doi: http://dx.doi.org/10.5007/1984-8412.2023.e85764

THE BRAZILIAN PORTUGUESE DIGRAPHS IN THE SPELLING OF STUDENTS OF THE FIRST YEARS OF ELEMENTARY SCHOOL

OS DÍGRAFOS DO PORTUGUÊS BRASILEIRO NA ESCRITA DE ALUNOS DOS ANOS INICIAIS DO ENSINO FUNDAMENTAL

LOS DÍGRAFOS DEL PORTUGUÉS BRASILEÑO EN LA ESCRITURA DE ALUMNOS DE LOS PRIMEROS AÑOS DE LA ESCUELA PRIMARIA

Ana Ruth Moresco Miranda*
Lissa Pachalski**
Lorenzo Steinhorst Richetti***
Federal University of Pelotas

ABSTRACT: This paper aims to describe and analyze the spelling of the digraphs of the Brazilian Portuguese orthographic system by children in the early years of Elementary School. For that, the (ortho)graphic error, the main object of analysis, is discussed considering a qualitative and quantitative-descriptive treatment, according to the variables' type of school (public and private), grade, type of digraph, and nature of the error. The results show that there is a decrease in errors as the grades advance and that there is a higher incidence of errors in the public schools. However, similarities are observed in relation to the quality of the error produced in both schools. Children give different treatment to digraphs according to the type of complexity involved in the phoneme-grapheme relationship. The digraphs that seem to be the greatest object of doubt are the heterosyllabic ones, considering the general error rates in relation to correct spellings and their almost categorically orthographic nature. Repercussions for teaching are discussed.

KEYWORDS: Digraphs. Literacy. Orthographic errors. Phonology and orthography.

* PhD in Linguistics and Letters from PUC-RS. Professor at UFPel. CNPq Research Productivity Scholar. Coordinator of GEALE (CNPq Research Group/UFPel). E-mail: anaruthmmiranda@gmail.com.

** PhD Student and Master in Languages and Literatures at UFPel. Degree in Education (Elementary Teaching License) from UFPel. Member of GEALE (CNPq Research Group/UFPel). E-mail: pachalskil@gmail.com.

*** Undergraduate in Education (Elementary Teaching License) at UFPel. CNPq Junior Researcher Scholar at GEALE (CNPq Research Group/UFPel). E-mail: lorenzo.richetti@gmail.com.

RESUMO: Este artigo busca descrever e analisar a grafia dos dígrafos do sistema ortográfico do português brasileiro (PB) por crianças dos anos iniciais do Ensino Fundamental. Para tanto, o erro (orto)gráfico é tido como objeto principal de análise, sendo discutido considerando-se um tratamento qualitativo e quantitativo-descritivo, segundo as variáveis tipo de escola (pública e particular), série, tipo de dígrafo e natureza de erro. Os resultados mostram que há decréscimo de erros à medida que avançam as séries e que há maior incidência de erros na escola pública. No entanto, são observadas similaridades em relação à qualidade do erro produzido em ambas as escolas. As crianças dão tratamento diferenciado aos dígrafos de acordo com o tipo de complexidade envolvida na relação fonema-grafema. Os dígrafos que aparentam ser maior objeto de dúvida são os heterossilábicos, considerando-se os índices gerais de erro em relação aos acertos e sua natureza quase categoricamente ortográfica. Repercussões para o ensino são discutidas.
PALAVRAS-CHAVE: Dígrafos. Aquisição da escrita. Erros ortográficos. Fonologia e ortografia.

RESUMEN: Este artículo busca describir y analizar la ortografía de los dígrafos del sistema ortográfico portugués brasileño (BP) por niños en los primeros años de la Enseñanza Fundamental. Para ello, se toma como principal objeto de análisis el error (orto)gráfico, siendo discutido considerando un tratamiento cualitativo y cuantitativo- descriptivo, según las variables tipo de escuela (pública y privada), grado, tipo de dígrafo y naturaleza del error. Los resultados muestran que hay una disminución de errores a medida que avanzan los grados y que hay una mayor incidencia de errores en las escuelas públicas. Sin embargo, se observan similitudes en relación a la calidad del error producido en ambas escuelas. Los niños dan un tratamiento diferente a los dígrafos según el tipo de complejidad implicada en la relación fonema-grafema. Los dígrafos que parecen ser objeto de mayor duda son los heterosilábicos, considerando los índices generales de error en relación con las respuestas correctas y su carácter casi categóricamente ortográfico. Se discuten también las repercusiones para la enseñanza.
PALABRAS-CLAVE: Dígrafos. Adquisición de la escritura. Errores ortográficos. Fonología y ortografía.

1 INTRODUCTION

The acquisition[1] of an alphabetic writing system involves the structuring and organization of knowledge of different natures. After dealing with the non-trivial task of understanding the alphabetic principle – graphemes represent phonemes – other challenges need to be solved by the learner, especially since there are few systems that keep most of the relationships transparent or biunivocal between phonemes and graphemes, as is the case with Finnish, for example. In general, as in Brazilian Portuguese (BP), there is greater opacity in these relationships, also called multiples, which are characterized by displaying more than one option to relate the phoneme to the grapheme, as well as the grapheme to the phoneme. Examples from Portuguese are the spelling of /s/ with the graphemes <s> and <c>, in words such as "seta" (arrow) and "cebola" (onion), or the use of the grapheme <x> that in words, such as "caixa" (box) and "extenso" (extense), represents different phonemes, a /ʃ/ and an /s/, respectively. In these cases, the spelling choice may or may not be determined by the context. This type of criterion, which defines levels of transparency and opacity based on the number of graphic options for each phoneme in a language, is known as the *orthographic depth* criterion (SEYMOUR; ARO; ERSKINE, 2003) and is related to orthographic complexities themselves, in the sense that there are rules of an arbitrary and/or contextual nature that regulate the relationships between phonemes and graphic options.

[1] The reader will notice the deliberate use of the expression *writing acquisition* in this article. Perhaps one wonders why the expression *learning to spell* is not being used, since it is a process that does not occur naturally and spontaneously, as is the case with speech acquisition, from an innatism point of view. It is understood, however, that this split between the acquisition and learning processes – the first associated with spontaneity and the second with instruction –, so necessary to the first studies with an innate basis, needs to be reviewed. Contribute to blurring the boundary between the terms acquisition and learning the book Psicogênese da Língua Escrita [The Psychogenesis of Written Language] (Ferreiro; Teberosky, 1999[1984]) and the Representational Redescription Model (according to Karmiloff-Smith, 1994), as well as many other studies developed within the scope of genetic epistemology, as they refer to the role of the subject in the construction of knowledge, that is, an active, biologically potent subject, a knowing subject, and also the loosening of the idea of strict modularity, as proposed by Fodor. The studies developed by GEALE opt, therefore, for the term *acquisition*, as they understand that the fruitful dialogue between innatism and constructivism, established over the last thirty years, is best translated by this term. In addition, the idea that spelling acquisition integrates the broader process of language acquisition is defended, with speech and writing being understood as alternative substances in the actualization of the language, even though, without a doubt, they require different processes for them to be acquired (MIRANDA, 2017).

Another complicating factor involved in the acquisition of orthography may be phonological, related to the representational complexities of certain units of the phonological grammar, from the internal structure of a segment or syllable to the level of the accent or prosodic word (MIRANDA, 2017, 2020). Thus, the spelling of these structures can also be a challenge in the acquisition process, even though, from the orthographic point of view, it involves a transparent phonographemic relationship, that is, the orthographic complexity is absent. An example is the representation of the palatal sonorants /ʎ/ and /ɲ/, complex segments from the point of view of their internal structure, but which, in BP orthography, present a biunivocal relationship between the phoneme-grapheme. Words with phonemes /ʎ/ and /ɲ/, such as "pa**lh**a" (straw) and "ma**nh**a" (dexterity), preferentially select the graphemes <lh> and <nh> in the orthographic system[2] (MIRANDA, 2014).

In the case of palatal sonorants, in addition to the inherent phonological complexity, another relatively common complexity in different writing systems is associated, which poses challenges for orthographic development: it is the representation of a phoneme by a grapheme composed of more than one letter, <nh> and <lh>. In Portuguese, in addition to palatal sonorants, the digraphs <ch> for /ʃ/ in "**ch**ato" (boring) or <rr> for /x/ in "ca**rr**o" (car), for example, make up this group. Digraphs constitute a complicating factor in the development of orthographic writing and do not have extensive descriptions in the literature, especially considering the acquisition of BP orthography, nor do they seem to fit directly into the two most common categories of complexities mentioned above, namely, orthographic (multiple relationships between phonemes and graphemes) and phonological (prosodic or segmental complexity).

In this context, the following questions arise: What would be the nature of the complexity of digraphs? Do they all behave similarly in the writings produced by children in their early years, simply because they are digraphs? The errors that children produce in the spelling of digraphs are motivated by what type(s) of knowledge(s)? What can the data reveal about what children know about these graphemes at different periods of orthographic development?

Motivated by such questions, this article[3] has the consonant digraphs of the Brazilian Portuguese orthographic system as a research topic, seeking to contribute to its description, especially from the point of view of spelling acquisition. To this end, the (ortho)graphic error, according to Miranda (2020), is taken as the main object of analysis. The term (ortho)graphic, with the use of parentheses isolating the composition element 'ortho-', seeks to demarcate the two main types of knowledge mobilized by children in the process of spelling acquisition, which also give rise to the main categories for interpretation of errors: one anchored in phonology and the other in orthography. In addition to considering orthographic and phonological aspects to explain errors, given their nature, the author also cites a third source capable of elucidating some writings produced by children: the phonographic source. This category, which will also be considered in this article, emerged from the analyses carried out by the Research Group and concerns spellings in which problems are observed in the tracing and sequencing of letters and syllables in contexts that do not imply phonological or orthographic complexity.

2 THE FUNCTIONING OF DIGRAPHS IN THE BRAZILIAN PORTUGUES (BP) ORTHOGRAPHIC SYSTEM

The treatment of digraphs in the Portuguese literature has been restricted to studies that describe the functioning of the orthographic system (LUFT, 1991; LEMLE, 2001; MORAIS, 2006; FARACO, 2015). The following table provides a summary of the digraphs and their functioning in BP according to the presentation of the aforementioned authors, especially Faraco (2015):

[2] In the case of palatal liquid, there is competition between the graphemes <lh>, <li> and <le> in a small group of words in which the palatal is pronounced, but the spelling is <li> and <le>, as in *família (family)*, *italiano (Italian)*, *auxílio (assistance)* and *óleo (oil)*, for example (MIRANDA, 2020).

[3]This article is supported by CNPq - Processes no. 312387/2020-2 and no. 423038/2021-4; Process 23110.012490/2021-88. We thank the team involved in data collection and research review, especially Nathalia Vitória Reinehr.

digraph	phoneme	relationship	context	competition	examples
<**qu**>	/k/	contextual	followed by /i/, /e/, /ɛ/	-	química (chemistry), querer (to want), quero (I want)
<**gu**>	/g/	contextual	followed by /i/, /e/, /ɛ/	-	guia (guide), gueto (ghetto), guerra (war)
<**rr**>	/x/	contextual	between vowels	-	carro (car)
<**ss**> <**sc/sç**> <**xc**> <**xs**>	/s/	contextual / arbitrary	between vowels	with each other and with the graphemes <c/ç>, <x>	gesso (plaster), nascer/nasça (to be born), exceção (exception), exsudar (to exude)
<**ch**>	/ʃ/	arbitrary	representing /ʃ/ in syllable initial position and between vowels	<x>	chave (key), achar (to find)
<**lh**>	/ʎ/	biunivocal	/ʎ/ between vowels; does not occupy the beginning of the word	, <le>	malha (mesh), lhama (llama), família (family), óleo (oil)
<**nh**>	/ɲ/	biunivocal	/ɲ/ between vowels; does not occupy the beginning of the word	–	manha (dexterity), nhoque (gnocchi)

Table 1: Characterization of consonant digraphs in the Brazilian Portuguese orthographic system
Source: own preparation

As can be seen in Table 1, in relation to the context of use, with the exception of the digraphs that represent the dorsal /k/ and /g/ and the voiceless palatal fricative /ʃ/, the others are restricted to the intervocalic position. The dorsal spellings are regulated by rules and are in a complementary distribution with the graphemes <c> and <g>, as shown in the following table:

phoneme	grapheme	context	example
/k/	<qu> <c>	before /ɛ, e, i/ before /a, ɔ, o, u/	quero (I want), queijo (cheese), quilo (kilo), casa (house), cola (glue), couve (kale), cuia (gourd)
/g/	<gu> <g>	before /ɛ, e, i/ before /a, ɔ, o, u/	guerra (war), gueixa (geisha), guiso (stew), garra (claw), gola (collar), goma (starch), gula (gluttony)

Table 2: Complementary distribution in the spelling of the phonemes /k/ and /g/.
Source: own preparation.

This complementary distribution, which offered sufficient information to the reader, as the presence of <qu> and <gu> before vowels [-posterior] indicated the presence of [k] and [g], underwent a shock after the regulation of the Orthographic Agreement of 1990 (BRAZIL, 2008). The umlaut, diacritical mark superimposed on the vowel to indicate that it should be pronounced as [kw] and [gw], in words like e "tranquilo" (calm) and "linguística" (linguistics), for example; or to change the vowel sound, as in the case of foreign words, normally indicating roundness of the lips to coronal vowels, as in "müller – mülleriano" (müller - müllerian), [y] remained only in the latter case. This change brought opacity to the reading by suppressing a rule that lent greater transparency to the system.

Still, in relation to the functioning of the digraphs for the dorsal consonants, it is observed that they share with those of the sonorants and the palatal fricatives, <lh>, <nh>, <ch> the same behavior related to translineation (breaking of lines), since all remain univocal, unlike what happens with the digraphs of the multiple vibrant, <rr>, (produced in most Brazilian dialects as a velar fricative) and the alveolar fricative, <ss>, <sç>, <sc>, <xc>, <xs>, which must be divided, one letter occupying the post-vocalic position and another, the initial position of the following graphic syllable. In Table 3, symmetry and asymmetry are observed between the phonic and graphic planes, in relation to the syllable, the basic prosodic unit:

examples	phonological syllables	graphic syllables
par**qu**e (park)	/paɾ.**k**e/	par-**qu**e
man**gu**e (mangrove forest)	/maN.**g**e/	man-**gu**e
fo**lh**a (leave, sheet)	/fo.**ʎ**a/	fo-**lh**a
fro**nh**a (pillowcase)	/fro.**ɲ**a/	fro-**nh**a
col**ch**a (quilt)	/kol.**ʃ**a/	col-**ch**a
morro (hill)	/mo.**x**o/	mo**r-r**o
passa (pass)	/pa.**s**a/	pa**s-s**a

nasce (be born)	/na.**s**e/	na**s**-ce
cresça (grow)	/kre.**s**a/	cre**s**-**ç**a
exceder (to exceed)	/e.**s**e.der/	e**x**-**c**e-der
exsudar (to exude)	/e.**s**u.dar/	e**x**-**s**u-dar

Table 3: Relations between phonological syllables and graphic syllables
Source: own preparation

The digraph for 'strong-r' has a specific context of use, the intervocalic position, and there is no competition in the orthographic system. The multiple vibrating phoneme, r- strong, is manifested with a word-initial <r> ("rato" – rat) and following a medial coda consonant ("melro" – blackbird), or with <rr> in the intervocalic context. A distribution that, despite the multiple relationships between phoneme and graphemes, is strictly regulated by the context.

The digraphs of the alveolar fricative /s/ establish the point of greatest opacity in the orthographic system of BP, thanks to the conjunction between the multiplicity of the phoneme-grapheme relationship and the availability of a dozen graphic options. In Table 1, in the column concerning the type of relationship, there are *contextual and arbitrary* specifications. What is defined as contextual is the conditioning of the digraph to the intervocalic context, an important identification so that there can be a reduction in the number of options observed for the spelling of /s/. It is also contextual the selection of <c> or cedilla <ç>, given the complementary distribution also observed between both, since the former has the vowels /ɛ, e, i/ in its sequence and the latter, /a, ɔ, o, u/. Digraphs with 'x', <xc>, and <xs>, in turn, are context-conditioned and are only used after <e> at the beginning of a word. In addition to this possible restriction of use through the observation of contexts, there is other morphological information that helps in the choice of the digraph, as shown by Luft (1991) when establishing a relationship between stems and derived words, such as ced – cess (ceder – cessão (assign - assignment)), prim – press (imprimir – impressão (print - print)), for example, or even to explain etymological reasons in words with erudite entries such as *florescere* (Latin) - florescer (to flourish).

The digraphs with H – <lh>, <nh>, and <ch> – have in common the fact that they represent the palatal consonants of Portuguese, which were not part of the Latin inventory and diachronically derived from sequences of consonants or from consonants and vowels. In the case of sonorants, according to Silva (1996), the etymology of the nasal is the sequence <ni>, while the lateral palatal comes from more varied sequences such as <li>, <lli>, <cl>, <gl> and <pl> by broad palatalization processes. The palatal fricative spelled with <ch> is an evolution of <cl>, <fl>, and <pl>, as shown by Williams (1973). For Cagliari (1999), our alphabet of Latin origin did not offer means for recording the palatals and that is why the digraphs with the letter H were created as a device capable of offering a graphic solution to represent phonemes that did not exist in Latin and that, therefore, had no correlation in the letters of the alphabet. According to the author, "the use of the wildcard H, forming digraphs, changed the principle of acrophonic in an intelligent manner, opening the possibility of new uses for the letters, without changing the alphabet" (CAGLIARI, 1999, p. 372).

In the following section, the results of studies that approach the acquisition of digraphs will be presented.

3 THE ORTHOGRAPHIC ACQUISITION OF DIGRAPHS

Nunes and Bryant (2014) present studies on digraphs and orthography, making comparisons between the spelling of English digraphs, a language that has a high frequency of digraphs and a greater number of studies, and Portuguese digraphs, which are still less studied . The authors seek to understand what happens beyond the initial period of understanding the alphabetic writing system, that is, how the spelling acquisition process takes place, once the alphabetic principle has been consolidated. The term grapheme is used by the authors "[...] when the written sign includes more than one letter" (NUNES; BRYANT, 2014, p. 35), as this would be,

according to them, a way to benefit the understanding/distinction between the digraphs and the other graphemes[4].

Assuming that the correspondence takes place between digraph-grapheme and letter-phoneme, the authors consider that learners at the beginning of literacy need to deal with correspondences based on this type of relationship. For the Portuguese orthographic system, Nunes and Bryant (2014) propose that the digraphs would be divided into mandatory, cases in which there is no other option of graphic representation in the system (<nh>, <lh>, <qu>, <gu>, <rr>), and alternatives, those that have other graphic options to represent the same phoneme (<ch>, <ss>, <sc>, <sç>, <xc>, <xs>). The authors take into account the contextual relationships observed in the first group and the arbitrary ones prevalent in the second. Based on the study by Treiman (1993), for the spelling of English digraphs, Nunes and Bryant (*op. cit.*, p. 36) defend the assumption that "[...] children become aware of the obligatory digraphs as orthographic units more easily than other digraphs, and also use them more consistently".

To address double-letter digraphs, the authors resume the study by Ehri and Soffer (1999) who performed tests with some English digraphs to assess learners' awareness of these units before and after explicit instruction. Regarding the results, which showed that, even before explicit instruction, children already demonstrated to treat double letters as graphophonemic units, Nunes and Bryant (*op. cit*) discuss the functioning of digraphs, making a distinction between those who use the same letter twice to which they call *geminates*[5], (<rr>, <ss>), and those that are composed of different letters (<ch>, <qu>).

The third type of digraph stipulated by the authors concerns extra digraphs or alternative spellings. They would be surplus digraphs, those that could be represented by a single letter, but which for etymological reasons, for example, are still recurrent in the language. The <ph> in English and the <ch> in Portuguese would be examples of this case. Nunes and Bryant's (2014) classification for digraphs, therefore, establishes three types of consonant digraphs, especially targeting the English orthographic system: obligatory vs. alternative; double consonants vs. different letters; extra digraphs.

Even though beginning learners generally operate with principles of simple correspondence between phonemes and graphemes, the studies presented by Nunes and Bryant (2014) indicate that less experienced children and young people perceive the need to represent some sounds with more than one letter early on and without explicit instruction. The data they analyzed also point to the possibility of greater success in spellings of mandatory digraphs due to the learning of orthographic rules compared to the learning of specific words necessary for the optional (extra) digraphs. In addition, low-proficiency readers seem to have more difficulty with required digraphs than younger readers of the same reading age.

Another study that addresses digraphs referring to BP is that of Ferreiro (2013), which takes up data from the interlinguistic project (Spanish, Italian and Portuguese) *Chapeuzinho vermelho aprende a escrever [*Red Riding Hood learns to write] (FERREIRO et al., 1996) and analyzes some trends in digraphs with H in the spellings of children speaking the three languages. Note that the <ch> digraph is the only one shared by all of them. In the Italian system, <ch> represents the phoneme /k/ before /i/, as in *occhi* (eyes) and *chiama* (flame); in Spanish spelling, <ch> corresponds to the spelling of /tʃ/ in words like *escuchar* (to listen) and *cuchillo* (knife); in Portuguese, on the other hand, the digraph <ch> corresponds to the palatal fricative /ʃ/ of encher (to fill) and c**h**ácara (country house), for example. The author, analyzing the sequencing of the letters that make up the digraph, observes the almost non-existence of deviant spellings: although frequent in Italian, there are no cases of permutation, that is, spellings such as 'hc' were not found in the data; in Spanish, children also have a high rate of correct answers, with only one deviant data being found, <sh> for <ch>. In

[4] This position in relation to the concept of grapheme differs from that adopted in this study, which follows the definition proposed by Scliar-Cabral (2003), according to which the grapheme corresponds to one or two letters that represent a phoneme. Thus, by this definition, graphemes correspond to phonemes and, therefore, digraphs relate to a grapheme, as do the letters of the alphabet that correspond to phonemes in the language. Other authors such as Treiman and Kessler (2014) prefer to use the term phonogram instead of grapheme, in view of this apparent inconsistency in the literature involving the definition of grapheme. The concept of phonogram as used by Treiman and Kessler (2014) corresponds to the concept of grapheme as assumed in this text, that is, it concerns a letter or group of letters (such as digraphs) that represent a phoneme of the language.

[5] It should be noted that such nomenclature has implications for phonology since gemination refers to the existence of two tenses for a segment, that is, two temporal units for a single root node, as occurs in Italian at the phonological level and also in the writing system, in word pairs such as ca**pp**ello - ca**p**ello (*hat* and *hair*) pa**ll**a and pa**l**a (*ball* and *shovel*). For BP, however, the presence of geminates in the language is controversial in the case of rhotic and not considered in the case of fricative.

Portuguese, considering the ordering of the letters, one child individually produced the sequence 'hc' in several words of their text.

The dorsal digraphs, <qu> and <gu>, were also analyzed by Ferreiro (2013). The digraph <qu> is also present in Spanish and, like <ch>, has high accuracy rates. According to the study, Brazilian children, in 1,456 occurrences, produced 99% of correct answers, and Spanish children, in 3,343 contexts for analysis, 98%. The errors found do not involve sequencing, but the use of <c>, <q>, or <k> in place of <qu>. The author also analyzes two possibilities of occurrence of <qu>, without <e> or <i>, representing the sequences /ke/ and /ki/. In data such as '*porqu*' for 'porque‟ (because) there would be two possible interpretations, according to the author: one can consider the omission of the vowel <e> or a change in the sound value of <u>, which would represent the sequence /ke/.

The GEALE[6] studies contemplated discussions regarding the digraphs of palatal sonorants and also of rhotic and alveolar fricative, based on data extracted from the Database (MIRANDA, 2001). Teixeira and Miranda (2008) specifically analyzed the spellings of palatal sonorants that in Portuguese only have their orthographic representation related to the digraphs <nh> and <lh>[7].

Errors are treated based on two main categories: *errors related to knowledge gaps regarding the use of digraphs and errors that evidence phonological processes*. In the first, data was included in which there was the deletion of 'l' or 'n' that make up the digraph or those in which the child changed 'nh' to 'lh' and vice versa; in the second, data was classified in which some similarity was observed with the processes verified in the phonology acquisition. The following examples illustrate the data analyzed by the authors:

Spelling of /ɲ/

a) mihoca – misspelling of minhoca (worm) (2nd grade)
b) milha - misspelling of minha (mine - possessive pron.) (2nd grade)
c) passarino - misspelling of passarinho (bird) (2nd grade)
d) vinheram - misspelling of vieram (they came) (3rd grade)

Spelling of /ʎ/

e) espantaho - misspelling of espantalho (straw man) (3rd grade)
f) finho - misspelling of fino (thin) (1st grade)
g) olios - misspelling of olhos (eyes) (1st grade)
h) filo - misspelling of filho (son) (3rd grade)

The examples of the first two lines show spellings interpreted by Teixeira and Miranda (2008) as related to orthography, because only the H was recorded or because there was a change in the digraph selection: where it should be <nh>, the child wrote <lh> and vice versa. In the last two lines, there are cases considered as phonological by the authors, since the spellings are related to processes observed in the phonological acquisition, which reinforce the idea that such consonants have a complex structure, that is, a primary consonant articulation and another vowel (according to CLEMENTS; HUME, 1995). Based on the idea of a complex consonant, it is possible to explain why /ʎ/ can be produced by children as [l], [lj], [i], [j], or even by the liquid [ɾ]; and /ɲ/ like [n] and [i], in forms that contemplate the two parts of the consonant separately or only one of them. The phonetic zero is also observed in the acquisition of speech and writing, as in “mioca” for “minhoca‟ (worm), an example that serves to explain why the inverse also occurs, as shown in the data in (d), “vinheram” for “vieram” (they came).

In addition to the study by Teixeira and Miranda (2008), the spellings of palatal sonorants in relation to the phonological acquisition were the subject of studies by Miranda (2012, 2014, 2020). In all computations, the error rate compared to the success rate is low.

[6] Study Group on Written Language Acquisition. Research group founded in 2001, registered with CNPq: <http://dgp.cnpq.br/dgp/espelhogrupo/3189572089608769.

[7] In Italian, <gn> and <gl> are used for nasal and liquid, respectively, and in Spanish, <ll> for liquid and <ñ> for palatal nasal.

According to Teixeira and Miranda (2008), of the 7,291 words with palatal words extracted from texts produced by children in the early years, only 416 had errors, equivalent to 6%. The highest incidence is related to the spelling of palatal liquids. Regarding the computation of dictation data, a specific instrument for the spelling of palatals, the errors reached an index of 13%, 413 occurrences in a universe of 3,202 contexts, with the first two grades responsible for the vast majority of errors and the liquid being more involved in errors when compared to the nasal.

Studies by GEALE also analyzed the spellings of the 'strong-r' and the coronal fricative /s/ and showed that errors involving the digraphs <rr> and <ss> are responsible for a high incidence of errors in the sample studied. The analysis by Araújo, Garcia, and Miranda (2006) on the spelling of 'strong-r' was carried out from the analysis of 3,160 words with context for 'strong-r' extracted from spontaneous texts. The incidence of errors in intervocalic contexts is on average 30% in the first two years of public school and 18% in the same years of private school.

Regarding the analysis of the use of digraphs to record /s/, Araújo, Garcia, and Miranda (2006) found 7,448 words in public school texts and 9,203 in private school texts with the context for recording this phoneme. The general error rates were around 25% in public schools and 18% in private schools. Concerning only errors in the use of the digraph <ss>, 35% of the data are of this spelling in both schools.

4 METHODOLOGICAL PROCEDURES

The data analyzed in this article were taken from texts that make up the Written Language Acquisition Text Database - BATALE (MIRANDA, 2001). BATALE is composed of spontaneous texts by Brazilian, Portuguese, and Mozambican children, collected from textual production workshops developed at GEALE. The workshops follow a pattern that consists of three different and interconnected moments: warm-up/motivation, individual writing, and sharing with the group. The textual productions, after being collected, are typed in Word, scanned in PDF, and encoded with information about the school, the school year, and the student. After that, they are stored in catalog folders in the research group room. Spontaneous writing is considered here in terms of Abaurre (2011), as records that demonstrate unique moments of interaction between the learner and the object of knowledge, which can help in the perception of the writing strategies used by each child.

To investigate the spelling of the digraphs, the sample considered was composed of the first 5 collections of Stratum 1 of the Database, carried out between 2001 and 2003[8]. At the time, the children attended the 1st to 4th grades of 8-year Elementary School (corresponding age group from 6 to 12 years), and their productions resulted in 979 texts – 479 from public schools and 500 from private schools. The digraphs considered were the consonant digraphs: <ch>, <lh, <nh>, <rr>, <ss>, <sc>, <gu>, and <qu>, being analyzed all the words that contained such graphemes in the target word. Data were organized into categorical variables: correct or incorrect answers; if errors, categorized according to Miranda's proposal (2020), in phonological, orthographic, or phonographic; digraph type (H, U, and heterosyllabic digraphs); school grade (1st to 4th), and type of school (public or private). Data analysis is quantitative-qualitative and descriptive. In this regard, data frequencies and proportions were collected according to the aforementioned variables, using the *RStudio software*, but no inferential statistical analysis was performed, so the results obtained and conclusions drawn are limited to this sample. The qualitative aspect of the research concerns especially the discussion about the nature of errors.

[8] BATALE's Stratum 1 has subsidized most of the GEALE research done to date. It is the oldest, most voluminous, and with the most complete treatment, with a total of 10 data collections distributed over 4 years, with 2024 texts collected and 24000 (ortho)graphic errors tabulated. For these reasons and because there is still no research focused exclusively on digraphs conducted in the research group, this was the stratum selected for analysis in this article, even though it has data produced by children who were still attending 8-year Elementary School at the time. It should be noted that, due to the theoretical-methodological orientation adopted by GEALE, whose emphasis is on the way in which the internalized knowledge already built on the mother tongue is taken up by the child in their initial spellings, the difference in the curricular organization and in the period in which collections were performed does not necessarily influence the results. This is what recent studies carried out by the group has allowed to corroborate, by showing a very similar distribution of errors analyzed from the linguistic approach adopted and also from the categories identified with the nature of the errors (according to PACHALSKI; MIRANDA, 2019; MIRANDA, 2020)

5 RESULTS AND DISCUSSION

In this section, the results of the computation performed will be presented, referring to the distribution of errors and correct spellings for each of the investigated digraphs, in the form of descriptive statistics. From the 979 texts, 8286 contexts were extracted for the spelling of consonant digraphs, with an average of 8.5 digraphs per text. Considering the type of school variable, the following distribution between errors and correct spellings was obtained:

school	**error**	**correct spelling**	**total occurrences**
public	343	3,041	3,384
	10.10%	*89.90%*	*100%*
private	306	4,596	4,902
	6.20%	*93.80%*	*100%*

Table 1: Ratio of frequency and proportion of errors and correct spellings **by school** (public and private)
Source: own preparation (research data)

The general computation of the data of errors in the spelling of the digraphs, by type of school, shows that the number of errors is higher in public schools. Miranda (2017, 2020), when analyzing the total data from the same stratum of BATALE, in a universe of 24000 errors surveyed, found similar results, that is, a higher incidence of errors in the sample consisting of data from public schools, in a 2/3 proportion for public and 1/3 for private. The author points out, however, that, despite the difference in the number of errors, the quality of errors produced is very similar between schools.

Table 2, below, presents the results of the distribution of errors, now considering the school grade in each of the schools studied:

grade	**public school**		**private school**	
	errors/contexts	*percentage of errors*	*errors/contexts*	*percentage of errors*
1 st	36/276	13.0%	93/642	14.5%
2 nd	127/1238	10.2%	117/1440	8.1%
3 rd	94/843	11.1%	60/1374	4.3%
4 th	86/1027	8.3%	37/1447	2.5%

Table 2: Frequency ratio and proportion of errors **by grade** for public and private schools
Source: own preparation (research data)

The distribution results show a gradual decrease in the number of errors found in the private school data and a different profile in the public school data. Both schools show a decrease in errors, from the first to the second grade. The public school reveals, however, an increase in the error rate in the third grade compared to the second grade and resumes the downward trend in the fourth grade.

There is also a marked difference in the rates of fourth grades between schools, keeping the rate relatively high in public school. Thus, it is possible that a closer look at the error, considering the type of digraph involved, may support interpretive assumptions for the distribution of the data in Table 2.

In order to compare the distribution of errors by type and by school, the digraphs were grouped into three blocks, following the characteristics mentioned above, in terms of their functioning, namely: digraphs with H for the palatal ones, heterosyllabic digraphs for the rhotic and coronal fricative, and digraphs with U for the dorsal ones. Charts 1 and 2, below, show the distribution of errors and correct spellings for each of the digraphs analyzed in both schools:

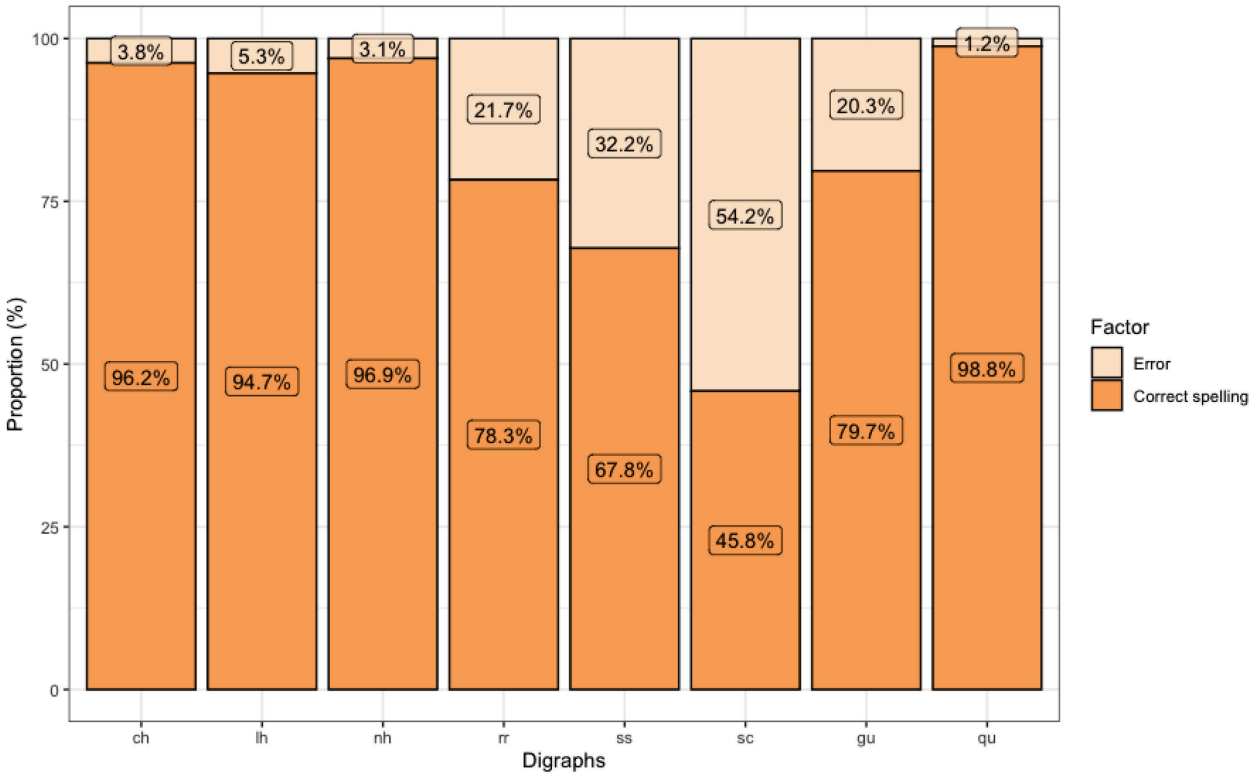

Chart 1: errors and correct spellings per digraph in **public schools**

Source: own preparation (research data)

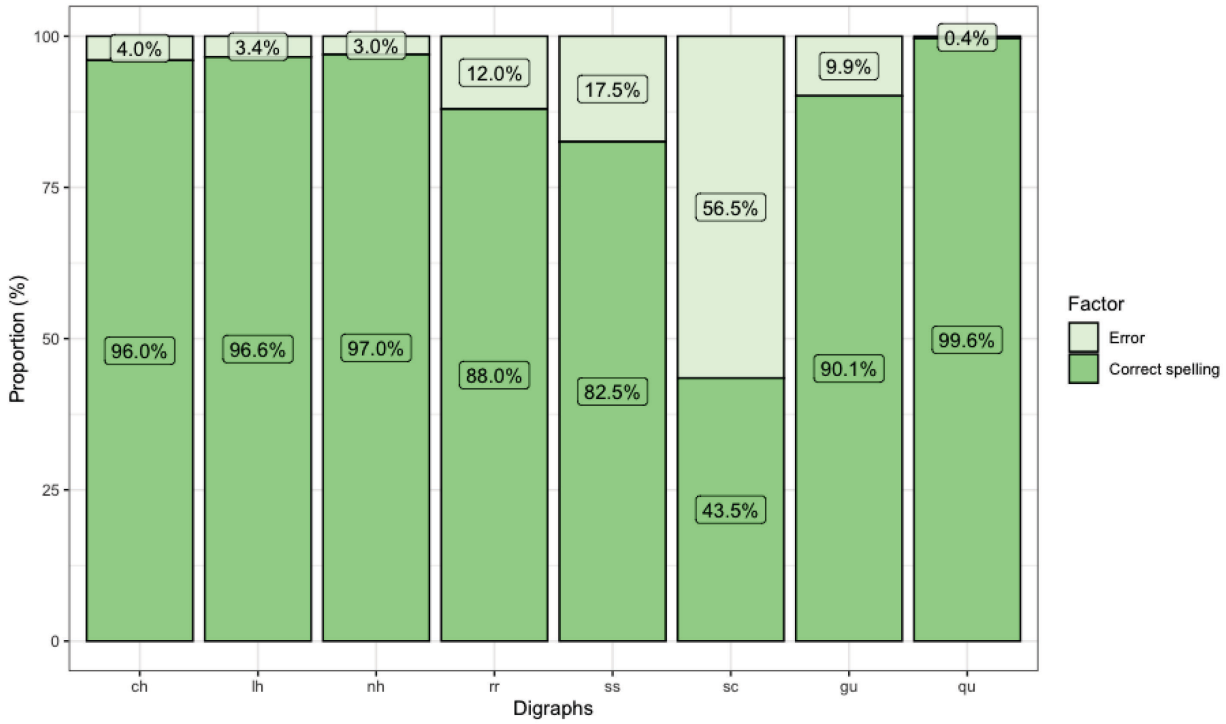

Chart 2: errors and correct spellings by digraph in **private schools**
Source: own preparation (research data)

The first aspect to be highlighted concerns the similarity between the results expressed in the two graphs presented. Low rates of errors for the palatal ones are observed, that is, for the digraphs with H; higher incidence of errors in the spellings of the digraphs for fricative and rhotic, with a predominance of errors in the fricative record, especially <sc>; and, among the dorsal consonants, more errors in the use of <gu> compared to <qu>.

Secondly, Charts 1 and 2 show a greater number of correct answers than errors in all consonant digraph records, except for <sc> whose percentage of correct answers is smaller than that of errors, computing 54.2% and 56.5%, in the public and private schools, respectively. It is important to take into account that the frequency of this digraph in the orthography of BP, compared to the others, is low. In the analyzed sample, only fifty contexts were found for <sc>, which correspond to five lexical items, namely: *nascer* (to be born), *crescer* (grow), *piscina* (swimming pool), *descer* (to descend) and *ressuscitar* (to resurrect), this last item with only two occurrences.

On the other hand, the heterosyllabic digraphs, <ss> and <rr>, seem to have different behavior when comparing schools. Although both present more errors in relation to the other digraphs, the public schools exhibit higher rates. The first round carried out with the amalgamated data, from the public and private schools, revealed an average of 81.1% of correct answers in <rr> and 76.4% in <ss>. When data is sorted by type of school, there is a difference in the percentages with higher incidences of errors in public schools, even though the order <ss> > <rr> is the same for both. The digraphs with H exhibit similar results to each other: <ch>, <lh> and <nh> present, respectively, 96.1%, 95.5%, and 96.8% of correct answers in this sample, when not considering the variable type of school which, when taken into account, reveals a small difference regarding the <lh> with a higher rate of errors in public schools. The dorsal digraphs also seem to be different in terms of the number of correct answers: the one with the most correct answers in this group and in the entire sample is <qu>, with 99.4% of the occurrences reaching the stipulated spelling of the target words, while <gu> presents 86% of correct answers, a percentage that changes again when the type of school is taken into account, as shown in Charts 1 and 2.

Table 3 shows, in a complementary way to Charts 1 and 2, the relationship of frequency and proportion of errors for the digraphs analyzed in this sample, in each of the grades, offering more details about the frequency of occurrence of the data as well as examples of errors found:

dígrafos	**escola pública**			**escola particular**			**exemplos**
	série	*erros/contextos*		*série*	*erros/contextos*		
<lh>	1ª 2ª 3ª 4ª	1/38 16/285 5/119 9/139	2% 5% 4% 6%	1ª 2ª 3ª 4ª	4/58 9/188 6/168 0/142	6% 4% 3% —	'espanta**li**o' para 'espanta**lho**'
<nh>	1ª 2ª 3ª 4ª	3/61 4/291 9/204 9/260	5% 1% 4% 3%	1ª 2ª 3ª 4ª	8/198 16/285 6/275 2/297	4% 5% 2% 0.6%	'mi**h**oca' para 'mi**nh**oca'
<ch>	1ª 2ª 3ª 4ª	13/122 6/271 8/179 2/201	**10%** 2% 4% 1%	1ª 2ª 3ª 4ª	13/86 11/307 8/251 3/232	**15%** 3% 3% 1%	'a**x**ou' para 'a**ch**ou'
<ss>	1ª 2ª 3ª 4ª	12/16 71/195 46/140 44/186	**75%** 36% 32% **23%**	1ª 2ª 3ª 4ª	46/69 50/259 30/270 22/250	**66%** **19%** 11% 8%	'di**c**e' para 'di**ss**e' 'va**s**ora' para 'va**ss**oura'
<sc>	1ª 2ª 3ª 4ª	1/1 5/6 3/8 5/9	100% 83% 37% 55%	1ª 2ª 3ª 4ª	8/9 3/4 0/3 2/7	88% 75% — 28%	'na**c**eu' para 'na**sc**eu'
<rr>	1ª 2ª 3ª 4ª	5/17 23/88 17/67 11/86	29% 26% 25% **12%**	1ª 2ª 3ª 4ª	5/25 21/106 8/91 6/110	20% 19% **8%** 5%	'soco**r**o' para 'soco**rr**o'
<qu>	1ª 2ª 3ª 4ª	1/18 0/89 1/107 2/122	5% — 0.9% 1%	1ª 2ª 3ª 4ª	2/183 2/254 0/271 0/356	1% 0.7% — —	'bos**kqu**e' para 'bos**qu**e'
<gu>	1ª 2ª 3ª 4ª	1/3 2/13 5/19 4/24	33% 15% 26% 16%	1ª 2ª 3ª 4ª	7/14 5/37 1/48 2/53	50% 13% 2% 3%	'**g**erra' para '**gu**erra'

Table 3: Distribution of errors by digraph and grade in public and private schools
Source: own preparation (research data)

A general assessment of the distribution by grade and school corroborates the low incidence of errors in all grades and in both schools regarding the spelling of digraphs with H, which has already been observed in the general distribution (Charts 1 and 2). The new information that arrives concerns a more specific difference observed between the palatals. While the sonorants present two-

way relationships between phoneme and grapheme, with <nh> and <lh> being the only options, the fricative is related to two graphemes that arbitrarily compete, <ch> and <x>, which gives it a greater degree of orthographic complexity and may explain the higher percentage incidence of errors in the first grade of both schools, 10%, and 15%.

The group of digraphs related to alveolar fricative and rhotic, specifically <ss> and <rr>, show a gap between the data of the two schools as the grades advance. In the case of <ss> there is a very high rate of errors in the first grade, 75% in the public school and 66% in the private school; such result, however, is expected for initial alphabetic spelling, since the orthography learning begins to be established after the consolidation of the alphabetic writing and the entry of information about orthographic rules via school practices. As it is the most complex spelling of the system, since /s/ has ten possible forms of representation, among which <ss> and <sc>, it is expected to observe the impact of school practices on writing, in the sense of gradually conforming to the norm. In this regard, there is a considerable drop in the number of errors in the private school and a slight decrease in the public school, which, in the fourth grade, presents an index closer to that observed in the second grade of the private school, 23% and 19%, respectively. As for the use of <rr>, which is worth noting that it's conditioned by a contextual rule, the data shows a decrease in errors throughout the grades, but with a slower decrease in public schools and more accentuated in private ones. The error rate of the public fourth grade is higher than that of the private third grade, 12%, and 8%, respectively.

The third group of digraphs, those with U, shows low rates of errors in both schools and in the four grades analyzed. It is worth noting that the public school has higher percentages in relation to <gu> errors, a result that converges to those already indicated in relation to the <rr> digraph, also regulated by contextual rule, and which shows slower acquisition in public schools.

After appreciating the results regarding the distribution of errors and correct answers in each digraph, considering school and grade, we now propose a look at the quality of the errors produced. To do so, we resort to the proposal of the Research Group for the nature of spelling errors. Thus, some specific considerations about the proposal will be briefly laid out, in order to explain the logic used in the data analysis proposed in this article.

Three major sources of knowledge mobilized in the heteroclite spellings observed in the early years are identified by the GEALE studies, namely: phonology, orthography, and phonography. Each of these sources corresponds to the complexity involved in the initial alphabetic-orthographic spellings and, when identified, can guide the analyst's or teacher's interpretation. The summary of the proposal is presented in Table 4:

Phonological	**Orthographic**	**Phonographic**
There is no orthographic complication, but phonological	There is no phonological complication, but rather orthographic	There is no orthographic or phonological complication
Segmental and prosodic	Contextual and arbitrary	Tracing, sequencing, omitting and inserting
Related to some kind of phonological complexity or even motivated by speech	Related to non-observance of contextual or arbitrary rules	Related to phoneme-grapheme processing

Table 4: GEALE error categories
Source: Miranda (2020, p.15)

Considering that phonology is at the base of the alphabetic system and that children use the phonological knowledge already acquired to produce their first spellings, it was postulated that phonological complexity, both at the segmental and prosodic levels, may be responsible for part of the errors observed, specifically when there is no orthographic complexity, only phonological. The second complicating factor taken into account for the analysis of errors is based on the orthographic system, that is, cases of non-direct relationships between phonemes and graphemes, the so-called multiple relationships (LEMLE, 1987). In addition to these two aspects that motivate errors, there is a third that occasionally operates on the initial spellings, namely, the phonographic. Phonography is understood here as the set of knowledge related to the alphabet (the letters, the names of the letters, the sounds, and their graphic correlates) and the automation of access to this information, as well as the mechanical execution of the writing of words.

Table 5 below shows examples of errors extracted from the sample studied with the respective interpretation according to their nature[9]:

	phonological	**orthographic**	**phonographic**
<ch>	**g**ifre for **ch**ifre (horn)	bi**x**os for bi**ch**os (animals)	**ce**h**g**o for **ch**egou (arrived)
<lh>	fi**li**a for fi**lh**a" (daughter)	*familha for família (family)	miho for mi**lh**o (corn)
<nh>	minoca for minhoca (worm)	–	tenlho for tenho (I have)
<rr>	*vago for varro (I sweep)	ca**r**inho for ca**rr**inho (small car)	*ca**rrr**inho for ca**rr**inho (small car)
<ss>	azim for assim (thus)	va**s**oura for vassoura (broom)	parinhos for pa**ss**arinhos (birds)
<sc>	*pe**z**ina for pi**sc**ina (swimming pool)	pe**ss**ina for pi**sc**ina (swimming pool)	*pe**cs**ina for pi**sc**ina (swimming pool)
<gu>	ni**qu**ei for nin**gu**ém (nobody)	ni**g**ei for nin**gu**ém (nobody)	comce**gr**iu for conse**gu**iu (was able to)
<qu>	**gu**e for **qu**e (that)	bos**c**e for bos**qu**e (woodland)	bos**kqu**e for bos**qu**e (woodlan por**q** for por**qu**e (because)

Table 5: examples of errors in the spelling of digraphs according to their respective natures

Source: own preparation (research data)

The analysis focused on the quality of errors, based on the categories proposed by GEALE, allows a vision of the strategies used by children when spelling the digraphs and gives clues about the assumptions they formulate about these graphophonic units.

Charts 3 and 4 show how the errors are distributed according to their nature for each of the analyzed digraphs:

[9] Examples with an asterisk (*) do not belong to the sample studied. They are, in fact, potential data that would be classified according to the proposal regarding the nature of errors, as adopted in this article. Note that palatal sonorants have a one-to-one relationship between phonemes and graphemes, with the exception of a few words with palatal liquid, as mentioned above. Thus, we chose to consider that <lh> may have spelling errors precisely because of the alternation <lh> vs. in words where /ʎ/ is at the base, which is not the case with <nh>, which is why the cell referring to <nh> is empty.

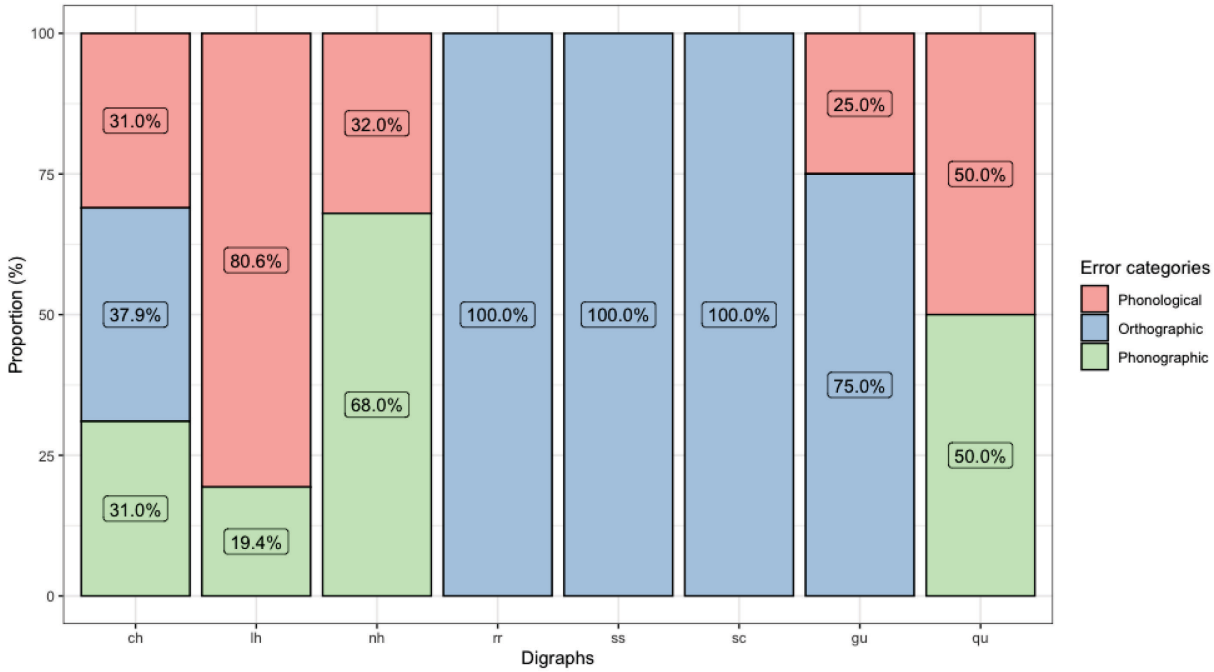

Chart 3: error categories per digraph in **public schools**
Source: own preparation (research data)

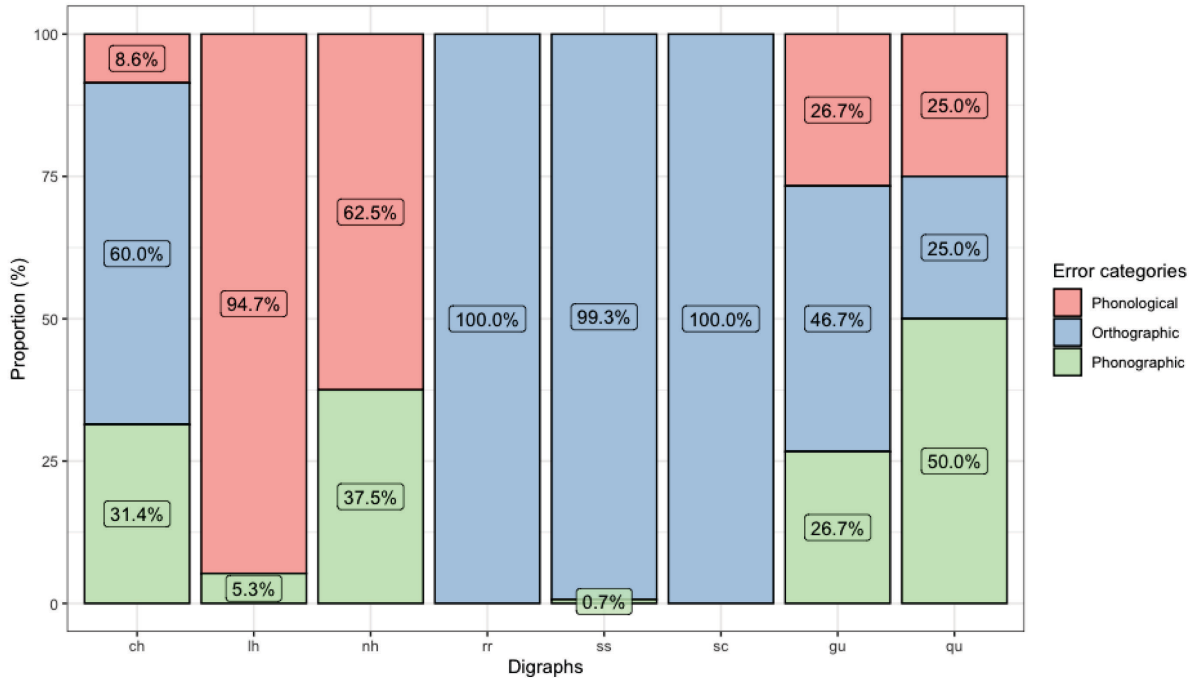

Chart 4: categories of error by digraph in the **private school**
Source: own preparation (research data)

When looking at Charts 3 and 4, it can be seen that the phonological errors are concentrated in the spellings of the digraphs with H and in the digraphs with U, while the orthographic errors in the spellings of the digraphs of rhotics and fricatives (/s/ and /ʃ/. The phonographic errors were found in the digraphs with H and with U.

In the spelling of digraphs with H, phonological errors are predominant in the records of the palatal liquid, 80.6% and 94.7% of errors in public and private schools, respectively. It is a complex consonant that, according to Bisol (1994) and Matzenauer (1994), has two articulations, a primary consonant, and a secondary vowel. The same complexity is proposed for the palatal nasal which, in

the data studied, reveals phonological and also phonographic error rates, the latter being more present in the public school data[10]. Note that palatal sonorants require the child to access the internal structure of a segment consisting of two nodes (according to CLEMENTS; HUME, 1995) and also the phonographic domain in the sense of selecting the letters that compose the digraph as well as its ordering. The expectation is that errors of this type occur in the earliest period of orthographic-alphabetic spelling. As for the <ch>, one must consider the effect of orthographic rules given the observed phonographemic relationship, which provides <x> and <ch> as options for the learner, which is reflected in the majority tendency of spelling errors.

The phonographic and phonological errors are also present in the data referring to the digraphs with U. In the case of <qu>, it is considered phonographic, for example, the motivation for the use of 'q', only, taking into account that there is a correspondence between the use of the letter and its name. Phonology is manifested in exchanges involving the feature [sonorous], in this case, the use of <g> to register /k/. The data "bosce" for "bos**qu**e" (woodland) was interpreted as orthographic based on the alternation existing in the system between <qu> and <c> for the registration of /k/, depending on the subsequent vowel

Finally, the digraphs <ss>, <sc> and <rr>, whose errors are almost entirely related to non-observance of orthographic rules, point to the effects of teaching orthography, or the lack of it, which can be best understood from the joint reading of Charts 3 and 4 and Table 3. As for rhotics, as it is a contextual rule capable of determining the representation of the 'strong-r' in all contexts, in this specific case, in intervocalic contexts, it is observed that in the public-school data there are still many errors in the fourth grade, which may be motivated by the difference in access to more effective reading and writing practices. Such results suggest the need to intensify classroom activities capable of drawing attention to the rules governed by context.

This statement is also supported by the results related to the use of <ss> as mentioned earlier. It is interesting to note that a more detailed analysis of the type of error for <ss> shows that in both schools there is a predominance of contextual errors, that is, the use of the grapheme <s> in an intervocalic position (as in "va**s**oura" to "va**ss**oura" (broom)), which corresponds to 75% of errors in public schools and 81% in private schools. These results may lead to a reflection on the role of explicit instruction in early-grade classrooms. Learning to see the functioning of the linguistic system, in this specific case of the orthographic system; giving attention to form; identifying the phonological neighborhoods, the positions of the segments, and the possible sequences and also the preferred ones; seem to be necessary tasks for both educational contexts, public and private, which may result in a considerable reduction of spelling errors and a consolidation of children's spelling lexicon.

6 FINAL CONSIDERATIONS

This article sought to describe and analyze the spelling of consonant digraphs in the Brazilian Portuguese orthographic system by children in the early years of Elementary School. To this end, the (ortho)graphic error, according to Miranda (2020), was taken as the main object of analysis, being discussed considering a qualitative and quantitative-descriptive treatment, and according to the variables type of school (public and private), grade (1st to 4th), type of digraph (digraphs with H, with U and heterosyllabic) and type of error (phonological, orthographic and phonographic). The contribution of this study lies, above all, in the approach to a topic that has been little studied in Brazil, especially from the point of view of spelling acquisition. In this regard, it is expected that it can give rise to future investigations, in order to expand the state of knowledge regarding the functioning and acquisition of digraphs in the orthographic system of BP, and with that, also, subsidize the educational practices in the Literacy Cycle.

More specifically, the results pointed, first, to a trend already observed in other GEALE studies, that is, that errors are more numerous in public school when compared to private schools, and that, even with the advance of school grades, the (expected) decrease in errors is always smaller in public schools. It should be noted, however, that when analyzing the quality of errors, the trends are very similar, which is attributed to the internalized linguistic knowledge common to the children of both groups, and also to the knowledge already built on the writing system, even though the latter, sometimes dependent on explicit instruction, develop

[10] In this study, the position adopted differs from that of Teixeira and Miranda (2008), since for the authors the use of H or another digraph with H to represent the sonorants was an orthographic error. In the analysis presented here, it is considered that such errors have a phonographic nature, since the digraphs <lh> and <nh> correspond to complex segments and, from the orthographic point of view, they present a two-way relationship between phoneme and grapheme.

at uneven paces between students from both types of schools, probably due to the difference in access to reading and writing practices in children's daily lives.

It was observed that children from both schools treat digraphs according to the nature of the complexity that is involved in each case of phoneme-grapheme relationship. In this regard, the groupings made – digraphs with H, heterosyllabic digraphs, and digraphs with U – proved to be relevant. The digraphs that seem to be the greatest object of doubt among children in the early years are the heterosyllabic ones, considering the general error rates in relation to correct spellings, both in public and private schools. The mark of this group is the orthographic complexity since <ss>, <sc>, and <rr> are regulated by arbitrary and contextual relationships in the system, a fact that is reflected in the nature of error recorded in almost 100% of the occurrences being orthographic, both in public and private schools, again. In digraphs with H and U, whose error rates are quite low compared to heterosyllabic ones, there is a greater influence of aspects of phonological and phonographic order. The <gu> digraph, specifically, has slightly higher error rates, which also seems to be linked to the orthographic complexity, since <gu> competes, in contextual relation, with <g> for the representation of /g/. No wonder (ortho)graphic errors are the majority in the record of this grapheme.

In view of these results, it can be assumed that aspects of the order of learning the rules that govern the multiple relationships between phonemes and graphemes, as well as the construction and consolidation of an orthographic lexicon, are the most influential in the spelling of these peculiar cases of the orthographic system when a phoneme is represented by a grapheme composed of more than one letter. Thus, the indispensability of explicit instruction, in the classroom, about the functioning of the orthographic system, with a proposal that includes, but goes beyond memorization, covering the identification of phonological contexts and the corresponding (ortho)graphic patterns is perceptible. In this regard, it is emphasized that the categorization of errors based on a proposal that considers their nature and the child's knowledge during spelling acquisition, can also provide the teacher with a valuable tool for mapping the hypotheses with which their students are operating, thus, indicating the specific aspects on which the educational action should focus at a given moment in the child's development.

REFERENCES

ABAURRE, M. B. M. A relação entre escrita espontânea e representações linguísticas subjacentes. *Verba Volant*, v. 2, n. 1, 2011. Disponível em: http://letras.ufpel.edu.br/verbavolant/segundo/abaurre2.pdf. Acesso em: 20 dez. 2021.

ARAÚJO, P.; GARCIA, M. A.; MIRANDA, A. R. M. O que revelam os textos espontâneos e um ditado com palavras inventadas sobre a grafia das consoantes róticas. *In*: CELSUL - CÍRCULO DE ESTUDOS LINGUÍSTICOS DO SUL, 7., 2006, Pelotas. *Anais* [...]. Pelotas, 2006. p. 1-10. Disponível em: http://www.leffa.pro.br/tela4/Textos/Textos/Anais/CELSUL_VII/dir1/12.pdf. Acesso em: 20 dez. 2021.

BISOL, L. Ditongos derivados. *D.E.L.T.A.*, v. 10, n. Especial, p.123-140, 1994. Disponível em: https://revistas.pucsp.br/delta/article/viewFile/45380/29973. Acesso em: 20 dez. 2021.

BRASIL. Decreto nº 6583, de 29 de setembro de 2008. Promulga o Acordo Ortográfico da Língua Portuguesa. *Diário Oficial da Uni*ão: Brasília, 30 set. 2008. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/decreto/d6583.htm. Acesso em: 22 dez. 2021.

CAGLIARI, L. C. *Alfabetizando sem o ba-be-bi-bo-bu***.** São Paulo: Scipione, 1999.

CLEMENTS, G. N.; HUME, E. V. The internal organization of speech sounds. *In*: GOLDSMITH, J. (org.). *The Handbook of Phonological Theory*. London: Blackwell, 1995.

EHRI, L. C.; SOFFER, A. G. Graphophonemic Awareness: Development in Elementary Students. *Scientific Studies of Reading*, v. 3, n. 1, p. 1-30. Disponível em: https://psycnet.apa.org/record/1999-00820-001. Acesso em: 22 dez. 2021.

FARACO, C. A. *Escrita e alfabetização*. 9. ed. São Paulo: Contexto, 2015.

FERREIRO. E. O *ingresso na escrita e nas culturas do escrito*: seleção de textos de pesquisa. São Paulo: Ed. Cortez, 2013.

FERREIRO, E. *et al. Chapeuzinho vermelho aprende a escrever:* estudos psicolinguísticos comparativos em três línguas. Juiz de Fora: Ática, 1996.

GARCIA, M. A. C.; ARAÚJO, P. R. M.; MIRANDA, A. R. M. Um estudo sobre a grafia do fonema /s/. *In*: CELSUL - CÍRCULO DE ESTUDOS LINGUÍSTICOS DO SUL, 7., 2006, Pelotas. *Anais* [...]. Pelotas, 2006. p. 299-299. Disponível em: http://www.leffa.pro.br/tela4/Textos/Textos/Anais/CELSUL_VII/dir1/18.pdf. Acesso em: 22 dez. 2021.

LEMLE, M. *Guia teórico do alfabetizador*. 17. ed. São Paulo: Ática, 2009.

LUFT, C. P. *Novo manual de português:* gramática, ortografia oficial, literatura brasileira e portuguesa, redação e testes de vestibular. 17. ed. São Paulo: Globo, 1991.

MATZENAUER, C. L. A Geometria de Traços na representação das palatais na aquisição do Português. *Letras de Hoje*, v. 29, n. 4, p. 159-167, 1994. Disponível em: https://revistaseletronicas.pucrs.br/ojs/index.php/fale/issue/view/745. Acesso em: 20 dez. 2021.

MIRANDA, A. R. M. *BATALE:* Banco de Textos de Aquisição da Linguagem Escrita. Faculdade de Educação. Universidade Federal de Pelotas. 2001. Disponível em: http://sistemavestigios.org/. Acesso em: 20 dez. 2021.

MIRANDA, A. R. M. Reflexões sobre a fonologia e a aquisição da linguagem oral e escrita. *Veredas Online*, v. 16, p. 122-140, 2012. Disponível em: https://periodicos.ufjf.br/index.php/veredas/article/view/25060. Acesso em: 20 dez. 2021.

MIRANDA, A. R. M. A fonologia em dados de escrita inicial de crianças brasileiras. Madri, *Linguistica*, v. 30, p. 45-80, 2014. Disponível em: http://www.scielo.edu.uy/scielo.php?pid=S2079-312X2014000200004&script=sci_abstract. Acesso em: 20 dez. 2021.

MIRANDA, A. R. M. Aquisição da escrita – as pesquisas do GEALE. *In*: MIRANDA, A. R. M.; CUNHA, A. P. N.; DONICHT, G. (org.). *Estudos sobre aquisição da linguagem escrita*. Pelotas: Editora da UFPel, 2017. Disponível em: http://guaiaca.ufpel.edu.br/handle/prefix/4391. Acesso em: 20 dez. 2021.

MIRANDA, A. R. M. Um estudo sobre a natureza dos erros (orto)gráficos produzidos por crianças dos anos iniciais. Belo Horizonte, *Educ. rev.*, v. 36, e221615, 2020. Disponível em: https://www.scielo.br/j/edur/a/Z3trLgDyBBZWB466FLHFXtz/abstract/?lang=pt. Acesso em: 20 dez. 2021.

MORAIS, A. G. *Ortografia: ensinar e aprender*. 4. ed. São Paulo: Ática, 2006.

NUNES, T.; BRYANT, P. *Leitura e ortografia*: além dos primeiros passos. Porto Alegre: Penso, 2014.

SEYMOUR, P. H. K.; ARO, M.; ERSKINE, J. M. Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, v. 94, p. 143-174, 2003. Disponível em: https://bpspsychub.onlinelibrary.wiley.com/doi/abs/10.1348/000712603321661859. Acesso em: 19 dez. 2021.

SILVA, R. V de M. *O português arcaico:* fonologia. 3. ed. São Paulo: Contexto, 1996.

TEIXEIRA, S. de M.; MIRANDA, A. R. M. Descrição e análise dos erros ortográficos referentes à grafia das soantes palatais e discussão sobre seu status fonológico. *In*: CELSUL - CÍRCULO DE ESTUDOS LINGUÍSTICOS DO SUL, 8., 2008, Porto Alegre. *Anais* [...]. Pelotas: EDUCAT, 2008. p. 1-9. Disponível em: https://wp.ufpel.edu.br/geale/files/2010/11/shimene_celsul1.pdf. Acesso em: 20 dez. 2021.

WILLIAMS, E. *Do Latim ao Português*. Rio de Janeiro: Tempo Brasileiro/INL, 1973.

Received on February 3, 2022. 03/02/2022. Approved on May 5, 2022.