Reading and Writing (2020) 33:1325–1349
https://doi.org/10.1007/s11145-019-10012-5

Portuguese spelling in primary grades: complexity, length and lexicality effects

Ana Mesquita[1] · Lénia Carvalhais[2] · Teresa Limpo[1] · São Luís Castro[1]

Published online: 2 January 2020

Abstract

This study investigates spelling abilities of 189 second, third, and fourth graders using a word and pseudoword dictation task in European Portuguese. We analyzed the effect of orthographic complexity on spelling accuracy and the moderating role of length (two vs. three syllables), lexicality (words vs. pseudowords), and grade (second, third, and fourth). Each item represented one of the following orthographic complexity categories: digraph, contextual consistency, position consistency, consonant cluster, stress mark, inconsistency, and silent letter ‹h›. Digraphs and position consistencies reached high levels of accuracy already in Grade 2, but stress marks, inconsistencies, and the silent letter ‹h› were not yet fully mastered by the end of primary school. Performance across complexities was more discrepant in Grade 2 than in Grades 3 and 4. Moreover, within each complexity, there were larger differences between Grades 2 and 3 than between Grades 3 and 4. Words were better spelled than pseudowords in position consistency and stress mark categories, and a shorter length improved accuracy only in consonant clusters and stress marks. These findings underline the importance of applying learning and teaching strategies in early education adapted to the properties of the writing system to be learned.

Keywords Length effect · Lexicality effect · Orthographic complexities · Portuguese spelling · Spelling acquisition

Ana Mesquita and Lénia Carvalhais are Joint first authorship.

✉ São Luís Castro
slcastro@fpce.up.pt

[1] Faculdade de Psicologia e de Ciências da Educação, Universidade do Porto, Rua Alfredo Allen, 4200-135 Porto, Portugal

[2] Department of Psychology and Education, Portucalense University, Porto, Portugal

Introduction

Writing is a core skill for effective human communication in professional as well as private settings. Skilled writers must be accurate spellers, and accurate spelling requires a solid knowledge of speech-to-print conventions. In alphabetic writing systems, the first step to achieve such knowledge is to acquire the alphabetic principle (Liberman, Shankweiler, & Liberman, 1989). This refers to the ability to match the speech sounds of a language (i.e., phonemes) with their accurate representation in written form (i.e., graphemes). A child who understands that each written word represents a structured sequence of individual speech sounds has assimilated the core principle of alphabetic systems (Liberman et al., 1989). However, to become a proficient writer the child has yet to master the constraints imposed by the orthography of his or her language. Two potential sources of constraints include orthographic depth and syllable structure (Seymour, Aro, Erskine, & the COST Action A8 network, 2003). Orthographic depth varies along a continuum, from shallow to deep orthographies (Katz & Frost, 1992), and comprises both the complexity and unpredictability, or inconsistency, of the correspondences between phonemes and graphemes (Schmalz, Marinus, Coltheart, & Castles, 2015). Shallow orthographies have simple and consistent relations between phonemes and graphemes (e.g., Finnish), whereas deep orthographies have complex and inconsistent sound-letter mappings (e.g., English). Another source of constraints is syllable structure, which refers to the distribution of consonants (C) and vowels (V) within a syllable. This distribution can be simple (e.g., CV) or complex (e.g., CCV). Groups of consonants in complex syllables can be difficult to spell, specially for children (e.g., Treiman, 1991). Overall, frequent complexities and inconsistencies in an orthographic system make spelling acquisition more challenging for beginning writers (e.g., Serrano et al., 2011).

Portuguese is an orthography of intermediate depth (Seymour et al., 2003) that has several complex as well as inconsistent spelling features. Studies focusing on the acquisition of these features by Portuguese children are scarce. In a study of first graders, Fernandes, Ventura, Querido, and Morais (2008) found a progression from reliance on phoneme-grapheme conversion in February, indicated by regularity and complexity effects, to the emergence of an orthographic lexicon in June, indicated by lexicality effects; as the authors pointed out, however, phonological mediation was probably preponderant even by the end of the year, as regularization errors increased from February to June. In a cross-language study of first graders, spelling acquisition seemed to progress more slowly in Portuguese than in Spanish, a shallower system, but faster than in French, a deeper system (Serrano et al., 2011). Studies with older children focused on spelling in children with dyslexia and the analysis of specific misspellings (Horta & Martins, 2004; Vale & Sousa, 2017). Horta and Martins (2004) carried out a longitudinal study with children in Grade 3 and then in Grade 4. Using a spelling-to-dictation task, the authors found that children's spellings became significantly more accurate from Grade 3 to Grade 4, but the proportion of the different types of errors (e.g., orthographic, phonetic) was comparable between grades. Vale and Sousa (2017)

compared spelling skills of fifth graders with dyslexia to those of typical readers from Grade 5 and Grade 3 (matched in reading level). Results from typical readers showed that children in Grade 5 had more correct word spellings than those in Grade 3, particularly in context-dependent and irregular spellings. Finally, a study by Almeida, Guerreiro, and Mata (1998) showed that Portuguese second graders struggled more in justifying the spelling of inconsistent than consistent words. Other studies have turned to pre-school children and addressed the impact of invented spelling programs in reading and spelling development (e.g., Martins, Albuquerque, Salvador, & Silva, 2013; Martins, Salvador, Albuquerque, & Silva, 2016). On the whole, then, an overall picture of how spelling develops across the primary school grades is lacking. Thus, the present study aimed to analyze how the complex and inconsistent features of the Portuguese orthography affect spelling acquisition beyond the first grade and how length and lexicality impact on this acquisition.

Characterization of the Portuguese spelling system

Portuguese is a Romance language from the Indo-European family. According to the Ethnologue database (Eberhard, Simons, & Fennig, 2019), it is the native language of more than 220 million people widespread across more than 10 countries. Here, we focus on European Portuguese, whose spelling and phonology do not fully overlap with other varieties such as Brazilian Portuguese (e.g., Sucena, Castro, & Seymour, 2009). Core features of the Portuguese spelling system are presented below, with phonemes indicated between slashes //, graphemes between angled brackets ‹›, and translations in italics.

Concerning the orthographic depth of sound-to-print correspondences, Portuguese has 37 phonemes to 67 simple and complex graphemes (Gomes, 2001). This proportion might suggest that one phoneme matches about two graphemes, but the actual speech-to-print distribution is not homogeneous. Some phonemes have a univocal correspondence to print (one phoneme to one grapheme), such as /p/ (‹pedra› *stone* /ˈ**p**ɛ.dɾɐ/) or /t/ (‹tela› *screen* /ˈ**t**ɛ.lɐ/). To correctly spell these sounds, children can rely on the simple rule that /p/ is always written ‹p› and /t/ is always written ‹t›. However, other phonemes have multiple correspondences (one phoneme to two, three, or four graphemes; see Moutinho, 2000), such as /i/, that can be represented by ‹e› (‹**e**feito› *effect* /**i**.ˈfɐj.tu/), ‹i› (‹b**i**co› *beak* /ˈb**i**.ku/) or ‹í› (‹**í**ris› *iris* /ˈ**i**.ɾiʃ/), and these can be of two types: complex consistencies and inconsistencies (Schmalz et al., 2015). Complex consistencies occur when the use of correct graphemes is predictable but involves a multi-letter representation and/or takes into account context-dependent rules. These rules rely on additional information about the target phoneme, such as its phonemic context (e.g., subsequent vowel) or its position within the word (e.g., syllable- or word-initial or final positions). Inconsistencies occur when no explicitly known rules indicate the correct spelling of the word.

Regarding syllable structure, Portuguese words are mostly composed of CV syllables (Gomes, 2001; Vigário & Falé, 1994). Although this is the most frequent pattern, Portuguese also uses more complex syllable structures, such as CCV and

CCVC (e.g., ‹prato› *dish* /ˈpɾa.tu/ and ‹flor› *flower* /ˈflor/, respectively). These syllables contain consecutive consonants (consonant clusters), which may be challenging for young learners to spell (e.g., Treiman, 1991).

In spelling, children may also process information at the level of morphemes, the smaller chunks of meaning. The Portuguese language includes lexical morphemes (radicals), as well as grammatical morphemes that allow the formation of new words originating in the radical (e.g., variations in gender, such as *-o* / *-a* in ‹gato› *cat* and ‹gata› *cat*, or number, such as *-s* in ‹gatos› *cats*; see Cunha & Cintra, 2014). Because words with a common radical share a common part, children may use this knowledge as a cue to spell unknown words (e.g., if the child knows that ‹humana› *human* is spelled with an initial silent ‹h›, he/she can infer that the word ‹humanidade› *humanity* begins also with an ‹h›).

In what follows we present features of the Portuguese spelling system that children need to master in order to become fluent readers and writers (for a description of expected landmarks in basic education for achieving proficient use of Portuguese as a native language, see, e.g., Sim-Sim, Duarte, & Ferraz, 1997).

Digraphs

A digraph refers to the use of two letters to represent a single phoneme. Digraphs do not exist in some languages, such as Hebrew (Kahn-Horwitz, Schwartz, & Share, 2011), though they are very common in others, such as French (Fernandes et al., 2008). Portuguese is somewhere in the middle: it includes two digraphs that establish biunivocal correspondences with phonemes, ‹nh› and ‹lh› (/ɲ/ and /ʎ/, respectively), as well as digraphs that are inconsistent, such as ‹ch›. The digraph ‹ch› represents the phoneme /ʃ/, but depending on the context /ʃ/ can also be spelled with ‹x›, ‹s›, or ‹z›. Portuguese also includes the digraphs ‹rr› and ‹ss›, which represent /ʁ/ and /s/ in intervocalic position; ‹qu› and ‹gu›, which are context-dependent digraphs (see next paragraph); and combinations such as ‹am› or ‹an› (i.e., a vowel plus ‹m› or ‹n›), which represent nasal vowels depending on context. Learning that two letters represent one sound can pose difficulties in the initial school grades. However, as learners develop their spelling ability, these letter combinations turn out to be processed as a unit (Tainturier & Rapp, 2004).

Contextual and position consistencies

In Portuguese, some phoneme-grapheme correspondences are context-dependent. Here, we use the term *contextual consistencies* to distinguish cases where a phoneme is represented by a specific grapheme depending on the vowel that immediately follows it, and *position consistencies* to refer to cases where the selection of the grapheme is determined by the position of the phoneme within the word (a particular case of context-dependent consistencies). Examples of contextual consistencies involve /k/ and /g/. If these phonemes are followed by the front vowels /e/, /ɛ/, /i/, /ẽ/, or /ĩ/, the semivowel /j/, or schwa (/ə/), the graphemic representations ‹qu› and ‹gu› should be used. For example, /kɐ.ˈfɛ/ (‹**c**afé› *coffee*) and /ˈgɔ.lɐ/ (‹**g**ola› *collar*) are spelled with ‹c› and ‹g›, whereas /ˈkɛ.dɐ/ and /ˈgi.zu/ are spelled ‹**que**da› *fall* and

‹**gui**zo› *rattle*. If ‹c› and ‹g› were used in the latter examples, they would be read as /s/, /ˈsɛ.dɐ/, and /ʒ/, /ˈʒi.zu/. Position consistencies occur in the representation of the phonemes /ʁ/ and /s/. In word initial position, the graphemes ‹r› (‹**r**ato› *mouse* /ˈʁa.tu/) and ‹s› (‹**s**aco› *bag* /ˈsa.ku/) are used. However, the same graphemes in coda position would be pronounced /ɾ/ (‹co**r**da› *rope* /ˈkɔɾ.dɐ/) and /ʃ/ (‹pa**s**ta› *folder* /ˈpaʃ.tɐ/). In intervocalic position, /ʁ/ and /s/ have to be represented with double letters (‹rr›, ‹ss›). Another position consistency relates to the nasalization of the vowels: the grapheme ‹m› should be used before /p/ and /b/, and the grapheme ‹n› before all other consonants. Overall, both contextual and positional consistencies require children to go beyond the single phoneme and take into account its neighborhood. In Spanish, which shares with Portuguese similar constraints regarding the spelling rules described above, children from Grades 1–4 achieved similar accuracy in contextual and position consistencies (Defior, Jiménez-Fernández, & Serrano, 2006, 2009). However, the acquisition of this type of complex consistencies seems to be influenced by the depth of the orthographic learning system. For example, when asked to spell words containing context-dependent consistencies, second graders learning a shallow orthography (Spanish) outperformed their peers learning a deeper orthography (French; Carrillo, Alegría, & Marín, 2013).

Consonant clusters

Consonant clusters are groups of consecutive consonants within a syllable, such as ‹dr› (‹**dr**ama› *drama* /ˈdɾɐ.mɐ/), ‹pl› (‹du**pl**a› *double* /ˈdu.plɐ/), and ‹cl› (‹te**cl**ado› *keyboard* /tɛ.ˈkla.du/). Although in spoken Portuguese three or more consecutive consonants can occur due to vowel reduction (Mateus & D'Andrade, 1998), in spelling consonant clusters have only two consonants and their frequency is relatively low (Vigário & Falé, 1994). A potential misspelling of consonant clusters is omitting one of the consonants. This error occurs in Portuguese (e.g., spelling the word ‹abutre› *vulture* /ɐ.ˈbu.tɾə/ as ‹abute›, where the second consonant of the cluster, ‹r›, is missing) as well as in other languages, such as English (e.g., Treiman, 1991) or Dutch (e.g., Van Bon & Uit de Haag, 1997). Another potential misspelling is the inclusion of an epenthetic vowel to provide support between consonants (e.g., misspelling the word ‹dupla› *double* /ˈdu.plɐ/ as ‹dupela› /ˈdu.pə.lɐ/). Interestingly, the epenthetic vowel is frequently observed not only in early spelling errors, but also in misarticulations of 3- to 5-year-old children (Castro & Gomes, 2000).

Stress marks

In some languages, diacritics can be used to convey word stress. In Portuguese, three diacritics mark stress and indicate whether the vowel is open or closed. The acute accent /´/ is used in words stressed in the antepenultimate syllable and in the penultimate syllable of words ending with ‹i› or ‹l›. The circumflex accent /^/ is used in closed vowels, such as ‹â›, ‹ê›, and ‹ô›. The tilde /~/ nasalizes the vowels ‹a› and ‹o› (e.g., ‹c**ã**o› *dog* /ˈkɐ̃w/, ‹aç**õ**es› *actions* /ˈɐ.sõ̞jʃ/) and it is typically used in stressed syllables. The most frequent stress mark in Portuguese is the acute accent, followed by the tilde and the circumflex accent (Gomes, 2001). In Spanish, stress marks are

one of the most difficult orthographic complexities to acquire (Defior et al., 2006, 2009); in Brazilian Portuguese, teachers also expect proficiency in stress marks to occur only in later phases of spelling development (Pinheiro, 1995).

Inconsistencies

Inconsistent phoneme-grapheme correspondences are not regulated by rules. In Portuguese, inconsistent mappings occur mostly in the representation of the phonemes /ʒ/ and /ʃ/. The sound /ʒ/ followed by the front vowels /e/, /ɛ/, /i/, /ẽ/, or /ĩ/, the semivowel /j/, or schwa (/ə/) can be accurately spelled with ‹j› or ‹g› (plus ‹e› or ‹i›, depending on the vowel). In reading, there is a rule determining that the letter ‹g› followed by ‹i› or ‹e› is always read /ʒ/; examples are ‹**gi**rassol› *sunflower* /ʒi.ʁɐ.ˈsɔɬ/, and ‹ti**ge**la› *bowl* /ti.ˈʒɛ.lɐ/, respectively. But in spelling the two possibilities are allowed. Therefore, children need to invoke lexical or morphological knowledge to correctly spell inconsistent words. For example, if children know the word ‹lojista› *shopkeeper*, their lexical knowledge in enough to spell the word. If children are not familiar with the word, but they are aware of the morpheme ‹loj–› (from ‹loja› *shop*), they can use this knowledge to correctly spell the derived form ‹lo**j**ista› (where ‹–ista› indicates the profession). In case of /ʃ/, there is also no rule to determine the grapheme that should be selected by the writer, namely, ‹ch›, ‹x›, ‹s›, or ‹z›. Again, children will only produce correct spellings if they rely on the lexicon or on morphological cues (e.g., ‹co**x**o› *lame* /ˈko.ʃu/, attending to the stem ‹cox–›, which is common to related words such as ‹co**x**ear› *to limp* /ku.ˈʃjaɾ/). Because phonological knowledge and orthographic rules cannot be applied, Portuguese children are expected to have more difficulty in mastering inconsistent over consistent correspondences, a finding already observed in their Brazilian Portuguese-speaking peers (Pinheiro, 1995).

Silent letter ‹h›

Silent letters are graphemes that do not have a phonemic correspondence. Silent letters can occupy different positions within the word and can be always silent or only in some cases. In Portuguese, the grapheme ‹h› is always silent. For example, ‹hera› *ivy* and ‹era› *was/era* are both pronounced /ˈɛ.ɾɐ/. In Spanish also, words written with an initial silent ‹h› do not follow any consistent rule, and children struggle to spell them correctly in the early grades (Defior et al., 2006, 2009). To successfully spell these words, children need to rely on lexical or morphological knowledge.

Length and lexicality effects

The difficulty of spelling a word or pseudoword is not uniquely dependent on the presence of orthographic complexities, such as those previously presented. One feature that can create additional demands on spelling is length. It is harder for children to spell longer items, regardless of length being assessed with number of phonemes or of letters (Treiman, 1993). Longer items can make spelling even more demanding

for individuals with acquired or developmental spelling difficulties. For example, in adults with acquired dysgraphia longer words elicited a higher number of spelling errors (Buchwald & Rapp, 2003; Tainturier & Caramazza, 1996). Furthermore, it was recently observed that length had an even more detrimental effect in spelling in children and adolescents with dyslexia compared to typically developing peers (Juul & Petersen, 2017). Because stimulus length can influence spelling accuracy, considering length effects in the study of spelling acquisition allows a more refined characterization of the development of each orthographic complexity.

Another feature that impacts spelling development is lexical status. A lexicality effect in spelling occurs when writers show better performance in writing words than non-lexical stimuli such as pseudowords (e.g., Defior et al., 2009). Lexicality effects provide important information for the study of spelling acquisition. On the one hand, they suggest that spelling strategies progress from a general phoneme-to-grapheme conversion procedure to the use of specific knowledge about words. In Portuguese, children from Grade 1 showed a lexicality effect in spelling simple words vs. simple pseudowords at the end of the school year, suggesting incipient acquisition of lexical knowledge (Fernandes et al., 2008). On the other hand, comparing word and pseudoword performance may inform about the consolidation of rules underlying orthographic complexities. For example, Defior et al. (2009) found that the advantage of word over pseudoword spelling in primary grades was more pronounced for digraphs, position consistencies, and stress marks than for contextual consistencies. In general, the study of orthographic consistencies in pseudowords is a useful indicator of whether children have generative knowledge of spelling rules and apply it to novel stimuli or if they just use them in words they memorized.

The present study

Our main goal was to examine how the complexities of the Portuguese orthography affect spelling acquisition of beginning writers in primary schools, taking into account stimulus length and lexical status. For this purpose, we assessed the spelling abilities of typically developing Portuguese children from Grades 2, 3, and 4 using a spelling-to-dictation task. This assessment included words and pseudowords requiring explicit or implicit knowledge of selected features of the Portuguese orthography. Our primary focus concerned the acquisition of the orthographic complexities of the Portuguese spelling system. Which features are developed earlier and possibly mastered at the end of primary school? We hypothesized that consistent features (digraph and contextual and position consistencies) and consonant clusters would be developed earlier than inconsistent features (inconsistency and silent letter ‹h› categories) and the stress mark. Because consistent features are predictable, we expected them to be learned before inconsistencies, as inconsistent features require word-specific knowledge and do not rely on general rules. In addition, the stress mark was expected to be a challenging feature. Data from Spanish children in Grades 1–4 (Defior et al., 2009) revealed that stress assignment was one of the most demanding features, not yet mastered at the end of primary school (fourth graders achieved only about 50% accuracy in stress-marked words, whereas in the remaining

complexities the minimum score was 63%). Our second focus was the moderating role of lexicality and length on the acquisition of orthographic complexities across grades. Specifically, we aimed to test whether complexities became more demanding when embedded in longer words. This would suggest a possible cumulative effect of length and complexity. Regarding lexicality, we aimed to explore if, and at which point of primary education, children have consolidated rule-based knowledge in order to apply it to novel stimuli (pseudowords).

Studying the effect of different sources of complexity in spelling acquisition within a language is important for different reasons. First, the specific prevalence of each complexity depends on language. Second, the contribution of each complexity to predict spelling performance may vary across grades (Willson, Rupley, Rodriguez, & Mergen, 1999). Finally, understanding how children acquire each complexity contributes to inform teaching practices (Defior et al., 2006). From a cross-linguistic perspective, the study of an orthography of intermediate depth such as Portuguese can improve our understanding of intermediate systems. In reading, for example, Portuguese elicited the use of flexible (task-dependent) decoding strategies (Lima & Castro, 2010), an adaptive pattern that may not be so present in orthographies with marked shallow or deep features.

Methods

Participants

Participants were 189 children, native speakers of European Portuguese, from three primary schools in the Aveiro area (center of Portugal): 71 second graders ($M = 94.66$ months, $SD = 3.90$; 38 girls), 49 third graders ($M = 105.43$, $SD = 5.07$; 26 girls), and 69 fourth graders ($M = 118.45$, $SD = 4.32$; 32 girls). Children came from public schools in urban middle-class neighborhoods, and none had known hearing, visual, intellectual, or behavioral disorders. Data collection took place between January and April (the second school term in Portuguese schools). All children were individually assessed for intellectual capacity measured with the Raven's Colored Progressive Matrices (Portuguese norms from Simões, 2000), reading proficiency measured with a reading age test (Teste de Idade de Leitura, TIL; Sucena & Castro, 2008; how many correct choices within 5 min for 36 sentences where the last word is missing and has to be chosen from five alternatives), and vocabulary skills measured with the Vocabulary subtest from the Wechsler Intelligence Scale for Children—III (WISC-III; Portuguese version from Simões, Rocha, & Ferreira, 2003). Averages by grade are provided in Table 1. None of the children showed impairments in these tests.

Instructional setting

During primary school years, foundational reading and spelling competencies are part of the Portuguese language school curriculum (Buescu, Morais, Rocha, & Magalhães,

Table 1 Children's characteristics by grade

Measures	Grade 2 ($n=71$)		Grade 3 ($n=49$)		Grade 4 ($n=69$)	
	M	*SD*	*M*	*SD*	*M*	*SD*
RCPM (percentile)[a]	70.56	22.56	58.16	26.19	57.25	22.06
TIL, reading age test (/36)	12.77	6.63	20.08	5.68	25.33	6.94
Vocabulary, WISC-III (raw score, /60)	16.68	5.69	21.10	5.36	24.25	6.21

RCPM, Raven's Colored Progressive Matrices; Teste Idade Leitura, Reading Age Test (see text); WISC-III, Wechsler Intelligence Scale for Children—Third Edition

[a]Raw scores were converted in percentiles according to age (years and months)

2015). In Grades 1–2, the teaching of reading and spelling is greatly focused on the development of basic phonological and orthographic skills through different activities (e.g., letter, grapheme and word/pseudoword naming and writing; phoneme deletion; syllable counting; writing-by-dictation of short texts). In Grades 3–4, teaching is more focused on developing complex reading and writing skills, particularly at the text level. Vocabulary development is fostered along the 4 years. Children contact with texts of different genres (e.g., narration, poetry) and learn to analyze their key components (e.g., main ideas, information about time, space, and characters' attitudes).

Tasks and measures

Based on the study by Defior et al. (2009), we designed a spelling-to-dictation task composed of two parts, first words and then pseudowords.

Word dictation task

This task is composed of 56 words, evenly divided into seven orthographic complexity categories with four two- and three-syllable words each ($7 \times 4 \times 2$): digraph, contextual consistency, position consistency, consonant cluster, stress mark, inconsistency, and silent letter ‹h›. The words were selected from Portulex (Teixeira & Castro, 2006), a lexical database containing words from Portuguese textbooks used in Grades 1–4. We selected words containing the most common syllable types in Portuguese (CV, CVC, V; Gomes, 2001; Vigário & Falé, 1994; see "Appendix"). Within each category, word absolute frequency ranged from low to medium values as described next: 1–36 for digraphs, 1–18 for contextual consistency, 1–30 for position consistency, 1–26 for consonant clusters, 2–50 for stress marks, 3–26 for inconsistencies, and 1–25 for silent ‹h›. Below we describe and provide examples of each complexity category.

Digraph

Digraphs are two letters representing one sound. We included the digraphs ‹lh› (/ʎ/), ‹nh› (/ɲ/), and ‹rr› (/ʁ/), each represented by three, three, and two words, respectively. All the selected phoneme-grapheme correspondences in this category are consistent, that is, only the target digraph could be used to correctly spell them. Examples are ‹ta**lh**o› *butchery* /ˈtɐ.ʎu/, ‹ba**nh**o› *bath* /ˈbɐ.ɲu/, and ‹de**rr**ota› *defeat* /də.ˈʁɔ.tɐ/. As noted in Introduction, the case of /ʁ/ spelled as ‹rr› could also be framed in position consistencies (/ʁ/ is ‹r› in word initial position, and ‹rr› between vowels). However, we included ‹rr› in the digraph category as it seems to be its most salient property, determining specific rules to be followed by children (e.g., separating the two ‹r› in translineation). This is also reflected in the Portuguese language school curriculum, where ‹rr› is specified as a digraph (see Buescu et al., 2015).

Contextual consistency

This complex consistency refers to cases where the selection of the grapheme depends on which vowel comes after the target phoneme. Words in this category included the phonemes /k/ or /g/ represented by ‹qu› or ‹gu›, respectively, before the vowels ‹e› or ‹i› (each representing /e/ or /ə/, or /i/ or /j/, respectively, depending on the word). We considered two words per each of the following orthographic representations: ‹que› (e.g., ‹tan**que**› *tank* /ˈtɐ̃.kə/), ‹qui› (e.g., ‹e**qui**pa› *team* /i.ˈki.pɐ/), ‹gue› (e.g., ‹fo**gue**te› *rocket* /fu.ˈge.tə/), and ‹gui› (e.g., ‹**gu**iador› *handlebar* /gjɐ.ˈdoɾ/).

Position consistency

This refers to cases where the selection of the grapheme depends on where the target phoneme occurs within the word or the syllable, namely initial vs. coda positions. Stimuli in this category represented one of the following cases: ‹r› in initial position (e.g., ‹**r**ima› *rhyme* /ˈʁi.mɐ/), ‹r› in coda position (e.g., ‹co**r**da› *row* /ˈkɔɾ.dɐ/), ‹s› in coda position (e.g., ‹pa**s**ta› *folder* /ˈpaʃ.tɐ/), or ‹m› in coda position before /p/ (e.g., ‹ta**m**pa› *lid* /ˈtɐ̃.pɐ/). We included two words per case.

Consonant cluster

A consonant cluster refers to consecutive consonants at the onset of a syllable (CCV syllable). In this category, the second consonant of the cluster was /ɾ/ or /l/. We included the following target spellings: ‹dr›, ‹fr›, ‹gr›, ‹pr›, ‹tr›, ‹bl›, ‹cl›, and ‹pl›. The cluster appeared in the beginning or in middle of the word (three and five items for each case, respectively).

Stress mark

Portuguese has three diacritics that mark stress and indicate the value of the vowel. We included three words with an acute accent (e.g., ‹j**ú**ri› *jury* /ˈʒu.ɾi/), three words

with a circumflex accent (e.g., ‹p**ê**ndulo› *pendulum* /ˈpẽ.du.lu/) and two words with a tilde (e.g., ‹capit**ã**o› *captain* /kɐ.pi.ˈtɐ̃w/).

Inconsistency

An inconsistency occurs when a phoneme/grapheme mapping is not predictable on the basis of contextual rules. To represent inconsistent mappings, we selected words with the phoneme /ʒ/ represented by ‹g› or ‹j› (e.g., ‹**g**ema› *yolk* /ˈʒe.mɐ/, ‹la**j**e› *slab* /ˈla.ʒə/), or the letter ‹x› representing /ʃ/ or /ks/ (‹co**x**o› *lame* /ˈko.ʃu/, ‹ane**x**o› *annex* /ɐ.ˈnɛk.su/).

Silent letter ‹h›

In Portuguese, the letter ‹h› as a single grapheme is used only in word initial position and has no phonemic value. Examples of our stimuli are: ‹hino› *anthem* /ˈi.nu/ and ‹humana› *human* /u.ˈmɐ.nɐ/.

Pseudoword dictation task

This task is composed of 40 pseudowords organized into five orthographic complexity categories, all but inconsistency and silent letter ‹h›. Each category included four two-syllable plus four three-syllable pseudowords. The pseudowords were formed by replacing one to two phonemes of the words in the word dictation task by another phoneme(s) of the same phonetic category (i.e., manner and/or place of articulation). For example, the word ‹talho› *butchery* /ˈ**t**ɐ.ʎu/, from the digraph category, resulted in the pseudoword ‹dalho› /ˈ**d**ɐ.ʎu/ by changing the first phoneme. In this case, both manner and place of articulation were kept, and the voiceless dental stop /t/ was replaced by the voiced homorganic /d/. Another example is the pseudoword ‹jagrado› /ʒɐ.ˈgɾa.du/, which was built from the word ‹sagrado› *sacred* /sɐ.ˈgɾa.du/ by changing place of articulation from alveolar to postalveolar. Using this procedure, we built pseudowords with the same syllable structure, length, and orthographic category as the corresponding words. To keep scoring criteria identical between words and pseudowords—only one correct spelling—we did not create pseudowords for the inconsistent categories (e.g., the pseudoword /ˈto.ʃu/ can be spelled as ‹toxo› or ‹tocho›, because both spellings represent correctly the phonological form).

Procedure

After gathering parental written consent and children's verbal assent, word and then pseudoword dictation tasks were administered to children in their classrooms. Each word and pseudoword was read once by the experimenter, who gave enough time for all children to spell it in a paper sheet. If a child did not understand the target word, it was repeated once.

Each child was tested individually on the background measures (Raven CPM, reading age and WISC-III Vocabulary) in a separate session.

Results

Complexity, length, lexicality, and grade effects

To examine the effects of complexity, length, lexicality, and grade on the spelling performance of primary grade children, we conducted a 5×2×2×3 (Complexity Category [digraph, contextual consistency, position consistency, consonant cluster, stress mark]×Length [two-, three-syllables]×Lexicality [word, pseudoword]×Grade [second, third, fourth]) Analysis of Variance (ANOVA), with repeated measures on the first three factors. As this analysis included multiple comparisons, we adopted a conservative approach by setting the alpha level to .001 to reduce the probability of Type I errors. The average percentage of correctly spelled two- and three-syllable words and pseudowords by complexity category and grade are presented in Table 2.

We found significant main effects of Complexity, Length, Lexicality, and Grade (see Table 3). Moreover, there were significant two-way interactions between Complexity and Grade, between Complexity and Lexicality, and between Complexity and Length. No significant three- or four-way interactions were found. The significant two-way interactions were decomposed with simple effect analyses, further detailed below.

Interaction between complexity and length

Differences between complexity categories

There were differences in children's performance according to complexity for both two-syllable, $F(4, 183) = 253.08$, $p < .001$, $\eta_p^2 = 0.85$, and three-syllable stimuli, $F(4, 183) = 407.82$, $p < .001$, $\eta_p^2 = 0.90$. Pairwise comparisons are detailed in Table 4. With the exceptions of position consistencies and consonant clusters, the pattern of differences across complexities was similar between two- and three-syllable stimuli. For both lengths, digraphs had the highest-accuracy categories (94% and 92% in two-and three-syllable stimuli, respectively), and stress marks the lowest and most discrepant scores in comparison to the remaining categories (44% and 36% in two- and three-syllable stimuli, respectively).

Differences between two- and three-syllable stimuli

Results showed that children performed better in two-syllable than in three-syllable stimuli in the consonant cluster and stress mark categories (length effect of about 8% for both categories), $F(1, 186) = 37.64$, $p < .001$, $\eta_p^2 = 0.17$, and $F(1, 186) = 27.58$, $p < .001$, $\eta_p^2 = 0.13$, respectively. No length effect was found for digraphs, $F(1, 186) = 6.95$, $p = .009$, $\eta_p^2 = 0.04$, contextual consistencies ($F < 1$) nor position consistencies, $F(1, 186) = 1.44$, $p = .23$, $\eta_p^2 = 0.01$.

Table 2 Descriptive statistics for the percentage of correctly spelled words and pseudowords by complexity category, length, and grade

Complexities	Grade 2				Grade 3				Grade 4			
	Words		Pseudowords		Words		Pseudowords		Words		Pseudowords	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Digraph												
Two syllables	93.31	16.88	90.14	20.91	95.92	10.64	94.90	11.39	93.84	13.83	95.29	10.74
Three syllables	86.62	23.08	88.03	22.30	97.96	8.59	92.35	14.62	92.75	14.29	92.03	13.89
Contextual consistency												
Two syllables	82.04	25.07	80.63	28.41	92.35	16.30	90.82	18.18	92.39	13.08	85.87	20.77
Three syllables	82.39	26.52	80.63	29.94	92.35	18.54	89.29	21.65	88.04	21.70	88.77	19.43
Position consistency												
Two syllables	88.03	19.30	80.99	21.32	93.88	12.00	87.76	16.24	96.38	9.85	92.03	13.89
Three syllables	88.03	19.76	83.45	20.24	97.45	7.65	88.78	19.14	93.84	15.10	93.48	13.32
Consonant cluster												
Two syllables	81.69	24.62	82.39	25.14	96.43	10.21	95.92	10.64	97.46	8.73	93.84	13.83
Three syllables	72.54	28.71	80.28	29.55	88.78	20.45	91.84	15.64	83.33	23.35	84.78	23.18
Stress mark												
Two syllables	29.93	20.97	29.93	20.97	48.98	24.45	41.84	20.66	57.25	28.47	55.43	28.40
Three syllables	27.82	16.12	25.70	15.22	37.24	16.24	32.65	17.83	53.62	24.36	41.30	25.66
Inconsistency												
Two syllables	52.11	23.81	–	–	55.61	24.08	–	–	56.16	21.61	–	–
Three syllables	47.54	22.42	–	–	58.16	22.47	–	–	60.14	19.33	–	–
Silent letter ‹h›												
Two syllables	14.44	21.82	–	–	46.43	33.46	–	–	41.67	31.41	–	–
Three syllables	25.35	27.87	–	–	53.06	25.84	–	–	56.88	25.68	–	–

Inconsistency and silent ‹h› categories do not apply to pseudowords

Table 3 Main effects and interactions of the Complexity × Length × Lexicality × Grade analysis of variance

Main effects and interactions	Pillai's V	*F*	*p*	η_p^2
Main effects				
Complexity	.91	453.14	< .001	0.91
Length	.17	38.81	< .001	0.17
Lexicality	.14	30.70	< .001	0.14
Grade	–	14.55	< .001	0.14
Two-way interactions				
Complexity × Length	.22	12.80	< .001	0.22
Complexity × Lexicality	.18	9.98	< .001	0.18
Complexity × Grade	.25	6.57	< .001	0.12
Length × Lexicality	.01	1.30	.26	0.01
Length × Grade	.02	2.00	.14	0.02
Lexicality × Grade	.03	2.87	.06	0.03
Three-way interactions				
Complexity × Length × Lexicality	.08	3.95	.004	0.08
Complexity × Length × Grade	.07	1.63	.11	0.03
Complexity × Lexicality × Grade	.08	1.93	.05	0.04
Length × Lexicality × Grade	.01	1.20	.30	0.01
Four-way interaction				
Complexity × Length × Lexicality × Grade	.12	3.00	.003	0.06

Alpha level set to .001 to minimize Type I errors due to multiple comparisons

Interaction between complexity and lexicality

Differences between complexity categories

Results revealed a complexity effect for words, $F(4, 183) = 374.88$, $p < .001$, $\eta_p^2 = 0.89$, and for pseudowords, $F(4, 183) = 345.10$, $p < .001$, $\eta_p^2 = 0.88$. Pairwise comparisons can be found in Table 4. For both words and pseudowords, here too digraphs had the highest accuracy rates (93% for words and 92% for pseudowords), whereas stress marks had the lowest (42% for words and 38% for pseudowords).

Differences between words and pseudowords

Children were more accurate when writing words than pseudowords in the position consistency and stress mark categories (lexicality effect of about 5% for both categories), $F(1, 186) = 34.61$, $p < .001$, $\eta_p^2 = 0.16$, and $F(1, 186) = 22.15$, $p < .001$, $\eta_p^2 = 0.11$, respectively. No lexicality effect was found for the remaining features: digraphs, $F(1, 186) = 2.10$, $p = .15$, $\eta_p^2 = 0.01$; contextual consistencies, $F(1,$

Table 4 Results of the pairwise comparisons among complexity categories

Comparison	Two-way interaction						
	Complexity × Length		Complexity × Lexicality		Complexity × Grade		
	Two-syllable	Three-syllable	Words	Pseudowords	Grade 2	Grade 3	Grade 4
Digraph (Dg) vs.							
Contextual consistency (CC)	Dg > CC	Dg > CC	Dg > CC	Dg > CC	Dg > CC	*ns*	*ns*
Position consistency (PC)	Dg > PC	*ns*	*ns*	Dg > PC	Dg > PC	*ns*	*ns*
Consonant cluster (Cl)	*ns*	Dg > Cl	Dg > Cl	Dg > Cl	Dg > Cl	*ns*	*ns*
Stress mark (SM)	Dg > SM	Dg > SM	Dg > SM	Dg > SM	Dg > SM	Dg > SM	Dg > SM
CC vs.							
PC	*ns*	*ns*	CC < PC	*ns*	*ns*	*ns*	*ns*
Cl	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
SM	CC > SM	CC > SM	CC > SM	CC > SM	CC > SM	CC > SM	CC > SM
PC vs.							
Cl	*ns*	PC > Cl	PC > Cl	*ns*	PC > Cl	*ns*	*ns*
SM	PC > SM	PC > SM	PC > SM	PC > SM	PC > SM	PC > SM	PC > SM
Cl vs.							
SM	Cl > SM	Cl > SM	Cl > SM	Cl > SM	Cl > SM	Cl > SM	Cl > SM

Significant comparisons at an alpha level of .001

186) = 5.38, $p = .02$, $\eta_p^2 = 0.03$; consonant clusters, $F(1, 186) = 2.63$, $p = .11$, $\eta_p^2 = 0.01$.

Interaction between complexity and grade

Differences between complexity categories

There was an effect of complexity category in Grade 2, $F(4, 183) = 230.63$, $p < .001$, $\eta_p^2 = 0.83$, in Grade 3, $F(4, 183) = 132.94$, $p < .001$, $\eta_p^2 = 0.74$, and in Grade 4, $F(4, 183) = 113.59$, $p < .001$, $\eta_p^2 = 0.71$. Pairwise comparisons are presented in Table 4. In Grades 2, 3 and 4, the lowest accuracy rates were in stress marks (28%, 40%, and 52%, respectively). Performance levels in digraphs, position and contextual consistencies, and consonant clusters were more discrepant in Grade 2 (ranging from 90% in digraphs to 79% in consonant clusters) than in Grades 3 and 4 (from 95% in digraphs to 91% in contextual consistencies, and from 94% in position consistencies to 89% in contextual consistencies, respectively). Third- and fourth-graders had a more homogeneous performance across those four categories than their second-grade peers.

Differences among grades

Results showed grade differences in the following categories: position consistencies, $F(2, 186) = 8.45$, $p < .001$, $\eta_p^2 = 0.08$, with more errors in Grade 2 than in Grade 4; consonant clusters, $F(2, 186) = 10.79$, $p < .001$, $\eta_p^2 = 0.10$, with more errors in Grade 2 than in Grades 3 and 4; and stress mark, $F(2, 186) = 29.79$, $p < .001$, $\eta_p^2 = 0.24$, with more errors in Grade 2 than in Grade 3, and in Grade 3 than in Grade 4. No grade effects were found for digraphs and contextual consistencies, $F(2, 186) = 3.15$, $p = .05$, $\eta_p^2 = 0.03$, and $F(2, 186) = 4.46$, $p = .01$, $\eta_p^2 = 0.05$, respectively.

Complexity, length, and grade effects

Because the inconsistency and silent letter ‹h› categories were only represented in words, these categories were not considered in the previously presented analyses. Thus, to examine whether grade and length affected performance in these word categories, and whether this performance differed from that in the other complexity categories, we conducted an additional 7 × 2 × 3 ANOVA (Complexity [digraph, contextual consistency, position consistency, consonant cluster, stress mark, inconsistency, silent letter ‹h›] × Length × Grade), with repeated measures on the first two factors. In this new analysis, we removed the Lexicality factor and added two extra levels in the Complexity factor. Following the previous analysis, we set the alpha level to .001.

As depicted in Table 5, results revealed significant main effects of Grade and Complexity. Furthermore, there were significant two-way interactions between Grade and Complexity, and between Length and Complexity, which were

Table 5 Main effects and interactions of the Complexity × Length × Grade analysis of variance

Main effects and interactions	Pillai's V	*F*	*p*	η_p^2
Main effects				
Complexity	.92	331.06	< .001	0.92
Length	.02	3.48	.06	0.02
Grade	–	23.48	< .001	0.20
Two-way interactions				
Complexity × Length	.35	16.29	< .001	0.35
Complexity × Grade	.36	6.75	< .001	0.18
Length × Grade	.003	0.25	.78	0.003
Three-way interaction				
Complexity × Length × Grade	.16	2.65	.002	0.08

Alpha level set to .001 to minimize Type I errors due to multiple comparisons

decomposed with simple effects analyses. No other significant interactions were found. To avoid repeating the results from the previous ANOVA, grade and length effects are only presented for the inconsistency and silent letter categories.

Interaction between complexity and length

Differences between complexity categories

Results revealed an effect of the complexity category for two- and three-syllable stimuli, $F(6, 181) = 215.41$, $p < .001$, $\eta_p^2 = 0.88$, and $F(6, 181) = 288.20$, $p < .001$, $\eta_p^2 = 0.91$, respectively. Regardless of word length, children performed more poorly in the silent letter ‹h› and inconsistencies than in digraphs, contextual and position consistencies, and consonant clusters.

Differences between two- and three-syllable words

Results showed a length effect for words in the silent letter ‹h›, $F(1, 186) = 34.07$, $p < .001$, $\eta_p^2 = 0.15$ (accuracy about 11% higher for three-syllable words), but not for the inconsistencies ($F < 1$).

Interaction between complexity and grade

Differences between complexity categories

There was an effect of complexity category in Grade 2, $F(6, 181) = 172.58$, $p < .001$, $\eta_p^2 = 0.85$, in Grade 3, $F(6, 181) = 93.19$, $p < .001$, $\eta_p^2 = 0.76$, and in Grade 4, $F(6, 181) = 89.04$, $p < .001$, $\eta_p^2 = 0.75$. Across grades, children's accuracy was significantly lower in the inconsistency and silent letter ‹h› categories than in digraphs, contextual and position consistencies, and consonant clusters. Additionally,

accuracy in the silent letter ‹h› was similar to the stress mark and inconsistency categories in Grades 3 and 4, but significantly lower than those two categories in Grade 2. Also, second- and third-graders produced more errors in stress marks than in inconsistencies, but fourth-graders performed similarly in the two categories.

Differences among grades

Results showed a grade effect for the silent letter ‹h›, $F(2, 186) = 31.98$, $p < .001$, $\eta_p^2 = 0.26$, with lower performance in Grade 2 than in Grades 3 and 4 that were similar; but not for inconsistencies, $F(2, 186) = 4.06$, $p = .02$, $\eta_p^2 = 0.04$, where there was no progress from Grade 2 to Grade 4.

Discussion

The present study was designed to analyze how orthographic complexities influence learning to spell in Portuguese, an orthography of intermediate depth. We examined the spelling accuracy of 189 children from Grades 2, 3, and 4, who performed dictation tasks composed of words and pseudowords. Words were chosen from a database of stimuli from Portuguese primary textbooks (Grades 1 to 4), and pseudowords were derived from these words. Each stimulus represented an orthographic complexity, assigned to one of seven categories: digraph, contextual consistency, position consistency, consonant cluster, stress mark, inconsistency, and the silent letter ‹h›. By focusing on a larger range of orthographic complexities and primary grade levels, we aimed to fill the gap in existing findings about spelling acquisition in Portuguese primary graders (e.g., Fernandes et al., 2008; Serrano et al., 2011). Specifically, we analyzed the effect of orthographic complexity on spelling accuracy and the moderating role of length (two vs. three-syllable stimuli), lexicality (words vs. pseudowords), and grade (second, third, and fourth). In general, results showed that there were differences in accuracy according to the complexity categories, and those differences were modulated by item length, item lexical status, and grade.

The comparison between different complexities revealed that children's best performance was for digraphs and position consistencies in words, and for digraphs in pseudowords. This result is not surprising because these categories represented entirely consistent orthographic features of Portuguese. Children can rely on unambiguous phoneme/grapheme mappings to correctly spell the phonological forms, no access to lexical knowledge required. However, it should be noted that a lexicality effect occurred for position consistencies (better performance in words than pseudowords). This result suggests that, in addition to knowledge on position-related regularities, children might have used word knowledge to spell accurately. In any case, regardless of the spelling strategy digraphs and position consistencies seem to be mastered in the very initial phases of learning: second graders showed high accuracy rates in both categories (about 90% for digraphs and 85% for position consistencies, words and pseudowords lumped together). These results complement well those observed by Fernandes et al. (2008) with first graders: children were more

accurate in spelling simple or complex regular words (containing one-letter graphemes or digraphs, respectively) than irregular words.

Regarding inconsistencies, the silent letter ‹h› and stress marks, results showed that these orthographic conditions were the most difficult for both beginning and advanced primary graders when compared to the remaining categories. The advantage of consistent over inconsistent categories replicates the regularity effect observed in first graders by Fernandes et al. (2008), and shows that it persists until Grade 4. This advantage is also congruent with instructional practices: consistencies can be taught explicitly through rules that may be memorized and trained by children, but inconsistencies do not. Thus, it is not surprising that children may face difficulties in explaining the underlying principles of inconsistent spellings (Almeida et al., 1998). Fostering reading practices to increment lexical knowledge can help children to master inconsistencies, which are essentially acquired by learning new words. It is therefore possible that examining children's reading habits, either in class or at home, could explain some variability in the inconsistency category. Concerning stress marks, these have already been shown to be a source of difficulty for children speaking other languages, such as Spanish (Defior et al., 2006, 2009) and Greek (Protopapas, Fakou, Drakopoulou, Skaloumbakas, & Mouzaki, 2013).

The examination of our results by item length has showed that length significantly modulated spelling performance in only two categories, consonant clusters and stress marks, where shorter stimuli (two-syllable) were spelled more accurately than longer ones (three-syllable). These results show that the impact of length on spelling is not homogeneous. In consonant clusters, the disadvantage of longer items can be due to the cumulative difficulty imposed by the cluster itself (the phonological representation) and the more complex trisyllabic structure. By adding an epenthetic vowel or omitting one of the consonants of the cluster (common errors), children break down the complex syllable structure turning it into a simpler sequence of CV syllables. In spelling longer (pseudo)words, this syllable simplification could be a useful strategy to handle stimuli with a higher load of information to process (more phonemes and syllables). Poor performance with the stress marks, especially in longer items, may be related not only to failing orthographic knowledge, but also to difficulties in dealing with the prosodic aspects of the language and in the awareness of lexical stress.

The analysis of differences from a grade-level perspective revealed that third and fourth graders performed similarly, and better than second graders. In other words, relevant improvements in spelling abilities occurred mostly from the second to the third grade. The exception was the stress mark category, where a significant gain in performance was also observed from Grades 3 to 4. The growth in spelling abilities beyond Grade 2 was also observed in another study with Portuguese children, where spelling accuracy in a text-to-dictation task improved from Grades 3 to 4 (Horta & Martins, 2004). Nevertheless, our data suggest that the middle stage of primary school is an important landmark for the development of spelling skills, though important learning steps are still to be achieved after the third grade. Specifically, at the end of primary school, the most difficult orthographic complexities have yet to be mastered, as revealed by the mean accuracy levels of fourth graders when writing stimuli with stress marks (about 52%, words and pseudowords together), and

the silent letter ‹h› and inconsistencies (about 49% and 58%, respectively). As noted before, stress marks rely on the articulation between orthographic and prosodic representations, which may be harder to acquire. Teaching stress marks pose additional demands to teachers who need to focus on non-orthographic aspects (e.g., prosody and awareness of lexical stress) and deal with the difficulties exhibited by children in mapping orthography and prosody. These additional demands may contribute to explain the slower acquisition of stress marks along primary school. Even though some teachers may consider adopting individualized approaches that could facilitate learning, these may consume extra time and may not always be feasible.

Limitations and future directions

The findings discussed above should be interpreted taking into account five limitations, which may be used to guide future research. First, as this study has a cross-sectional design, it does not allow direct conclusions about children's development. Future research on spelling acquisition in Portuguese could adopt a longitudinal design across primary school years. Alternatively, a cross-sectional study with an extended range of school grades beyond the fourth grade would also be an important development, as the most difficult orthographic complexities are not yet learned after 4 years of instruction.

Second, it was not possible to collect detailed data about the socioeconomic level of children and about the teaching methods adopted in each class. Future research into the acquisition and development of spelling would benefit from a refined characterization of the social, cultural and economic level of children's families, as well as of teaching methods. Information on these methods may help to explain some of the reported findings.

Third, as we assessed spelling abilities using dictation tasks with pre-defined items, the performance in those tasks may not fully resemble the spelling performance of children in spontaneous writing (e.g., a composition). Nevertheless, as the type of errors given in dictation tasks may overlap with those occurring in spontaneous writing (e.g., children's misspellings of consonant clusters in handwriting compositions; Treiman, 1991), our spelling-to-dictation task remains a valuable tool to assess the spelling abilities of Portuguese beginning writers. Moreover, although we employed a scoring method using a dual criterion (correct vs. incorrect) based on orthographic accuracy, it is possible to use other methods, such as those based on phonological accuracy. The use of more fine-grained analysis of children's spelling, either at the level of phonemes or letters, can also provide relevant information (see Treiman, Kessler, & Caravolas, 2019, for a recent analysis of the power of different scoring methods to predict spelling ability).

Fourth, due to practical reasons, we did not examine the moderating role of frequency, a variable known to impact spelling performance (Fernandes et al., 2008; Sprenger-Charolles, Siegel, & Bonnet, 1998). The inclusion of additional frequency conditions would increase task complexity, namely in number of stimuli, and this would not be feasible to include in children's assessments. Future studies targeting the systematic analysis of frequency and complexity effects

together would provide a deeper characterization of the orthographic properties studied here. To keep assessment feasibility, this could be achieved either by planning a more complex task divided in two parts given to children in separate moments, or by not including length-related conditions.

A final limitation is that our spelling-to-dictation task does not cover all the possible complexities of the Portuguese orthography. For example, we did not directly assess the influence of morphological aspects in spelling. Also, the selected items do not provide an exhaustive representation of the different variations within each orthographic complexity. For instance, although we studied the most common consonant clusters in Portuguese (those with /ɾ/ or /l/ as the second consonant), there are other, less frequent forms that can also be studied (e.g., the clusters with /n/ or /s/ as the second consonant, as in ‹gnomo› *gnome* /ˈgno.mu/ and ‹psicólogo› *psychologist* /psi.ˈkɔ.lu.gu/, respectively; examples from Cunha & Cintra, 2014).

As explicit instruction is beneficial for the development of spelling skills (Graham & Santangelo, 2014), a concrete understanding of how these skills are developed in an orthography is relevant to bring light upon aspects that may merit special attention in teaching. In this regard, our study was an important step to track how children learn both easier and more difficult spelling features of an orthography of intermediate depth.

Conclusion

This study provides a cross-sectional perspective of spelling acquisition in Portuguese. We showed that complex consistencies (digraphs and context-dependent consistencies) and consonant clusters do not seem to pose difficulties for most beginning writers as early as Grade 2; across grades, these complexities were successfully mastered by at least four out of five children. Greater improvements occurred mostly from Grade 2 to Grade 3, but by the end of Grade 4 stress marks and inconsistencies were still in need of further learning and experience to be efficiently mastered. As students benefit from systematic instruction of spelling (Graham & Santangelo, 2014), our focus in how spelling abilities develop in primary grades in the context of Portuguese orthography may be informative for adapting learning and teaching practices in early school grades. Furthermore, our study is important for cross-linguistic research. By providing data on spelling acquisition in relation to grade and orthographic complexity in European Portuguese, an orthography of intermediate depth, the present study paves the way for comparisons with other orthographies varying in degree of consistency and inconsistency.

Acknowledgements This work was supported by Grants CPUP UID/PSI/00050/2013 and SFRH/BPD/65831/2009 from the Portuguese Foundation for Science and Technology. We are thankful to Isabel Santos and to the participating schools, teachers and children.

Appendix

Words and pseudowords used in the spelling task by complexity category and length

Category	Length			
	Two syllables		Three syllables	
	Word	Pseudoword	Word	Pseudoword
Dg	talho /ˈtɐʎu/	dalho /ˈdɐʎu/	molhado /muˈʎadu/	nolhado /nuˈʎadu/
	parra /ˈpaʀɐ/	tarra /ˈtaʀɐ/	agulha /ɐˈguʎɐ/	aculha /ɐˈkuʎɐ/
	linha /ˈliɲɐ/	zinha /ˈziɲɐ/	tamanho /tɐˈmɐɲu/	gamanho /gɐˈmɐɲu/
	banho /ˈbɐɲu/	panho /ˈpɐɲu/	derrota /dəˈʀɔtɐ/	terroda /tɔˈʀɔdɐ/
CC	tanque /ˈtɐ̃k(ə)/	danque /ˈdɐ̃k(ə)/	jaqueta /ʒɐˈketɐ/	zaqueta /zɐˈketɐ/
	quilo /ˈkilu/	quifo /ˈkifu/	equipa /iˈkipɐ/	equima /iˈkimɐ/
	Guida /ˈgidɐ/	guiba /ˈgibɐ/	guiador /gjɐˈdoɾ/	guiator /gjɐˈtoɾ/
	sangue /ˈsɐ̃g(ə)/	jangue /ˈjɐ̃g(ə)/	foguete /fuˈget(ə)/	voguete /vuˈget(ə)/
PC	rima /ˈʀimɐ/	rina /ˈʀinɐ/	retina /ʀəˈtinɐ/	redila /ʀəˈdilɐ/
	corda /ˈkɔrdɐ/	gorta /ˈgɔrtɐ/	perfume /pərˈfum(ə)/	bermufe /bərˈmuf(ə)/
	pasta /ˈpaʃtɐ/	dasca /ˈdaʃkɐ/	estufa /(i)ʃˈtufɐ/	escufa /(i)ʃˈkufɐ/
	tampa /ˈtɐ̃pɐ/	sampa /ˈsɐ̃pɐ/	campino /kɐ̃ˈpinu/	tampino /tɐ̃ˈpinu/
Cl	fruta /ˈfrutɐ/	frata /ˈfratɐ/	abutre /ɐˈbutɾ(ə)/	aditre /ɐˈditɾ(ə)/
	drama /ˈdʀɐmɐ/	drana /ˈdʀɐnɐ/	sagrado /sɐˈgradu/	jagrado /jɐˈgradu/
	dupla /ˈduplɐ/	tupla /ˈtuplɐ/	teclado /tɛˈkladu/	peclato /pɛˈklatu/
	prego /ˈprɛgu/	preco /ˈprɛku/	tablete /tɐˈblɛt(ə)/	gablete /gɐˈblɛt(ə)/
SM	fértil /ˈfɛɾtiɬ/	vértil /ˈvɛɾtiɬ/	pêndulo /ˈpɐ̃dulu/	têngulo /ˈtɐ̃gulu/
	júri /ˈʒuɾi/	zúri /ˈzuɾi/	ânimo /ˈɐnimu/	âmino /ˈɐminu/
	condão /kõˈdɐ̃w/	pondão /põˈdɐ̃w/	último /ˈuɬtimu/	úldimo /ˈuɬdimu/
	maltês /maɬˈteʃ/	naltês /naɬˈteʃ/	capitão /kɐpiˈtɐ̃w/	gapidão /gɐpiˈdɐ̃w/
In	gema /ˈʒemɐ/	–	fugido /fuˈʒidu/	–
	laje /ˈlaʒ(ə)/	–	tigela /tiˈʒɛlɐ/	–
	giro /ˈʒiɾu/	–	lojista /luˈʒiʃtɐ/	–
	coxo /ˈkoʃu/	–	anexo /ɐˈnɛksu/	–
SH	hino /ˈinu/	–	hípico /ˈipiku/	–
	hiper /ˈipɛɾ/	–	Helena /iˈlenɐ/	–
	haste /ˈaʃt(ə)/	–	humilde /uˈmiɬd(ə)/	–
	hiena /iˈenɐ/	–	humana /uˈmɐnɐ/	–

Dg digraph, *CC* contextual consistency, *PC* position consistency, *Cl* consonant cluster, *SM* stress mark, *In* inconsistency, *SH* silent letter ‹h›

References

Almeida, M. M., Guerreiro, M., & Mata, L. (1998). O desenvolvimento de competências ortográficas e as interações sociais [The development of orthographic competencies and social interactions]. *Análise Psicológica, 2,* 321–329.

Buchwald, A., & Rapp, B. (2003). The orthographic representation of consonant-vowel status: Evidence from two cases of acquired dysgraphia. *Brain and Language, 87,* 120–121. https://doi.org/10.1016/S0093-934X(03)00228-1.

Buescu, H. C., Morais, J., Rocha, M. R., & Magalhães, V. F. (2015). *Programa e metas curriculares de português do ensino básico* [Program and curriculum goals of Portuguese in basic education]. Retrieved from the Directorate-General for Education of the Portuguese Ministry of Education and Science website: https://www.dge.mec.pt/sites/default/files/Basico/Metas/Portugues/pmcpeb_julho_2015.pdf.

Carrillo, M. S., Alegría, J., & Marín, J. (2013). On the acquisition of some basic word spelling mechanisms in a deep (French) and a shallow (Spanish) system. *Reading and Writing: An Interdisciplinary Journal, 26,* 799–819. https://doi.org/10.1007/s11145-012-9391-6.

Castro, S. L., & Gomes, I. (2000). *Dificuldades de aprendizagem da língua materna [Language difficulties in children]*. Lisboa: Universidade Aberta.

Cunha, C., & Cintra, L. (2014). *Nova gramática do português contemporâneo [New grammar of contemporary Portuguese]* (21st ed.). Lisboa: Edições João Sá da Costa.

Defior, S., Jiménez Fernández, G., & Serrano, F. (2006). Spelling acquisition: A transversal study of Spanish children. *International Journal of Learning, 12,* 293–299. https://doi.org/10.18848/1447-9494/CGP/v12i10/48218.

Defior, S., Jiménez-Fernández, G., & Serrano, F. (2009). Complexity and lexicality effects on the acquisition of Spanish spelling. *Learning and Instruction, 19,* 55–65. https://doi.org/10.1016/j.learninstruc.2008.01.005.

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2019). *Ethnologue: Languages of the world* (22nd ed.). Retrieved August 1, 2019, from https://www.ethnologue.com/.

Fernandes, S., Ventura, P., Querido, L., & Morais, J. (2008). Reading and spelling acquisition in European Portuguese: A preliminary study. *Reading and Writing: An Interdisciplinary Journal, 21,* 805–821. https://doi.org/10.1007/s11145-007-9093-7.

Gomes, I. (2001). *Ler e escrever em português europeu [Reading and writing in European Portuguese] (Unpublished doctoral dissertation)*. Porto: Faculty of Psychology and Education Sciences, University of Porto.

Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing: An Interdisciplinary Journal, 27,* 1703–1743. https://doi.org/10.1007/s11145-014-9517-0.

Horta, I. V., & Martins, M. A. (2004). Desenvolvimento e aprendizagem da ortografia: Implicações educacionais [Orthography development and learning: Educational implications]. *Análise Psicológica, 1,* 213–223. https://doi.org/10.14417/ap.142.

Juul, H., & Petersen, D. K. (2017). Length effects in pseudo-word spelling: Stronger in dyslexic than in non-dyslexic students. *Annals of Dyslexia, 67,* 369–382. https://doi.org/10.1007/s11881-017-0149-3.

Kahn-Horwitz, J., Schwartz, M., & Share, D. (2011). Acquiring the complex English orthography: A triliteracy advantage? *Journal of Research in Reading, 34,* 136–156. https://doi.org/10.1111/j.1467-9817.2010.01485.x.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In L. Katz & R. Frost (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 67–84). Amsterdam: Elsevier Science Publishers.

Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1–33). Ann Arbor, MI: University of Michigan Press.

Lima, C., & Castro, S. L. (2010). Reading strategies in orthographies of intermediate depth are flexible: Modulation of length effects in Portuguese. *European Journal of Cognitive Psychology, 22,* 190–215. https://doi.org/10.1080/09541440902750145.

Martins, M. A., Albuquerque, A., Salvador, L., & Silva, C. (2013). The impact of invented spelling on early spelling and reading. *Journal of Writing Research, 5,* 215–237. https://doi.org/10.17239/jowr-2013.05.02.3.

Martins, M. A., Salvador, L., Albuquerque, A., & Silva, C. (2016). Invented spelling activities in small groups and early spelling and reading. *Educational Psychology, 36,* 738–752. https://doi.org/10.1080/01443410.2014.950947.

Mateus, M. H., & D'Andrade, E. (1998). The syllable structure in European Portuguese. *Documentação de Estudos em Lingüística Teórica e Aplicada [Documentation of Studies in Theoretical and Applied Linguistics], 14*(1), 13–32. https://doi.org/10.1590/S0102-44501998000100002.

Moutinho, L. (2000). *Uma introdução ao estudo da fonética e fonologia do português [An introduction to the study of phonetics and phonology of Portuguese]*. Lisboa: Plátano Edições Técnicas.

Pinheiro, A. (1995). Reading and spelling development in Brazilian Portuguese. *Reading and Writing: An Interdisciplinary Journal, 7,* 111–138. https://doi.org/10.1007/BF01026950.

Protopapas, A., Fakou, A., Drakopoulou, S., Skaloumbakas, C., & Mouzaki, A. (2013). What do spelling errors tell us? Classification and analysis of errors made by Greek schoolchildren with and without dyslexia. *Reading and Writing: An Interdisciplinary Journal, 26,* 615–646. https://doi.org/10.1007/s11145-012-9378-3.

Schmalz, X., Marinus, E., Coltheart, M., & Castles, A. (2015). Getting to the bottom of orthographic depth. *Psychonomic Bulletin & Review, 22,* 1614–1629. https://doi.org/10.3758/s13423-015-0835-2.

Serrano, F., Genard, N., Sucena, A., Defior, S., Alegria, J., Mousty, P., et al. (2011). Variations in reading and spelling acquisition in Portuguese, French and Spanish: A cross-linguistic comparison. *Journal of Portuguese Linguistics, 10*(1), 183–204. https://doi.org/10.5334/jpl.106.

Seymour, P. H. K., Aro, M., Erskine, J. M., & The COST Action A8 network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94,* 143–174. https://doi.org/10.1348/000712603321661859.

Simões, M. R. (2000). *Investigações no âmbito da aferição nacional do teste das Matrizes Progressivas Coloridas de Raven (M.P.C.R.)* [Research in the context of the national adaptation of the Raven's Colored Progressive Matrices test (R.C.P.M.)]. Lisboa: Fundação Calouste Gulbenkian/Fundação para a Ciência e a Tecnologia.

Simões, M., Rocha, A. M., & Ferreira, C. (2003). *Escala de Inteligência de Wechsler para Crianças – Terceira Edição [Wechsler Intelligence Scale for Children – Third Edition] [Measurement instrument]*. Lisboa: CEGOC-TEA.

Sim-Sim, I., Duarte, I., & Ferraz, M. J. (1997). *A língua materna na educação básica: Competências nucleares e níveis de desempenho [The native language in basic education: Core competencies and performance levels]*. Lisboa: Ministério da Educação.

Sprenger-Charolles, L., Siegel, L. S., & Bonnet, P. (1998). Reading and spelling acquisition in French: The role of phonological mediation and orthographic factors. *Journal of Experimental Child Psychology, 68,* 134–165. https://doi.org/10.1006/jecp.1997.2422.

Sucena, A., & Castro, S. L. (2008). *Aprender a ler e avaliar a leitura. O TIL: Teste de Idade de Leitura [Learning how to read and the assessment of reading]*. Coimbra: Almedina.

Sucena, A., Castro, S. L., & Seymour, P. (2009). Developmental dyslexia in an orthography of intermediate depth: The case of European Portuguese. *Reading and Writing: An Interdisciplinary Journal, 22,* 791–810. https://doi.org/10.1007/s11145-008-9156-4.

Tainturier, M.-J., & Caramazza, A. (1996). The status of double letters in graphemic representations. *Journal of Memory and Language, 35,* 53–73. https://doi.org/10.1006/jmla.1996.0003.

Tainturier, M. J., & Rapp, B. C. (2004). Complex graphemes as functional spelling units: Evidence from acquired dysgraphia. *Neurocase, 10,* 122–131. https://doi.org/10.1080/13554790490497238.

Teixeira, C., & Castro, S. L. (2006). *Portulex: Uma base de dados lexical do vocabulário escrito do 1.º ao 4.º anos do ensino básico* [Portulex: A lexical database of written vocabulary for the 1st to 4th grades]. Paper presented at the VI National Symposium of Research in Psychology, Évora, Portugal.

Treiman, R. (1991). Children's spelling errors on syllable-initial consonant clusters. *Journal of Educational Psychology, 83,* 346–360. https://doi.org/10.1037/0022-0663.83.3.346.

Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. New York, NY: Oxford University Press.

Treiman, R., Kessler, B., & Caravolas, M. (2019). What methods of scoring young children's spelling best predict later spelling performance? *Journal of Research in Reading, 42,* 80–96. https://doi.org/10.1111/1467-9817.12241.

Vale, A. P., & Sousa, J. (2017). Tipo de erros e dificuldades na escrita de palavras de crianças portuguesas com dislexia [Type of errors and difficulties in word spelling of Portuguese children with dyslexia]. *Da Investigação às Práticas, 7*(3), 61–83. https://doi.org/10.25757/invep.v7i3.141.

Van Bon, W. H. J., & Uit de Haag, I. J. C. A. F. (1997). Difficulties with consonants in the spelling and segmentation of CCVCC pseudowords: Differences among Dutch first graders. *Reading and Writing: An Interdisciplinary Journal, 9,* 363–386. https://doi.org/10.1023/A:1007942911807.

Vigário, M., & Falé, I. (1994). A sílaba no português fundamental: Uma descrição e algumas considerações de ordem teórica [The syllable in fundamental Portuguese: A description and some theoretical considerations]. In Associação Portuguesa de Linguística (Ed.), *Actas do IX Encontro da Associação Portuguesa de Linguística* (pp. 465–478). Lisboa: Colibri.

Willson, V. L., Rupley, W. H., Rodriguez, M., & Mergen, S. (1999). The relationships among orthographic components of word identification and spelling for grades 1–6. *Literacy Research and Instruction, 39,* 89–102. https://doi.org/10.1080/1938807990955.