating Events in Narrative. Vol*, *2*, 219-257.

Soanes, C., & Stevenson, A. (2002). *The Concise Oxford English Dictionary*. J. Pearsall (Ed.). Oxford: Oxford University Press.

Sprenger-Charolles, L., & Siegel, L. S. (1997). A longitudinal study of the effects of syllabic structure on the development of reading and spelling skills in French. *Applied Psycholinguistics*, *18*, 485-506.

Sulzby, E., Barnhart, J. & Hieshima, J. (1989). *Forms of writing and writing connections*.Needham Heights, MA: Allyn and Bacon, Tech. Report No. 20: 31–63, Center for the Study of Writing, University of California, Berkeley, CA.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms.*Language typology and syntactic description*, *3*, 57-149.

Tolchinsky Landsmann, L. (1996). Three Accounts of Literacy and the role of the environment. *Children's Early Text Construction. Hillsdale, NJ: Lawrence Erlbaum Associates*, 99-101.

Tolchinsky, L. (2003). *The Cradle of Culture and What Children Know About Writing and Numbers Before Being*. Psychology Press.

Tolchinsky, L. (2004). The nature and scope of later language development. In *Language development across childhood and adolescence*, 233-247.

Tolchinsky, L. (2006). The emergence of writing. *Handbook of writing research*, 83-95.

Tolchinsky, L. (2013). L'aprenentatge de l'escriptura i les seves dificultats [Learning to write and its difficulties]. In L. Andreu i Barrachina (ed.). *Trastorns d'aprenentatge de l'escriptura i de les matemàtiques* [Writing and math learning disorders]. Barcelona:UOC.

Tolchinsky, L., Bigas, M., & Barragan, C. (2012). Pedagogical practices in the teaching of early literacy in Spain: Voices from the classroom and from the official curricula. *Research Papers in Education*, *27*(1), 41-62.

Tolchinsky, L., Levin, I., Aram, D., & McBride-Chang, C. (2012). Building literacy in alphabetic, abjad and morphosyllabic systems. *Reading and Writing*,*25*(7), 1573-1598.

Torrance, M., & Galbraith, D. (2006). The processing demands of writing.*Handbook of writing research*, 67-80.

Traweek, D., & Berninger, V. W. (1997). Comparisons of beginning literacy programs: Alternative paths to the same learning outcome. *Learning Disability Quarterly*, *20*(2), 160-168.

Treiman, R. (1991). Children's spelling errors on syllable-initial consonant clusters. *Journal of Educational Psychology*, *83*(3), 346.

Treiman, R. (1993). *Beginning to spell*. New York: Oxford University Press.

Treiman, R., & Bourassa, D. C. (2000). The development of spelling skill.*Topics in language disorders*, *20*(3), 1-18.

Treiman, R., & Cassar, M. (1996). Effects of morphology on children's spelling of final consonant clusters. *Journal of Experimental Child Psychology*.

Treiman, R., & Cassar, M. (1997). Spelling acquisition in English. *Learning to spell: Research, theory, and practice across languages*, 61-80.

Treiman, R., Cohen, J., Mulqueeny, K., Kessler, B., & Schechtman, S. (2007). Young children's knowledge about printed names. *Child Development*, *78*(5), 1458-1471.

Ure, J. (1971). Lexical density and register differentiation. *Applications of linguistics*, 443-452.

Vaessen, A., Bertrand, D., Tóth, D., Csépe, V., Faísca, L., Reis, A., & Blomert, L. (2010). Cognitive development of fluent word reading does not qualitatively differ between transparent and opaque orthographies. *Journal of educational psychology*, *102*(4), 827.

van Dijk, T. (1985). Semantic discourse analysis. In T. Van Dijk (ed.), *Handbook of discourse analysis, vol. II* (pp. 103-136). London: Academic Press.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading*, *11*(1), 3-32.

Venezky, R. L. (1970). *The structure of English orthography* (Vol. 82). De Gruyter Mouton.

Vennemann, T. (1987). *Preference Laws for Syllable Structure: And the Explanation of Sound Change with Special Reference to German, Germanic, Italian, and Latin*. De Gruyter Mouton.

Verhoeven, L., Aparici, M., Cahana-Amitay, D., Hell, J. V., Kriz, S., & Viguié-Simon, A. (2002). Clause packaging in writing and speech:: A cross-linguistic developmental analysis. *Written Language & Literacy*, *5*(2), 135-161.

Vernon, S. A. (1993). Initial sound/letter correspondences in children's early written productions. *Journal of Research in Childhood Education*, *8*(1), 12-22.

Vernon, S. A., & Ferreiro, E. (1999). Writing development: A neglected variable in the consideration of phonological awareness. *Harvard Educational Review*,*69*(4), 395-416.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*,*101*(2), 192-212.

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.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology, 30*(1), 73-87.

Wallace, D. L., & Hayes, J. R. (1991). Redefining revision for freshmen. *Research in the Teaching of English, 25*, 54-66.

Wechsler, D. (2001). *WPPSI-III: Escala de Inteligencia de Wechslerpara Preescolar y Primaria* [WPPSI-III: Wechsler Preschool and Primary Scale of Intelligence]. Madrid, Spain: Tea.

Wechsler, D. (2003). *Wechsler Preschool and Primary Scale of Intelligence—Third UK Edition (WPPSI-III*UK*)*. Oxford, England:Psychological Corp.

Wheaton, B., Muthén, B., Alwin, D., F., and Summers, G. (1977). Assessing Reliability and Stability in Panel Models. *Sociological Methodology*, *8* (1), 84-136.

Whitaker, D., Berninger, V., Johnston, J., & Lee Swanson, H. (1994). Intraindividual differences in levels of language in intermediate grade writers: Implications for the translating process. *Learning and Individual differences*,*6*(1), 107-130.

Wimmer, H. (2006). Don't neglect reading fluency!. *Developmental Science*,*9*(5), 447.

Ziegler, J. C., & Goswami, U. (2006). Becoming literate in different languages: similar problems, different solutions. *Developmental science*, *9*(5), 429-436.

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., ... & Blomert, L. (2010). Orthographic Depth and Its Impact on Universal Predictors of Reading A Cross-Language Investigation. *Psychological Science*, *21*(4), 551-559.

Appendix 1. Transcription samples

T4_Word_9	T4_Intended_9	T4_Level_9	T4_Segment_9	T4_Case_9	T4_CA_9	4_LetterLength_	T4_PoS_9	T4_Morph_9
we	we	9	4	2	2	2	PR1	0
pet	pet	9	4	2	2	3	N	0
Holday	holiday	9	4	1	1	7	N	0
and	and	9	4	2	2	3	CJ2	0
and	and	9	4	2	2	3	CJ2	0
tv.	TV	9	4	1	2	2	N	0
my	my	9	4	2	2	2	POSS	0
~tromp1Len.	trampoline	9	1	1	1	10	N	0
in	in	9	4	2	2	2	ADV	0
At	At	9	4	2	2	2	PRE	0
won	won	9	4	2	2	3	V	0
dinusar	dinosaur	9	4	2	1	8	N	0
foot_ball	football	9	2	2	1	8	N	0
is	is	9	4	2	2	2	V	0
and	and	9	4	2	2	3	CJ2	0
it	it	9	4	2	2	2	PR3	0
bruthr~	brother	9	1	2	1	7	N	0
the	the	9	4	2	2	3	D	0
house	house	9	4	2	2	5	N	0
his	his	9	4	2	2	3	POSS	0
I	I	9	4	2	2	1	PR1	0
~house.	house	9	1	2	2	5	N	0

Individual words transcription sample. Each row indicates a different child. T4 = end-Year 1; 9 (in column headers) = ninth word in the text. From left to right: the "Word" column shows the literal transcription of the child; "Intended" shows the meaning that the child wanted to convey. The remaining columns were used to classify each word as to its level of representation (Level); whether it showed conventional word boundaries (segment); whether it showed appropriate use of case (case); whether it showed conventional spelling (CA); automatic calculation of word length in letters (LetterLength); the part of speech (PoS); and representation of morphology, where appropriate (Morph).

T4_Fulltext_Original	T4_Fulltext_Intended
I went Pond diPing I got suM minibests. I bilt a bug hotel. I sur my daddy we went to the Park I sur Lydia and Emily H and and Emma and Zoe and Lydia. and romns in a romn wor.	I went pond dipping I got some mini-beasts I built a bug hotel I saw my daddy we went to the park I saw Lydia and Emily H(olmes) and and Emma and Zoe and Lydia And Romans in a Roman world
On Saturday I went to my club and sandy Lane. On Sunday I watch my mum~ ~and sister at race for Life. we had Lunch. We went shoping. I went home and saw my DaD. we played. a gaMe.	On Saturday I went to my club and Sandy Lane On Sunday I watched my mum and sister at Race for Life We had lunch We went shoping I went home and saw my dad We played a game
I went to chestnonMods to get a Medel. I had a hotdog I Wone a Medel it woos in a exBox IMe going to get a troFey I had aa grat~ ~time.	I went to Chester Nomads to get a medal I had a hot dog I won a medal it was in an X-Box I'm going to get a trophy I had a great time
I wet BI~ ~rI -- #~ ~cxn I wet~ ~Y1o~ -- ~Pc wit~ ~my~ ~cxn I~ ~wt~ ~o~ ~v --	I went bike riding with my cousin I went to the park with my cousin I went to [the] [park]
I went to My Atey srao and I saw My frend Harvey I plaod football and I -- -- -- trampiln	I went to my auntie Sarah's and I saw my friend Harvey I played football and I went on the trampoline
I had an BBQ at my nannas and gran_dady and owen and Ellie + Joe and RoBBie and mum and dad and oila I play Joe and then I went Home two Bed	I had a BBQ at my nana's and grandaddy and Owen and Ellie and Joe and Robbie and mum and dad and Olivia I played Joe and then I went home to bed
I~ ~weht1~ ~camping~ ~with Mummy daddy daniel Max baniel~ ~deeh~ ~sick he~ ~went home~ ~at~ ~mibnit1	I went camping with mummy daddy Daniel Max Daniel being sick he went home at midnight
on sunday I went to~ ~nothgat and I showd my Dabby my swi!gnin after that I~ ~want To Burger!King and I	On Sunday I went to Northgate and I showed my daddy my swimming after that I went to Burguer King and I

Complete text transcription sample. Words were transcribed and coded one by one, but they could later be combined, so as to read the entire text transcribed. The left column displays the literal transcription and the right column, the intended meaning. The symbol ~ indicates that the word was attached to either the preceding or following word (or to both).

Appendix 2. Teacher questionnaires

Reception Year

1. How much time do you devote <u>specifically</u> to phonics instruction?
    a. Less than 15 minutes a day
    b. Around 15 minutes a day
    c. More than 15 minutes a day

2. Which of these methods in phonics instruction do you usually follow?
    a. Jolly phonics
    b. Progression in phonics/Letters-to-sound programme
    c. Another method (specify): ………………………………………………

3. Which reading scheme(s) do you use? Please, indicate title(s) and publisher(s).
    ....
    ....
    ....

4. How often do you read a story to the whole class (story time)?
    a. Daily
    b. Every other day
    c. Less than twice a week

5. How often do children take part in guided reading?
    a. Daily
    b. Every other day
    c. Less than twice a week

6. Do you have an in-class library?
    a. Yes
    b. No
    c. Other (specify): ....

7. If you do have an in-class library, how do you make use of this library?
    a. Children are allowed to consult the books almost anytime
    b. There's a specific moment daily in which children are asked to use it.
    c. There's a specific moment during the week, but not every day, in which children are asked to use it.

8. How often do children read out loud to the teacher or teaching assistant individually?
    a. Never
    b. Once a week
    c. More than once a week

9. How much time do you devote to handwriting (letter-formation instruction) activities?
    a. Less than 15 minutes a day
    b. Around 15 minutes a day
    c. More than 15 minutes a day

10. How much time do you devote to children writing on their own?
    a. Less than 15 minutes a day
    b. Around 15 minutes a day
    c. More than 15 minutes a day

11. How much time do you devote to teaching sight-word vocabulary?
    a. It's done on a daily basis
    b. Every other day
    c. About twice a week or less

12. How often do you give reading homework to children?
    a. Never
    b. Once a week
    c. More than once a week

13. How often do you give writing homework to children?
    a. Never
    b. Once a week
    c. More than once a week

14. Please, prioritise the following literacy-related activities in terms of their importance, from highly important (1) to less important (5) in the reception class.
    Phonics
    Sight-word vocabulary
    Autonomous book reading
    Autonomous writing
    Letter-formation instruction

15. Please, prioritise the following literacy-related activities from highly important (1) to less important (5), in terms of the importance that is given to them in the literacy programme that you currently follow in the reception class.
    Phonics
    Sight-word vocabulary
    Autonomous book reading
    Autonomous writing
    Letter-formation instruction

Year 1

1. How much time do you devote daily to phonics instruction?
   (please specify minutes or hours/day).

   ....

2. Which of the following methods of phonics do you use? (please tick more than one if applicable)
   ☐ Jolly Phonics
   ☐ THRASS
   ☐ Progression in phonics/Letters-to-sound programme
   ☐ Other methods (please, specify):

   ....

3. Which reading scheme/s do you use?
(please tick more than one if applicable)
   ☐ Oxford Reading Tree
   ☐ Ginn All Aboard
   ☐ Other methods (please, specify):

   ....

4. How much time do the children spend in guided reading & writing group sessions? (please indicate minutes or hours/week for each)

Reading:
....

Writing:
....

5. How much time do you devote daily to storytelling (story time)? (Please, specify minutes or hours/day).

   ....

6. Do you use the specific list of words (Vocabulary taught list) that NLS recommends for Year 1?
   ☐ Yes
   ☐ No
   ☐ No, but I use a different list (please specify):

   ....

7. Please tick any of the literacy skills taught in the Year 1 curriculum:
   ☐ Letter naming
   ☐ Phoneme awareness
   ☐ Letter-sound knowledge
   ☐ Letter writing
   ☐ Single word reading
   ☐ Nonword reading
   ☐ Spelling
   ☐ Nonword writing
   ☐ Reading comprehension
   ☐ Text writing
   ☐ Other (please, specify):

   ....

8. Do you have an in-class library and, if so, how do you make use of it?
   ☐ We don’t have an in-class library.
   ☐ We have an in-class library and children are allowed to consult the books almost anytime.
   ☐ We have an in-class library and there’s a specific moment daily in which children are asked to use it.
   ☐ We have an in-class library and there’s a specific moment sometimes during the week, but not every day, in which children are asked to use it.

9. What kind of homework do you usually give to children and with what frequency?

	Never	1 x week	+1 x week
reading	☐	☐	☐
spelling	☐	☐	☐
hand writing	☐	☐	☐
text writing	☐	☐	☐
other (specify):	☐	☐	☐

....

10. Could you please indicate how much time (if any) do you devote daily to the writing activities listed below? (Please, specify min/day)

☐ I don't devote any time to writing activities.
☐ handwriting: .... mins/day.
☐ guided text-writing: .... mins/day.
☐ free text writing: .... mins/day.
☐ spelling: .... mins/day.

11. Please, indicate the extent to which you follow the NLS Year 1 curriculum to the teaching of reading and writing on a 1-5 scale (1=I don't follow this curriculum at all; 5=I follow this curriculum completely).

1 **2** **3** **4** **5**

12. By the end of Year 1, what do you consider the ideal level of attainment to be for each of the following skills? (you may tick more than one option).

	Not at all	**Basic level**	**Partial mastery**	**Full mastery**
single word reading	☐	☐	☐	☐
single word writing	☐	☐	☐	☐
sentence reading	☐	☐	☐	☐
text reading	☐	☐	☐	☐
text writing	☐	☐	☐	☐
sentence comp.	☐	☐	☐	☐
text compr.	☐	☐	☐	☐

13. Do you teach spelling patterns?

☐ Yes

☐ No

If you answered YES above, please indicate the method used:

☐ spelling exercises (e.g., worksheets)
☐ reading tasks
☐ free writing
☐ other (please, specify):

....

14. Please, indicate if you teach any of the grammar rules listed below:

☐ plural formation (e.g., one *foot* vs. two *feet*, one *duck* vs. two *ducks*)
☐ production of possessive *–s* (e.g., This is the *boy's* ball)
☐ third-person singular formation (e.g., Mary *runs* fast)
☐ past simple formation (e.g., John *ate* an apple, Mary *played* the piano).
☐ derivatives formation (e.g., Someone who paints walls is a *painter*).
☐ None of the above (if other, please specify):

....

15. Does your school offer any reading and/or writing courses/workshops outside the school hours?

☐ Yes
☐ No

If you answered YES above, please provide further details:

....
....
....

Appendix 3. Assessment of teacher practices

Table A3.1

Results of the Kruskal-Wallis tests to determine teacher effects on literacy activities

	English		Spanish	
Time devoted to literacy activities	*Chi*^{2} *(df)*	*p*	*Chi*^{2} *(df)*	*p*
Writing_Reception	8.00 (8)	.433	4.33 (4)	.364
Reading_Reception	8.00 (8)	.433	4.45 (4)	.349
Storytelling_Reception	0.00 (8)	1.000	5.72 (4)	.221
Phonics_Reception	8.00 (8)	.433	5.72 (4)	.221
Writing_Year 1	6.93 (7)	.436	8.41 (5)	.135
Reading_Year 1	7.93 (7)	.339	9.44 (5)	.093
Storytelling_Year 1	7.73 (7)	.357	8.43 (5)	.134
Phonics_Year 1	7.73 (7)	.357	7.73 (5)	.172
Writing_Year 2	5.00 (5)	.416	6.28 (5)	.280
Reading_Year 2	5.00 (5)	.416	8.78 (5)	.118
Storytelling_Year 2	3.00 (3)	.392	7.73 (4)	.102
Phonics_Year 2	4.00 (4)	.406	7.37 (4)	.118

Appendix 4. Two-way repeated-measures ANOVAs controlling for age and parents' education

A4.1 Objectives and Description of the Sample

The analyses reported in Section 4.3.1 did not control for the effect of age and parents' education, which had been found to differ across languages (Section 3.2). In this section I report the results controlling for these variables, and describe the discrepancies between these two sets of analyses. Finally, an explanation for the observed discrepancies is provided. The criteria for conducting the analyses in Section 4.3.1 was identical to those followed to run the present set—e.g., type of contrasts, calculations of effect sizes, etc. (see Section 4.3.1 for more details). Note, however, that sample sizes were reduced considerably, since parents' education information was only available for about half the sample in each language group. The English sub-group was thus constituted of 78 children (39 boys), mean age at Time 1: 60.22 months (SD = 3.57). The Spanish sub-sample was constituted of 88 children (46 boys), mean age at Time 1: 66.45 months (SD = 3.70).

It was important to ensure the comparability of the selected sub-sample (children with parents' education information) to the unselected sub-sample (children with no information on parents' education) in each language. The distribution of male and female participants was similar across language sub-groups, $\chi^2(1) = 0.09$, $p = .770$, as was the case with the complete samples (Section 3.2). A significant difference in the mean age between the two sub-groups was found, whereby Spanish children were, on average, older than English children, $t(164) = 12.19$, $p < .001$, $d = -1.95$, which is consistent with the difference found for the larger groups (Section 3.2). No significant differences were found between the mean age of the selected and the unselected sub-samples in English, $t(186) = 0.17$, $p = .867$, or Spanish, $t(188) = 1.08$, $p = .284$. Tables A4.1, A4.2, and A4.3 show the descriptive statistics of the sub-samples and the full sample for each dependent variable and language group at Times 3, 4, and 5, respectively.

Table A4.1

Mean scores and standard deviations of selected and unselected sub-samples, and of the full sample in each language group at Time 3

	English			Spanish		
	full sample *M* (*SD*)	selected sample *M* (*SD*)	unselect. sample *M* (*SD*)	full sample *M* (*SD*)	selected sample *M* (*SD*)	unselect. sample *M* (*SD*)
TSPELL	47.34 (23.77)	53.49 (23.05)	40.77 (22.90)	55.30 (26.66)	57.07 (26.91)	51.70 (26.09)
SEG	48.41 (36.58)	58.27 (36.15)	37.49 (34.03)	54.88 (34.22)	58.27 (33.62)	47.93 (34.77)
CASE	65.43 (26.31)	69.97 (24.31)	60.32 (27.52)	64.71 (32.65)	64.93 (33.44)	64.26 (31.36)
WORD	15.36 (8.52)	16.71 (8.62)	14.48 (8.32)	14.36 (8.41)	14.84 (8.77)	13.37 (7.63)
LETT	56.60 (31.21)	60.77 (32.31)	52.14 (29.56)	53.64 (29.83)	54.52 (29.99)	51.84 (29.75)
PUNCT	5.41 (8.68)	6.05 (8.76)	4.71 (8.61)	1.29 (4.49)	1.35 (4.35)	1.17 (4.84)
CONN	4.07 (6.27)	5.26 (7.08)	2.80 (5.02)	6.42 (7.32)	6.45 (7.40)	6.35 (7.26)
WCL	5.08 (1.30)	5.05 (1.40)	5.12 (1.18)	4.97 (2.52)	4.98 (2.61)	4.96 (2.34)
SUBORD	5.77 (11.42)	6.50 (12.28)	4.98 (10.45)	6.59 (14.05)	7.10 (14.78)	5.54 (12.50)
ADJADV	2.26 (4.15)	2.80 (4.63)	1.69 (3.51)	1.94 (3.58)	2.27 (3.81)	1.23 (2.96)
LEXDEN	46.48 (16.78)	46.27 (13.45)	46.72 (19.83)	48.78 (17.29)	49.52 (17.54)	47.25 (16.87)

Note. TSPELL: spelling accuracy; SEG: word segmentation; CASE: use of case; WORD: number of words; LETT: number of letters; PUNCT: punctuation; CONN: connectors; WCL: words per clause; SUBORD: percentage of subordination; LEXDEN: lexical density; ADJADV: percentage of adjectives and adverbs.

Table A4.2

Mean scores and standard deviations of selected and unselected sub-samples, and of the full sample in each language group at Time 4

	English			Spanish		
	full sample *M* (*SD*)	selected sample *M* (*SD*)	unselect. sample *M* (*SD*)	full sample *M* (*SD*)	selected sample *M* (*SD*)	unselect. sample *M* (*SD*)
TSPELL	64.08 (21.15)	68.85 (19.43)	58.98 (21.83)	81.16 (12.28)	81.94 (12.38)	79.56 (12.06)
SEG	68.58 (33.37)	74.01 (30.57)	63.12 (35.37)	87.79 (13.79)	89.95 (12.72)	83.37 (14.97)
CASE	75.28 (21.57)	78.51 (19.90)	72.03 (22.80)	91.70 (8.36)	92.52 (7.71)	90.03 (9.43)
WORD	25.08 (12.57)	25.92 (13.27)	24.19 (11.81)	24.04 (8.55)	24.22 (8.75)	23.66 (8.23)
LETT	92.15 (46.15)	96.26 (49.77)	87.77 (41.85)	88.79 (30.35)	90.56 (30.95)	85.16 (29.11)
PUNCT	6.03 (6.53)	6.30 (6.53)	5.75 (6.57)	4.07 (5.43)	4.20 (5.47)	3.80 (5.40)
CONN	5.73 (5.98)	6.55 (6.28)	4.85 (5.55)	9.34 (6.63)	9.34 (6.60)	9.34 (6.79)
WCL	5.83 (1.90)	5.99 (1.74)	5.65 (2.06)	5.79 (2.03)	5.45 (1.55)	6.48 (2.65)
SUBORD	5.78 (10.70)	6.91 (11.75)	4.58 (9.40)	9.04 (14.31)	10.49 (15.81)	6.09 (10.14)
ADJADV	4.73 (5.71)	5.41 (6.38)	4.01 (4.82)	3.69 (4.84)	3.52 (4.60)	4.04 (5.34)
LEXDEN	43.96 (11.05)	43.28 (10.59)	44.68 (11.55)	44.17 (8.15)	44.72 (8.23)	43.03 (7.97)

Note. TSPELL: spelling accuracy; SEG: word segmentation; CASE: use of case; WORD: number of words; LETT: number of letters; PUNCT: punctuation; CONN: connectors; WCL: words per clause; SUBORD: percentage of subordination; LEXDEN: lexical density; ADJADV: percentage of adjectives and adverbs.

Table A4.3

Mean scores and standard deviations of selected and unselected sub-samples, and of the full sample in each language group at Time 5

	English			Spanish		
	full sample *M* (*SD*)	selected sample *M* (*SD*)	unselect. sample *M* (*SD*)	full sample *M* (*SD*)	selected sample *M* (*SD*)	unselect. sample *M* (*SD*)
TSPELL	69.48 (19.37)	71.92 (17.92)	66.87 (20.61)	85.25 (9.58)	86.06 (8.79)	83.61 (10.93)
SEG	79.02 (21.17)	81.08 (25.09)	76.94 (28.33)	92.13 (10.19)	93.14 (8.95)	90.06 (12.22)
CASE	77.19 (18.15)	80.10 (16.89)	74.18 (19.56)	94.10 (5.53)	93.94 (5.59)	94.43 (5.49)
WORD	31.90 (15.89)	33.47 (16.89)	30.23 (14.67)	34.47 (13.29)	33.97 (13.29)	35.52 (13.38)
LETT	117.79 (57.82)	123.14 (60.94)	112.07 (54.11)	129.07 (50.46)	126.56 (50.02)	134.21 (51.56)
PUNCT	6.53 (6.28)	7.18 (6.74)	5.83 (5.73)	5.45 (4.99)	5.49 (5.11)	5.35 (4.81)
CONN	8.13 (5.72)	8.07 (5.58)	8.19 (5.91)	10.55 (5.89)	10.32 (5.56)	11.01 (6.56)
WCL	5.54 (1.20)	5.53 (1.24)	5.56 (1.17)	5.66 (1.57)	5.58 (1.46)	5.82 (1.78)
SUBORD	12.43 (14.97)	14.85 (16.33)	9.84 (12.98)	12.58 (14.78)	12.47 (14.97)	12.80 (14.57)
ADJADV	6.18 (5.26)	6.61 (5.32)	5.71 (5.19)	4.83 (4.34)	4.38 (4.12)	5.75 (4.67)
LEXDEN	46.02 (7.91)	43.28 (10.59)	44.68 (11.55)	46.02 (7.91)	43.99 (6.41)	44.64 (7.13)

Note. TSPELL: spelling accuracy; SEG: word segmentation; CASE: use of case; WORD: number of words; LETT: number of letters; PUNCT: punctuation; CONN: connectors; WCL: words per clause; SUBORD: percentage of subordination; LEXDEN: lexical density; ADJADV: percentage of adjectives and adverbs.

Mean scores for most measures across the two Spanish sub-samples were equivalent, except for segmentation at Time 4, $t(129) = 2.62$, $p = .010$, $d = .46$, where the selected sub-sample outscored the unselected one; and for words per clause at Time 4, $t(129) = 2.35$, $p = .022$, $d = .41$, where the reverse pattern was found (selected < unselected). In English, in contrast, significant differences across sub-samples were found for spelling accuracy at Time 3, $t(149) = 3.40$, $p = .001$, $d = .56$, and Time 4, $t(149) = 2.94$, $p = .004$, $d = .48$; segmentation at Time 3, $t(149) = 3.63$, $p < .001$, $d = .59$, and Time 4, $t(149) = 2.02$, $p = .046$, $d = .33$; case at Time 3, $t(149) = 2.29$, $p = .024$, $d = .38$, and Time 5, $t(149) = 2.03$, $p = .044$, $d = .33$, and percentage of subordination at Time 5, $t(149) = 2.09$, $p = .038$, $d = .34$. Effect sizes were in the small to moderate range. Moreover, in all cases of significant differences, mean scores for the selected sub-sample were higher than those of the unselected one, while the mean score for the complete sample lay somewhere in the middle. The slightly better performance by the selected sample may be a result of the very fact by which they were "selected": These children were the ones whose parents filled in and returned the questionnaires. Parents' response to the questionnaire might reflect their degree of involvement in school activities, which has, in turn, been suggested to predict other types of involvement and, ultimately, to have an impact on children's academic outcomes (e.g. Hoover-Dempsey, Bassler, & Brissie, 1992; Reynolds, 1992). The differences observed between the two sub-samples should be taken into account when interpreting divergences (and similarities) across the two sets of analyses—those controlling for age and parents' education and those that do not control for these effects.

A4.2 Exploration of Control Variables

Partial correlations were run between the age and parents' education and all criterion measures, controlling for the effect of language for each time point separately. Tables A4.4 through A4.6 show the partial correlations between all text-based measures with age and parents' education. All associations were nonsignificant, except for age with number of words and number of letters at Time 3, both of which were of small size. Moreover, it was important to test the degree of overlap between the factor of interest to the study, language or orthographic consistency, and the control variables. The correlation between language and parents' education was small, $r = -.20$, $p < .011$, but the correlation between age and language was very high, $r = .69$, $p < .001$.

Therefore, a potential source of discrepancies between sets of analyses may be related to the fact that (1) no robust links were found between age and parents' education and all text-based measures, which may give rise to spurious effects of the control variables; and (2) age and language/orthographic consistency overlapped to a great extent and may be "competing" for the same variance.

Table A4.4

Partial correlations of age and parents' education with text-based measures at Time 3 controlling for the effect of language

	Age at Time 1 (N = 282)	Parents' education (N = 166)
TSPELL	.152	-.059
SEG	.154	.032
CASE	.062	.037
WORD	.227*	-.139
LETT	.217*	-.130
PUNCT	.062	-.018
CONN	.004	-.007
WCL	.102	-.043
SUBORD	.091	-.051
LEXDEN	-.079	.023
ADJADV	.152	.041

Note. TSPELL: spelling accuracy; SEG: word segmentation; CASE: use of case; WORD: number of words; LETT: number of letters; PUNCT: punctuation; CONN: percentage of connectors; WCL: words per clause; SUBORD: percentage of subordination; LEXDEN: lexical density; ADJADV: percentage of adjectives and adverbs.
* $p < .01$

Table A4.5

Partial correlations of age and parents' education with text-based measures at Time 4 controlling for the effect of language

	Age at Time 1 (N = 282)	Parents' education (N = 166)
TSPELL	.063	-.078
SEG	.047	-.129
CASE	.041	-.044
WORD	.099	-.122
LETT	.125	-.090
PUNCT	.061	.139
CONN	.027	-.072
WCL	-.005	-.084
SUBORD	.060	-.063
LEXDEN	-.052	.016
ADJADV	-.085	-.037

Note. TSPELL: spelling accuracy; SEG: word segmentation; CASE: use of case; WORD: number of words; LETT: number of letters; PUNCT: punctuation; CONN: percentage of connectors; WCL: words per clause; SUBORD: percentage of subordination; LEXDEN: lexical density; ADJADV: percentage of adjectives and adverbs.

* $p < .01$

Table A4.6

Partial correlations of age and parents' education with text-based measures at Time 5 controlling for the effect of language

	Age at Time 1 (N = 282)	Parents' education (N = 166)
TSPELL	-.033	-.003
SEG	-.035	-.055
CASE	.014	.028
WORD	.111	-.042
LETT	.110	-.035
PUNCT	.022	-.003
CONN	.008	-.095
WCL	.015	-.035
SUBORD	.114	.038
LEXDEN	-.047	.055
ADJADV	.002	.028

Note. TSPELL: spelling accuracy; SEG: word segmentation; CASE: use of case; WORD: number of words; LETT: number of letters; PUNCT: punctuation; CONN: percentage of connectors; WCL: words per clause; SUBORD: percentage of subordination; LEXDEN: lexical density; ADJADV: percentage of adjectives and adverbs.

* $p < .01$

A4.3 Mixed ANOVAs with Age and Parents' Education as Covariates

When age and parents' education were entered as covariates, results changed to some extent. The main effect of time, which had been significant for all measures in the first set of analyses, was lost for most measures, except spelling accuracy, $F(1.50, 241.75) = 4.50$, $p = .020$, segmentation, $F(1.39, 223.94) = 10.25$, $p < .001$, and words per clause, $F(1.90, 305.83) = 5.32$, $p = .006$. Similarly to the first set of analyses, improvement over time was stronger for the Time 3-Time 4 contrast—$F(1, 161) = 3.68$, $p = .057$, $r = .148$, for spelling accuracy; $F(1, 161) = 11.35$, $p = .001$, $r = .257$, for segmentation—than for the Time 4-Time 5 contrast—$F(1, 161) = 1.24$, $p = .268$, $r = .089$, for spelling accuracy; $F(1, 27619) = 0.30$, $p =$

.582, r = .045. For the words-per-clause measure, planned contrasts showed a stronger effect of time in the first contrast, $F(1, 161) = 9.78$, $p = .002$, $r = .239$, than in the second one, $F(1, 161) = 2.94$, $p = .088$, $r = .134$.

The main effect of language was significant in the first set of analyses for all word-level measures (Spanish > English), punctuation (Spanish < English), connectivity (Spanish > English), and percentage of adjectives and adverbs (Spanish < English). After controlling for age and parents' education, language was nonsignificant for word-level variables, though it remained significant for punctuation, $F(1, 161) = 19.01$, $p < .001$, $r = .326$, and adjectives and adverbs, $F(1, 161) = 9.09$, $p = .003$, $r = .230$. In both cases, English children outperformed Spanish ones, as was the case in the first set of analyses.

In contrast with previous analyses, a small, though significant, main effect of language was found for text length measured in words, $F(1, 161) = 4.25$, $p = .041$, $r = .161$, which favoured the English group. Note that no significant effect of language was attested for text length in letters, which must be regarded as a more refined measure of text productivity across languages, provided the average word length reported for Spanish (Section 2.2.2), as well as attested for the content words in the texts produced by the participants in this study (Section 3.8.15).

The first set of analyses showed that time and language effects were moderated by a significant interaction that affected word-level variables, productivity, and punctuation. In this second set of analyses, the interaction was nonsignificant for productivity measures and for punctuation. However, it remained significant for word-level variables: spelling accuracy, $F(1.50, 241.75) = 10.39$, $p < .001$; segmentation, $F(1.39, 223.94) = 12.51$, $p < .001$; and case, $F(1.33, 216.69) = 12.84$, $p < .001$. The simple effect of language was significant for spelling accuracy only at Time 5, $F(1, 162) = 16.48$, $p < .001$, $\eta_p^2 = .092$; and for case at Time 4, $F(1, 162) = 11.99$, $p = .001$, $\eta_p^2 = .069$, and Time 5, $F(1, 162) = 20.55$, $p < .001$, $\eta_p^2 = .113$. For segmentation, there was only a marginally significant difference at Time 5, $F(1, 162) = 5.41$, $p = .021$, $\eta_p^2 = .032$. Follow-up repeated measures ANOVAs to test the simple effect of time were nonsignificant for both language groups.

The control variables were rarely significant. No significant effects were attested involving parents' education, either as a main effect or as part of an interaction. A main effect of age was significant for spelling accuracy, $F(1, 161) = 5.21$, $p = .024$; segmentation, $F(1, 161) = 4.16$, $p = .043$; case, $F(1, 161) = 4.01$, $p = .047$; and punctuation, $F(1, 161) = 4.06$, $p = .046$. Finally, a significant Age by Time interaction was attested for segmentation, $F(1, 161) = 5.99$, $p = .003$, and for words per clause, $F(1, 161) = 4.58$, $p = .012$.

The main discrepancies between the first (Chapter 4) and the present set of analyses were (1) the loss of the main effect of time for most measures—both word- and text-level; (2) the changes affecting the main effect of language, which was lost or reduced in some measures or became significant in the case of length in words, which had shown no effect of this factor. Importantly, however, significant interactions between the two main variables remained.

The overlap between language and age reported above is the likely cause for the lack of language effects in variables that had shown sensitivity to it in the first set of analyses; including age as a covariate may provoke suppressor effects. Notably, the only measures for which age was a significant factor, word-level variables and punctuation, did not see their respective patterns of results altered in this second set of analyses. All in all, however, the pattern of results for word-level variables was akin to that obtained in the first set of analyses, since here too the effect of language became stronger with time, as shown by the significant interaction between language and time.

An explanation for the loss of the main effect of time and for the newfound main effect of language in the number of words measure was less straightforward. It was possible that age was suppressing, not just the effect of language, but also growth over time. On the other hand, and although in both sets of analyses productivity in letters was unaffected by language, a more refined analysis of the specific contribution of each factor to explaining variance in text length in words might shed light on the interplay between the main factors and the covariates.

A4.4 Multiple Regressions

The discrepancies observed between the two sets of analyses deemed it necessary to disentangle the specific contribution of each factor to explaining variance in text-based measures. This issue was explored via a series of multiple-regressions, for which the “enter” method was used. Regression models for Time 3 text-based measures included three predictor variables: language (or orthographic consistency), age, and parents’ education (measured in number of years). Models for Time 4 and Time 5 dependent measures also included the corresponding autoregressor. In all cases, the autoregressor was the same measure from the immediately preceding time point. This procedure allowed determining the specific contribution of each of the predictors, particularly with a view to determining the relative contribution of the control variables—age and parents’ education—, over and above that of the variable of interest: language or orthographic consistency.

Table A4.7 shows the standardized beta coefficients of each predictor variable for the Time 3 measures, as well as fit indexes (R^2 and adjusted R^2) for all models. It is apparent from the table that no single predictor or a combination of them explained a considerable amount of variance in any of the mid-Year 1 text-based measures. Similarly to the repeated measures ANOVAs reported above, parents’ education never made a significant contribution to the models. Language had a minor significant contribution for case, length in words and letters, punctuation, percentage of adjectives and adverbs, and lexical density. Age, on the other hand, also had small significant contributions in spelling accuracy, word segmentation, text length measures, words per clause, and percentage of adjective and adverbs. In sum, leaving aside the null effect that parents’ education had in the present sub-sample, language and age at Time 1 were poor predictors of text-based measures at Time 3. The reduced role of language as an explanatory variable of Time 3 writing outcomes is consistent with the findings reported in Chapter 4, by which text features sensitive to variations in orthographic consistency were

Table A4.7

Standardized beta coefficients and R^2 *of the Time 3 models*

	Language/Orthographic consistency		Age at Time 1		Parents' education		Final model fit
	β	t	β	t	β	t	R^2/Adjusted R^2
TSPELL	-.13	-1.25	.27	2.77*	.03	0.40	.043 / .027*
SEG	-.17	-.1.67	.28	2.90*	.08	1.11	.051 / -.035*
CASE	-.22	2.19*	.19	1.97	.01	0.09	.030 / .014
WORD	-.40	-4.15*	.35	3.73*	-.12	-1.64	.097 / .083*
LETT	-.35	-3.58*	.30	3.12*	-.11	-.1.51	.074 / .058*
PUNCT	-.41	-4.24*	.12	1.29	-.02	-0.32	.111 / .097*
CONN	-.05	0.53	.00	0.01	-.03	-0.34	.005 / -.012
WCL	-.15	-1.52	.19	1.98*	-.06	-0.81	.024 / .008
SUBORD	-.02	-0.23	.11	1.06	-.01	-0.06	.008 / -.008
LEXDEN	.23	2.23*	-.16	1.64	.07	0.99	.028 / .012
ADJADV	-.26	-2.60*	.29	3.05*	.08	1.01	.061 / .046*

Note. TSPELL: spelling accuracy; SEG: word segmentation; CASE: use of case; WORD: number of words; LETT: number of letters; PUNCT: punctuation; CONN: connectors; WCL: words per clause; SUBORD: percentage of subordination; LEXDEN: lexical density; ADJADV: percentage of adjectives and adverbs; β: standardized beta coefficient of final model (with all predictors); t: t value.

* $p < .05$

Table A4.8

Standardized beta coefficients and R^2 *of the Time 4 models*

	Time 3 autoregressor		Language/Orthographic consistency		Age at Time 1		Parents' education		Final model fit
	β	*t*	*β*	*t*	*β*	*t*	*β*	*t*	R^2/Adjusted R^2
TSPELL	.42	6.45*	.32	3.52*	.06	0.67	-.014	-0.22	.333 / .317*
SEG	.52	8.32*	.33	3.75*	-.05	-0.59	-.12	1.87	.380 / .365*
CASE	.26	3.89*	.41	4.31*	.07	0.72	-.01	-0.07	.262 / .242*
WORD	.39	5.20*	-.07	-0.63	.01	0.11	-.06	-0.78	.165 / .145
LETT	.40	5.47*	-.09	-0.92	.07	0.66	-.03	-0.39	.178 / .158*
PUNCT	.05	0.68	-.27	-2.46*	.23	2.23*	.14	1.86	.080 / .058*
CONN	.19	2.55*	.17	1.58	.01	0.07	-.09	-1.21	.089 / .067*
WCL	.09	1.18	-.04	0.33	-.20	-1.91	-.12	-1.47	.059 / .037*
SUBORD	.04	0.56	.07	0.66	.04	0.38	-.06	-0.79	.020 / -.003
LEXDEN	.19	2.40*	.05	0.43	.02	0.20	.07	0.86	.042 / .019
ADJADV	.20	2.56*	-.11	-1.00	-.10	-0.99	-.06	-0.77	.071 / .049*

Note. TSPELL: spelling accuracy; SEG: word segmentation; CASE: use of case; WORD: number of words; LETT: number of letters; PUNCT: punctuation; CONN: connectors; WCL: words per clause; SUBORD: percentage of subordination; LEXDEN: lexical density; ADJADV: percentage of adjectives and adverbs; *β*: standardized beta coefficient of final model (with all predictors); *t*: t value.

* $p < .05$

Table A4.9

Standardized beta coefficients and R^2 *of the Time 5 models*

	Time 4 autoregressor		Language/Orthographic consistency		Age at Time 1		Parents' education		Final model fit
	β	*t*	β	*t*	β	*t*	β	*t*	R^2/Adjusted R^2
TSPELL	.61	10.98*	.21	2.84*	.00	0.06	-.01	-0.18	.525 / .515*
SEG	.75	14.69*	.06	0.86	-.01	-0.15	.02	0.37	.588 / .579*
CASE	.44	7.25*	.32	3.90*	.08	0.98	.07	1.25	.449 / .437*
WORD	.47	7.08*	.07	0.72	.02	0.21	.01	0.20	.219 / .203*
LETT	.48	7.24*	.10	1.03	-.01	-0.15	.02	0.29	.226 / .209*
PUNCT	.27	3.78*	-.18	-1.79	.10	1.05	-.02	-0.27	.103 / .083*
CONN	.19	2.54*	.15	1.45	-.01	-0.06	.00	0.03	.066 / .045*
WCL	.06	0.81	-.03	-0.32	.08	0.81	-.12	-1.62	.024 / .003
SUBORD	.08	1.10	-.24	-2.36*	.27	2.68*	.09	1.23	.056 / .035*
LEXDEN	.27	3.88*	-.12	-1.22	-.07	-0.71	.09	1.18	.112 / .093*
ADJADV	.17	2.34*	-.26	-2.54*	.10	1.05	.05	0.62	.083 / .063*

Note. TSPELL: spelling accuracy; SEG: word segmentation; CASE: use of case; WORD: number of words; LETT: number of letters; PUNCT: punctuation; CONN: connectors; WCL: words per clause; SUBORD: percentage of subordination; LEXDEN: lexical density; ADJADV: percentage of adjectives and adverbs; β: standardized beta coefficient of final model (with all predictors); *t*: t value.

* $p < .05$

more so at later stages of writing development. For example, effect sizes calculated for the simple effect of language in word-level variables were larger at Times 4 and 5, but language had a small or nonsignificant effect at Time 3. Finally, the standardized beta coefficients for language and age, in those cases where a significant contribution was observed, were always of opposite sign, thus supporting the claim that age has suppressor effects on the language effect.

Tables A4.8 and A4.9 show the Time 4 and Time 5 models. In these models language had a significant contribution in explaining variance in word-level variables, punctuation, and percentage of adjectives and adverbs, consistent with findings reported in Chapter 4[15]. Age, on the other hand, did not contribute to explaining variance in text-based measures over and beyond the contribution of the autoregressor and language, with the exception of punctuation at Time 4 and percentage of subordination at Time 5. The role of parents' education was, again, negligible in the current sub-sample.

A4.5 Concluding Remarks

Age and parents' level of education had no reliable impact on the criterion variables. Including them in the analyses, as covariates, resulted in discrepancies between the results of the repeated-measures ANOVAs reported in Chapter 4, which did not control for those variables. Such discrepancies mainly involved the loss of the main effect of time and/or its interactions with language for many variables. It has been now shown that age and parents' level of education did not have significant associations with text-based measures as indicated both by (a) the partial correlations between control variables and text-based measures at every time point, which were generally nonsignificant; and (b) the multiple regressions conducted for each time point and dependent variable, in which neither control variable had a clear impact over and above language and autoregressor effects.

[15] A significant contribution was also attested in this new set of analyses for percentage of subordination at Time 5; however, note that the mean for the English sub-sample was significantly higher than the rest of the sample.