



## Orthographic effects in speech production: A psycholinguistic study with adult Brazilian-Portuguese English bilinguals

### *Efeitos ortográficos na produção da fala: um estudo psicolinguístico com adultos bilíngues falantes de Português Brasileiro e Inglês*

Alison Roberto Gonçalves

Universidade Federal do Paraná (UFPR), Curitiba, Paraná / Brazil

alisonrobertogoncalves@gmail.com

<https://orcid.org/0000-0003-0959-7053>

Rosane Silveira

Universidade Federal de Santa Catarina (UFSC/CNPq), Florianópolis, Santa Catarina / Brazil

rosanesilveira@hotmail.com

<https://orcid.org/0000-0003-0329-0376>

**Abstract:** The present study inquired whether orthography affects phonological processing of English as an L2. To do so, a lexicon that simulated opaque and transparent grapho-phonetic English relations in nuclear position was developed (e.g., *keet*, *deit*, *toud*). Bilingual speakers of Brazilian Portuguese and English were compelled to learn this new lexicon through a repeated-exposure training paradigm in which they were introduced to the lexicon phonological forms associated with their visual forms, and then to the phonological forms associated with their visual and orthographic forms. After undergoing training, subjects were tested with a Timed Picture Naming task to investigate orthographic recruitment in spoken production. Results suggested that orthography influenced naming of the trained words, indicating that the process of converting a visual input into its phono-articulatory representations for production involves orthographic activation. Such a finding was interpreted as a frequency effect of the grapho-phonetic combination, which resulted in lack of skill to compute this operation in the sublexical route. Overall, the presence of orthographic effects in this task can be interpreted as evidence for such a system to function as a strategic mechanism that

aids lexical encoding and, consequently, influences lexical access in initial stages of instructed language acquisition.

**Keywords:** phonological acquisition; orthography; psycholinguistics.

**Resumo:** Este estudo investigou se a ortografia afeta o processamento fonológico do inglês como L2. Para tal, um léxico que simulava as relações grafo-fônicas opacas e transparentes do inglês em posição nuclear (e.g., *keet*, *deit*, *toud*) foi desenvolvido. Bilíngues falantes de português brasileiro e de inglês participaram de um treinamento para adquirir este novo léxico com o paradigma de exposição repetida, através do qual foram introduzidas as formas fonológicas deste léxico associadas às suas formas visuais e, depois, as formas fonológicas associadas às suas formas visuais e ortográficas. Após a fase de treinamento, os participantes foram testados com uma tarefa temporalizada de nomeação de figuras para investigar efeitos do recrutamento ortográfico na produção da fala. Os resultados sugeriram que a ortografia influenciou a nomeação das palavras aprendidas no treinamento, indicando que o processo de conversão de uma representação visual para suas representações fonoarticulatórias na produção da fala em L2 envolve a ativação ortográfica. Este resultado foi interpretado como um efeito de frequência da combinação grafo-fônica, que resultou em inabilidade para executar esta operação na rota sublexical. Assim, a presença de efeitos ortográficos nessa tarefa pode ser interpretada como evidência de que o sistema ortográfico pode funcionar como um mecanismo estratégico que auxilia na codificação lexical e, conseqüentemente, influencia o acesso lexical nos estágios iniciais da aquisição da linguagem em meios instrucionais.

**Palavras-chave:** aquisição fonológica; ortografia; psicolinguística.

Submitted on January 20th, 2020

Accepted on April 04th, 2020

## 1 Preliminary remarks

Research in the realm of psycholinguistics long ago made a strong case for investigating the effects of orthography on word processing abilities, both on the auditory and on the visual domains of language (FROST; KATZ, 1989). In this vein, studies have also looked into how orthography poses a major influence to speech perception and production (DAMIAN; BOWERS, 2003, 2009; QU; DAMIAN, 2019a; RASTLE *et al.*, 2011; ZIEGLER; FERRAND, 1998; ZIEGLER; FERRAND; MONTANT, 2004) in the light of the effects that literacy impinges on human cognition (KOLINSKY; PATTAMADILOK; MORAIS, 2012).

When it comes to acquiring an additional language later in life, it has been well documented that many learning hurdles arise from the fine-grained similarities and dissimilarities between different sound systems in contact that impinge on the perception of non-native sounds, leading to miscategorization and difficulties in production (ESCUADERO, 2011; SMILJANIC, 2011). More recently, orthography also started to be modeled as part of the knowledge that underlies L2 speech acquisition and processing (BASSETTI; ATKINSON, 2015; ESCUDERO, 2011; QU; DAMIAN, 2019b; VEIVO; JÄRVIKIVI, 2013; VEIVO *et al.*, 2018).

In language teaching, researchers have conceded that orthographic input should be treated as an empirical variable in the study of L2 phonological acquisition (BASSETTI; ESCUDERO; HAYES-HARB, 2015; SILVEIRA, 2012), as various traits of learners' phonological development, as observed in perception and production, can be traced back to their L1 orthography, thus rendering orthographic input to "filter" aural input (YOUNG-SCHOLTEN; LANGER, 2015). Hence, it is paramount to note that orthography, a cultural artifact resulting from human creation, plays a significant role in the constitution of phonological representations in the lexicon when subjects are schooled (VENTURA *et al.*, 2001), and introduced to L2 instructional contexts, being constantly exposed to copious amounts of written input (VEIVO; JÄRVIKIVI, 2013). Research has also revealed that new words are lexicalized, that is, they are integrated into the lexicon in functional manner, after exposure to their orthographic forms (SALETTA; GOFFMAN; BRENTARI, 2016), attesting to the powerful influence of orthographic knowledge on the adult lexicon.

The scope of the present investigation arises from the scenario described above. To observe whether orthography is recruited for L2 speech production with a sample of adult Brazilian-Portuguese bilingual speakers of English as an additional language, our study employed an exposure-based training paradigm, which consisted of study and verification blocks as exposure to an artificial lexicon that simulated English graphophonemic relations. Research subjects initially learned the phonological forms of a new lexicon through associations composed by pictures and their auditory forms to subsequently be introduced to their orthographic information. The new lexicon contained single-syllable words that compose experimental and control items, which differ in the consistency of the spelling-to-sound association in syllable nuclear

position. To observe whether orthographic effects arose in production, subjects completed a Timed Picture Naming task.

In the following section, the influence of orthography on phonological development is accounted for and studies investigating the role of orthography in L2 speech production with alphabetic languages are reviewed. Next, the complete study method is presented, including all criteria considered for stimuli preparation and the instruments and procedures involved in the data collection. Finally, results are presented and discussed.<sup>1</sup>

## 2 The role of orthography in phonological development

According to Veivo and Järvikivi (2013), there have been two main explanations to account for the activation of orthography during speech processing. One is regarded as the on-line *co-activation account*, which posits that orthographic and phonological representations co-exist and are strongly linked at both pre-lexical and lexical levels. As representations are linked, they can be activated automatically. The other account is the *restructuring account*, which claims that there are no separate representations for each of the systems. Instead, phonological representations that are pre-existing fundamentally change when one learns to read an alphabetic script. Thus, these representations, in nature, are abstract and

[...] amalgamate both orthographic and phonological information. As a consequence, orthographic effects during spoken word processing are taken as arising within the phonological system and resulting from these abstract phonological representations influenced by orthography (VEIVO; JÄRVIKIVI, 2013, p. 865).

Notwithstanding, Veivo and Järvikivi (2013) argue that a more plausible account is the *co-structuring account*, in which orthographic information contributes in parallel to the formation of lexical categories, along with phonological information. These representations then become co-structured with orthographic information because a functional link is established between orthographic and phonological representations with the attainment of literacy (KOLINSKY, 2015).

---

<sup>1</sup> The present study reports on the results of one of the experiments carried out by the first author in his PhD study (GONÇALVES, 2017).

Veivo and Järvikivi's (2013) claim can also be extended to the case of learning an L2. Their postulate allows one system to dominate over the other in specific cases, as with early learners in instructional settings, when orthography is believed to be more robust due to great amounts of written input, leading orthography to be regulatory over phonological encoding. Especially with an L2, the orthographic forms are learned either before or simultaneously to phonological forms, hence both of these systems are able to contribute to the formation of lexical entries, even if one is less autonomous than the other.

In this vein, Cutler (2015) argues that phonological representations in the lexicon are not compiled only from experience with speech perception. These representations are also distinguished by nonspeech, metalinguistic information such as visual-articulatory information (BERTELSON; VROOMEN; GELDER, 2003). The author has argued that "L2 learners exploit every type of help they can get with the language-learning task, and one result is that they set up phonological representations in the lexicon that include information that they have not extracted from the input" (CUTLER, 2008, p. 1607).

Another source of metalinguistic information is the recruitment of orthography to aid the construction of separate entries in the lexicon. Cutler (2015) posits that auditory perception of spoken items is not the only source for the storage of lexical distinctions, as some contrasts that are indistinguishable in perception can be recognized with the assistance of orthographic information, prompting learners to attempt a distinction in production. Relatedly, Saletta *et al.* (2016) attested for Cutler's claim when submitting that orthography induces the process of lexicalization of word forms because participants produced pseudowords more accurately after reading them, but not after just hearing them. This would show that new words would be integrated into the lexicon after subjects' repeated exposure to their written forms.

However, when lexical representations are implemented without their bondage to speech perception, the use of orthography can also represent a hindrance. Cutler (2015, p. 120) discusses that "incorporating a distinction at the lexical level without being able to perceive it in the lexicon works substantially against the learner's interest, in that it increases competition". For instance, when presented with novel auditory information that has not been previously mapped onto orthographic forms, the word <deaf> [dɛf] can compete with the prefix def- [dɛf]

and with daff- [dæf], giving rise in activation to a number of words that begin with such structures.

In general lines, orthographic effects can both benefit and hinder the formation of the lexicon. In perception, if distinctions are implemented in the lexicon solely based on orthographic information without a functional link to perception, this might hinder processing by increasing competition in processes of word recognition. However, orthography may also be advantageous for it compels the learner to attempt a developmental distinction in production that later may also aid the perception of that specific distinction.

### **3 Orthographic effects on L2 speech production**

Many studies have documented orthographic influences of different types in the domain of perception (CUTLER; DAVIS, 2012; ESCUDERO; WANROOIJ, 2010; ESCUDERO; SIMON; MULAK, 2014; ESCUDERO, 2015; HAYES-HARB; NICOL; BARKER, 2010; PYTLYK, 2011; SIMON; CHAMBLESS; ALVES, 2010) and little attention has been cast on production (ERDENER; BURNHAM, 2005; HAN; KIM, 2017) from a psycholinguistic perspective. Both studies that looked at cross-linguistic influences in production, with learners from alphabetic languages, are reviewed below.

Erdener and Burnham (2005) tested for the influence of auditory-visual stimuli on speech production by examining the production of pseudowords in a transparent<sup>2</sup> (Spanish) and in an opaque language (Irish) by Australian English and Turkish speakers. Native speakers of the tested languages recorded the stimuli and had their facial expressions video-recorded for the experiment. These participants were asked to repeat words as quickly as possible. They were tested within four conditions, including audio only, audio-visual (facial expressions), and audio plus orthography. In the conditions including orthography, they were also asked to write down the target item. The addition of visual information

---

<sup>2</sup> Orthographic systems can be classified either as transparent, when connections are isomorphic, i.e., a graphemic node is connected to only one phonological node, or as opaque, when a graphemic node is connected to several phonological alternatives (KATZ; FROST, 1992, 2001). Other terms have been employed, such as 'deep' and 'shallow', but we opted for 'transparent' and 'opaque' to guarantee consistency.

decreased the number of phoneme errors, irrespective of language background. However, such a result was only possible in the absence of orthography. One of the reasons for such a finding is, as Erdener and Burnham (2005) claim, the fact that the gesture is redundant enough to facilitate speech production, as they consider auditory-visual speech perception an ecologically valid process. On the other hand, the symbolic representations that connect speech to orthography are powerful enough that they affect basic auditory processes.

Erdener and Burnham (2005) also found that Turkish participants made fewer errors when the orthographic information presented to them was transparent. Yet, when such information was opaque, they were outperformed by their Australian English counterparts. This was also corroborated by their findings with the writing task in which Turkish participants made fewer spelling errors when they were presented with Spanish words than when they were shown Irish words. This suggests that these participants tend to process this type of input on the basis of grapheme-to-phoneme conversions. In contrast, the Australian participants, speakers of English, performed better with the Irish words, as both of these languages share opaque orthographies. Overall, these results show that participants from a transparent language background are likely to be misled when the orthographic information displayed does not match phonology straightforwardly. The authors conclude by saying that “when the target language has an opaque orthography, it seems better not to provide the learners with orthographic input, at least in the initial stages of exposure [...], and especially if they themselves have experience only with a transparent orthography” (ERDENER; BURNHAM, 2005, p. 219-220).

Han and Kim (2017) investigated the effect of orthography on the production of Korean allophones by Mandarin learners of the language. Sixty participants, who were paid to participate, were trained with 20 nonwords that contained /h/ in the syllabic onset of the second syllable where it can be produced similarly to [h̥] (a voiced fricative), [w] (a labial-velar approximant), or deleted. Training consisted of sound-picture associations split into nine sessions in four days. In each session, each target word was heard 10 times in randomized order, eight of which were realized with a deleted /h/, and the other two being the variants [h̥] and [w]. After the end of training, the orthographic representations of the nonwords were introduced in a single session in which participants would be exposed to the spellings along with the corresponding pictures.

Participants were grouped into three different groups that were exposed to specific spellings. The ‘ø-letter’ group was exposed to the deletion case in which the spelling <∅> corresponded to the null consonant; the ‘h-letter’ group were exposed to the spelling corresponding to the segment [h], <h>, and the other group was the ‘no-letter’ group, which received auditory input only. Stimulus presentation in this session was not timed, thus participants could spend as much time as they needed looking at the spellings. Testing took place with a picture naming task and a spelling recall task in which participants were required to write the spelling of each nonword.

For the analyses, participants were grouped differently according to two proficiency levels (beginner and advanced), which was measured through their experience studying Korean (length of learning). Results from picture naming demonstrated that participants tended to produce tokens influenced by the spelling they had been exposed to. The deletion group produced the novel words with allophonic variation only 10% of times, regardless of proficiency level. On the other hand, in the ‘no-letter’ group, participants with lower proficiency produced the nonwords with allophones only 5% of times, whereas participants of higher proficiency were able to reproduce allophonic variation in their speech 19% of times. However, the group which received exposure to the symbols that represented the allophones produced the novel words with the greatest allophonic variation, 20% of time - 23.5% for the beginners and 22% for advanced learners. Therefore, the spellings reinforced the allophonic variation that was presented auditorily during the training phase, thus influencing such a group of participants to produce it more often. This also shows that orthography can reinforce recently acquired phonological representations, as in the case of the allophones in their study. Bearing in mind the effects that orthography might pose to spoken production, we shall move on to the next section where the complete method of the present study will be presented.

#### **4 Method**

The guiding objective of the present study was to investigate whether speech production evokes orthographic knowledge, bearing in mind that the bilinguals’ first language consisted of a transparent script (Brazilian Portuguese), whereas their L2 was a language of an opaque

script (English). If orthography is recruited for naming, the study will be able to demonstrate that linguistic systems can act in an encapsulated manner, thus working strategically to assist speech production in naming. Therefore, we inquired: *Does orthography influence speech production?* If so, we hypothesized that response latencies from the Timed Picture Naming task would be affected by the level of orthographic transparency of the pseudowords used. Hence, to conduct the data analysis, the statistical model created included a continuous variable (response latencies from the naming task) and a categorical variable, orthographic consistency, with two levels (consistent and inconsistent).

The method in this study is a conceptual replication of previous studies which employed training with an artificial lexicon (ESCUDEIRO *et al.*, 2014; HAN; KIM, 2017; TAMMINEN; DAVIS; RASTLE, 2015; TAYLOR; DAVIS; RASTLE, 2017; RASTLE *et al.*, 2011), but specially that of Rastle *et al.* (2011), given that the design of their training tasks is replicated here. In this section, we explain all the criteria involved in the creation of the stimuli, both auditory and visual. Also, we present the study design and discuss the development of the experiments.

#### **4.1 Stimuli: The criteria to create the pseudowords**

This section explains in detail the factors that were taken into account for the creation of the words in this study, namely, the phonotactics, the target percepts and their spellings, and the phonological and orthographic neighborhood density of the words. Furthermore, the section presents the criteria and the steps used to create the picture and the auditory stimuli.

##### **4.1.1 Phonotactics and word length**

In the present study, all pseudowords created adhere to English phonotactics (BAUER, 2015) and have the same underlying syllabic structure, CVC. As previous research found no effects in naming latencies for mono and disyllabic words (DAMIAN *et al.*, 2010), the experimenters decided not to include word length as a factor to be tested. However, this factor is considered to be balanced in the stimuli, as words range from 3 to 5 letters in their written register, and all the targets have 3 phonemes. Moreover, the metric of orthographic depth is manipulated at the nucleus of these pseudowords. This was done so because English vowel sounds at

nucleus position may be spelt with a number of digraphs, thus enabling manipulations with different graphophone combinations to guarantee orthographic transparency or opacity, and also to guarantee the study reliability regarding its main focus of analysis.

#### 4.1.2 The target percepts and their spellings

To choose the percepts that integrate the nucleus of the target pseudowords, likely graphophonetic mappings that each vowel may have in English were considered. Two percepts were chosen: /i/, a vowel that can be mapped onto <ee>, <ea>, <ei>, and <eo>; and /ʌ/, which can be mapped onto <u>, <ou>, and <oo>. As the digraph <ee> is frequently associated with the tense high front vowel, it was considered a consistent pattern, therefore, a dominant spelling (ZIEGLER *et al.*, 2004), which is used as a control. It has been attested that the doubling <ee> reinforces duration as an acoustic trait that aids the detection of this vowel in English words by learners of different backgrounds (ESCUADERO, 2015; RAUBER, 2006). Thus, this digraph would give learners an advantage. The phoneme /ʌ/ is tested as control when it is consistently mapped onto a single grapheme, <u>.

To manipulate the consistency metric, the digraphs <ei> and <eo> for the percept /i/ were selected as the experimental opaque mappings, as these are less frequently associated with the target percept (ZIEGLER *et al.*, 2004). The digraph <ea> was not included for being also very frequently mapped onto /i/. As for /ʌ/, <ou> and <oo> were selected as the opaque experimental mappings. All of these four experimental digraphs can normally be mapped onto different vowels, which adds up to their degree of opacity in the experiment. As argued by Schmalz *et al.* (2016), multi-letter rules slow down sublexical processing because of a conflict between single-letter and grapheme pronunciations.

The acoustic proximity of the two vocalic categories in nuclear position assigned to the pseudowords was also considered. Selecting two percepts that were positioned close in the vocalic space of these learners and could somehow resemble each other would certainly make the learning experiment more difficult and would likely hinder the acquisition of the trained lexicon. Therefore, a high front vowel and a mid-central vowel were selected. The level of difficulty that these vowels generally pose for Brazilian learners was also observed. Rauber (2006) claims that the high front vowel pair (/i-I/) is the best distinguished in perception,

and the second best in production. Thus, at least when tested without its orthographic information, the tense high front vowel is a percept that generally poses little difficulty to Brazilians who hold a certain level of proficiency in the language. As in the case of the mid central vowel, Baptista (2006) explains that this category was the one that was most difficult for Brazilian learners to acquire in a target-like fashion when learning English in a naturalistic environment.

All words were presented in a balanced lexical environment. Each percept was orthographically represented by three different combinations of graphemes. Each combination was used in three different words, adding up to a total of 18 target items. Other four items containing the vowel /ɔ/ in nuclear position were used as distractors.<sup>3</sup> Thus, 22 items composed the stimuli in the learning phase, which can be seen in Appendix A. In order to guarantee that these were not actual English words, CLEARPOND<sup>4</sup> (MARIAN; BARTOLOTTI; CHABAL; SHOOK, 2012) was used. CLEARPOND is a user-friendly, access-free database, available in five languages, that allows for the identification of densities for both real words and pseudowords. Therefore, all pseudowords had their “non-word” status confirmed by searches on this database.

#### 4.1.3 Phonological and Orthographic Neighborhood density

All words that are phonologically similar to a certain item integrate their phonological neighborhood. Words with larger cohorts take longer to retrieve or might be recognized with more delay, given that many competitors might be activated by their phonological similarity. According to Fernández and Cairns (2011, p. 196), such a mechanism of retrieval is delayed as “more phonological information is required to specify uniquely a word from a dense neighborhood than from a sparse neighborhood”. No pseudowords used in training are homophones<sup>5</sup> to real words in order to avoid inadequate lexical selection and not to trigger direct lexical competition between two items that carry the same pronunciation.

---

<sup>3</sup> To know in more detail about the creation of the stimuli and piloting of the experiments, refer to Gonçalves (2017).

<sup>4</sup> Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities: <http://clearpond.northwestern.edu/>

<sup>5</sup> Homophony refers to words that have the same pronunciation, but differ in spelling and meaning (e.g., pair – pear).

Orthographic neighborhood is also controlled for. Any word integrates the orthographic neighborhood of a target word when it differs from it by a single letter, respecting length and letter position (VAN HEUVEN; DIJKSTRA; GRAINGER, 1998). During word recognition, different words can be activated non-selectively across languages in the lexicon when they share orthographic similarities, but not necessarily the same phonological characteristics. To identify phonological and orthographic neighbors, CLEARPOND was used.<sup>6</sup> Measures of lexical frequency are also available in CLEARPOND, as provided by the Subtlex<sup>7</sup> databases.

#### 4.1.4 Picture stimuli

Previous studies that grappled with orthographic effects have also successfully employed training paradigms in which subjects are compelled to associate pictures to pseudowords (BARTOLOTTI; MARIAN, 2017; ESCUDERO *et al.*, 2014; HAYES-HARB *et al.*, 2010; RASTLE *et al.*, 2011; SIMON *et al.*, 2010). Such a technique is advantageous for enabling testing without any sort of written exposure, thus providing an unbiased environment for observing the influence of orthography.

For the development of the visual stimuli that represent the pseudowords used in the experiments, initially three factors were taken into account. First, the drawings could not be so abstract in a way that remembering them would become too effortful. Second, the picture could not easily remind the learner of any other existing object, thus it should be something new. If any picture directly resembled any existing object, it would prompt learners of a clue for that specific word. Last, pictures could not be colored, as colors may lead to better memory performance with certain items, specially due to the fact that certain color combinations can produce higher levels of contrast, which influences memory retention (DZULKIFLI; MUSTAFAR, 2013). Figure 1 below is an example of a picture developed for the present study.

---

<sup>6</sup> For further details on how neighborhood density values were obtained, see Gonçalves (2017).

<sup>7</sup> Available at <http://crr.ugent.be/programs-data/subtitle-frequencies>

FIGURE 1 – The visual representation of the word “seeg”



Source: Gonçalves (2017, p. 87).

#### 4.1.5 Auditory stimuli

The auditory stimuli used in training and testing phases were recorded by a female native speaker of English who was invited to do the recordings, for which no compensation was involved. She was a 25-year-old from Herndon, Virginia (USA), who had been living in Brazil on a federal internship program. The recording session took place in the acoustic booth at *Laboratório de Fonética Aplicada* (FONAPLI). All stimuli were digitally recorded by using OCENAUDIO<sup>8</sup> version 2.0.14, at a sampling frequency of 44100 Hz in mono channel, with 16 bits resolution. The microphone was a dynamic, unidirectional SHURE (model SM48-LC). The computer used was an iMac 9.1.

The informant was instructed to read in natural speaking style. She was also explicitly instructed on how each set of words should be read in order to guarantee phonetic consistency in the recordings. Along with the words for reading aloud, the computer screen presented a note that informed real words to which the targets would be analogous, e.g., “geib” and “seeg” were analogous to “beat” and “beet”. She was allowed to rehearse before reading. To make sure phonetic consistency was guaranteed, each target word underwent an auditory and visual inspection on PRAAT carried out by the experimenters. If the word was appropriately produced by the speaker, the stimulus was edited on the same software and saved separately from other words.

---

<sup>8</sup> Cf. [Ocenaudio.com.br](http://Ocenaudio.com.br).

## 4.2 The study design

This study consists of two phases: a training and a testing phase. The training phase introduces participants to new spoken and written words that encompass transparent and opaque graphophonic mappings, before the testing phase. The testing phase consists of a timed picture naming experiment to measure orthographic effects in spoken production.

### 4.2.1 The training phase

In this phase, participants took part in study and verification blocks in which they were introduced to the study stimuli. Stimuli presentation was controlled with DMDX (FORSTER; FORSTER, 2003), version 5.1.3.6. (April 2016). Participants took the study and verification blocks in a quiet room, while sitting in front of a computer with a headset on.

The training session consisted of eight study and eight verification blocks. Each study block presented the stimuli three times, in three different sets. Among each set, participants were offered a short break. A verification block presented the stimuli twice, in two different sets, between which participants were offered a short break. The design of this scheme can be visualized in Table 1 below.

TABLE 1 – Study experimental design

Type of exposure	Design of treatment	
	<b>Phase 01: Training</b>	
<b>Auditory only</b>	Part 1: Study block 1 (three training sets)	Verification block 1 (two testing sets)
	Part 2: Study block 2 (three training sets)	Verification block 2 (two testing sets)
	Part 3: Study block 3 (three training sets)	Verification block 3 (two testing sets)
	Part 4: Study block 4 (three training sets)	Verification block 4 (two testing sets)

<b>Auditory + orthographic</b>	Part 5: Study block 5 (three training sets)	Verification block 5 (two testing sets)
	Part 6: Study block 6 (three training sets)	Verification block 6 (two testing sets)
	Part 7: Study block 7 (three training sets)	Verification block 7 (two testing sets)
	Part 8: Study block 8 (three training sets)	Verification block 8 (two testing sets)
<b>Phase 2: testing</b>		
<b>No input offered</b>	Timed Picture Naming Task	

Source: Gonçalves (2017, p. 89-90)

In study blocks, participants were shown a picture of a novel object while listening to its spoken form over headphones. They needed to repeat the object’s name after each trial. This was to guarantee articulatory encoding of the new lexical representations and to observe whether participants were paying attention to the stimuli presentation. In order to familiarize the participant with the procedure, three trials were provided as a familiarization block. The stimuli consisted of the 22 new words (Appendix A) which were presented twelve times each, adding up to a total of 264 trials split into eight different study blocks during training.

A participant firstly took part in three training sets in one study block, with a total of 66 trials, which were then followed by a verification block with two testing sets. Each trial presentation in a study block lasted 2000ms to allow for object recognition and phonological encoding. This duration is comparable to previous research involving training on new lexical items (2000ms: BARTOLOTTI; MARIAN, 2017; ESCUDERO, 2015; SIMON *et al.*, 2010). The participants were explicitly instructed to repeat each spoken form while paying attention to the visual form that was presented simultaneously on the computer screen. For the final data collection, no response was registered from study blocks.

After each study block, each subject took part on two testing sets in a verification block. Verification blocks consisted of a Picture Identification Task in which participants needed to choose, from two pictures displayed on the computer screen, the one that matched the stimuli heard. Feedback was given immediately for wrong responses with the message “*Wrong response! Try harder!*”. Each trial was available for

5000ms before time out occurred in case the participant did not respond. In such a case, the message “No response” was displayed on the screen before the next trial came up. Four practice trials were provided to familiarize the participant with the experiment before the presentation of the verification block started. Each verification block in the Picture Identification Task contained 44 trials, divided into two testing sets.

Beginning with the fifth study block, participants were exposed to the lexicon written forms in conjunction with the spoken forms and the picture in study blocks. The procedure was very similar to the protocol followed with the first four study blocks. Each trial lasted 2200ms to allow for picture recognition, and orthographic and phonological encoding. These extra 200ms were allowed to present orthographic input, thus entailing in one additional process that was not present in the first four parts of training. After three study sets in a study block, participants were required to take the Picture Identification Tasks with two testing sets in a verification block in which they needed to select the target, from two pictures displayed on the screen, which matched the stimulus heard. Feedback was given immediately for wrong responses with the message “*Wrong response! Try harder!*”. Each trial was available for 5000ms, before time out occurred in case the participant did not respond. In such a case, the message “No response” was displayed on the screen before the next trial came up.

Responses from the Picture Identification Tasks were used to observe how well participants performed in each stage of the training phase. Participants were informed of their progress as soon as they completed each Picture Identification Task and were explicitly told that they needed to reach at least 80% of correct responses in order to move to the testing phase.

#### **4.2.2 The testing phase: Timed picture naming task**

In picture naming, participants are required to generate a matching spoken form to the picture presented on the computer screen. Any delay or wrong responses might be due to intervening factors that influenced the retrieval and encoding of that word form from long-term memory. Jiang (2012) states that picture naming involves three major cognitive processes, namely: object recognition, conceptual activation, and lexical access and production. The author also claims that this task has been used “to examine common and unique properties of lexical

access in L2” (p. 148), as is the case of the present study. Accordingly, participants are instructed to name all 22 pictures seen in training as rapidly and as accurately as possible. Stimuli presentation and recording of voice responses is done with DMDX, which takes .rtf script files as input.

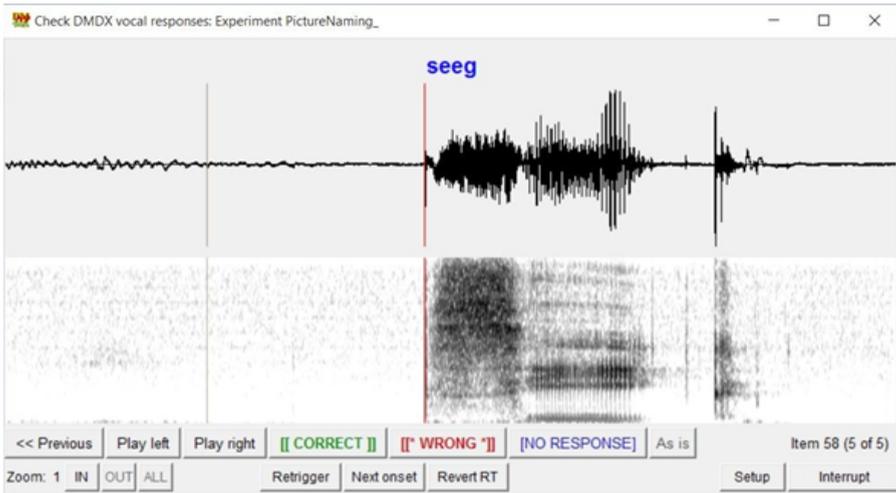
To familiarize the participant with the procedure, this experiment consists of a practice block, with six trials, each with a word from the study. In sequence, there are four different blocks, with 22 trials each, when the 22 words used in the study are presented in the automatic randomized order DMDX applies. After each block, the participant might choose to take a short break. A trial consists of a fixation point and a target picture. The fixation point lasts for 500ms, which is then followed by a target stimulus that stays on screen for 2500ms, when the time out occurs. Pictures are positioned to the center of the screen. The recording of voice responses is done by the headset the participant wears while taking the experiment. As Jiang (2012) advises, sensitivity of the voice key is adjusted to a medium level because normal vocalization can provide enough energy to trigger the voice key. If the sensitivity is too high, low-volume noise may stop the timer, causing RTs to be very short.

To analyze naming data, the oral responses were scored offline with CheckVocal (PROTOPAPAS, 2007). This is a software developed to help process the results of naming tasks from DMDX. It checks for accuracy (correct, wrong or no responses) and timing (to see if the voice key is properly triggered). The program takes three different files as input. The .azk file that DMDX generates with the RTs from a given participant, along with a previously prepared .txt file containing the answers in written form for each trial number. For instance, if the picture naming has four blocks with 22 trials each, the answer file needs to present 88 answers, each in a different line, properly numbered according to the trial to which it belongs. The last input file CheckVocal takes is the DMDX script that is written to run the naming experiment.

As shown in Figure 2 below, CheckVocal displays each naming waveform and spectrogram with the voice-key mark. On the top of the screen, it also shows the expected answer for any trial in written form. The experimenter needs to observe if the timing mark is properly placed on the onset of the voice response. If not, in case it has been mistriggered to a premature onset because of lip smacking, or late-triggered because of low-volume in the oral response, the experimenter can click on the

waveform to reset the voice-key trigger. The software also presents an option that automatically re-triggers the mark to a subsequent onset. Figure 2 displays the inspection of the word “seeg” on CheckVocal.

FIGURE 2 – Inspection of “seeg” on CheckVocal



Source: Gonçalves (2017, p. 96)

After inspecting the placing of the timing mark, the experimenter needs to check for the accuracy of the participant’s response. The software displays three buttons on the bottom of its screen, for “correct”, “wrong” and “no response” options. When the response is wrong, a negative mark is assigned to that RT. When there is no response, that RT is automatically set to -2500ms in the data list answer-file it generates. In order to define for correctness of voice responses, all phonemes of a given word need to be correct in initial, medial, and final position. Slips of the tongue are to be considered as wrong responses. When this inspection is over, the output of CheckVocal is a data list .txt file that shows different rows with each trial number followed by its corrected RT.

### 4.3 Participants

Thirty-six participants took part of the final data collection for this study. They all volunteered and were mostly recruited from undergraduate classes of the *Letras* program at Universidade Federal de

Santa Catarina. Some participants were also recruited through personal contacts of the experimenters. In this phase, participation consisted of one data collection encounter which, consecutively, started with a training phase whose objective consisted of participants' learning the artificial lexicon, followed by a testing phase.

The participants were thirteen men and twenty-three women, whose ages varied from 18 to 47 ( $M: 26,1$ ). They all had normal speech and hearing, and normal or corrected-to-normal vision. Participants spoke Brazilian Portuguese as their first language and learned English as an additional language. They all self-reported to have a minimum intermediate proficiency level in the additional language. It was also required that participants were right-handed, given that response times were registered by using particular keys for the dominant hand across a series of different tasks. More detailed information about participants has been provided in Appendix B.

#### 4.4 Procedures

This study was approved by the Ethics Research Board of the university where it was conducted.<sup>9</sup> Participants encountered individually with the experimenter. The data collection took place in a quiet room, with participants sitting in a comfortable chair. The headset volume was adjusted to a comfortable listening level. A Microsoft LifeChat headset was used for auditory presentation and the recording of oral responses, and an Avell notebook was used to administer all the experiments. Firstly, participants were given the Consent Form and took part in the first training session. Next, they moved onto the second training session. Finally, participants were tested with Timed Picture Naming task, besides being given a questionnaire that gathered some information on how they learned English and other additional foreign languages. At the beginning of all encounters, it was emphasized that answers should be given as quickly and as accurately as possible for when they were tested. Data collection sessions varied from 100 to 140min, depending on the amount of breaks

---

<sup>9</sup> This research was certified in *Sistema Nacional de Informações sobre Ética em Pesquisa Envolvendo Seres Humanos* (SISNEP) under the register 54197716.5.0000.0121. It was approved by UFSC Ethics Committee as attested by the register 1.518.285 issued on April 26<sup>th</sup>, 2016.

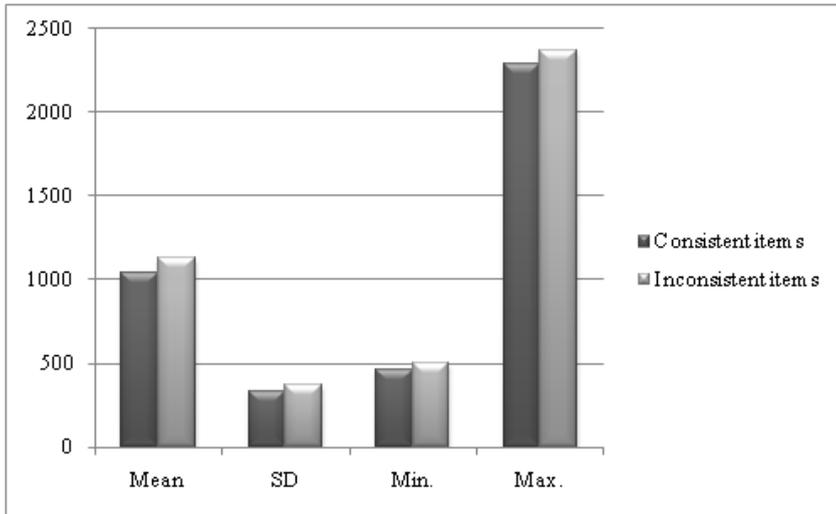
participants decided to take. All participants received a certificate for hours of participation.

## **5 Results and discussion**

To conduct the data analysis for the Timed Picture Naming task, the statistical model included a continuous variable (response latencies), along with orthographic consistency, with two levels (consistent and inconsistent). The data spreadsheets were inspected for any data cases with negative and “no response” latencies (latencies that are automatically set to 2500ms by CheckVocal), which were excluded. Participants scored 68% of valid responses ( $N$ : 2141 data cases containing correct responses with no time out values). Participants timed out on 13% of trials (4,29% with consistent items and 9,62% with inconsistent items). Missing values were unchanged, and the data were analyzed with multi-level statistical models, which are better equipped to analyze missing values (LACHAUD; RENAUD, 2011).

Results displayed in Figure 3 below show that participants scored a mean time of 1085ms to produce oral responses, with latencies ranging from 464 to 2366ms in response time. The average standard deviation reached 354ms. To inspect any differences in performance according to orthographic condition, data were analyzed separately for control (consistent) or experimental (inconsistent) items. Figure 3 demonstrates that items from the control condition, were named faster than items from the experimental condition (consistent: 1038ms; inconsistent: 1134ms), as seen in previous studies with naming (CORTESE; SIMPSON, 2000). Moreover, standard deviation values did not vary as much in the control condition when compared to the experimental condition, attesting that participants varied to a lesser extent when producing consistent words (consistent: 332ms; inconsistent: 370ms).

FIGURE 3 – Descriptive statistics for the Timed Picture Naming task (in ms)



Source: Authors

Tests of normality indicated that response latencies did not achieve normal distribution across the two conditions ( $p: .000$ ). Thus, a Man-Whitney U test was run to observe whether orthographic consistency was affecting subjects' performance in naming the lexicon learned during training. The probability value achieved significance ( $Z: -6,343$ ;  $p: .000$ ), thus demonstrating that orthography influenced picture naming with this sample of subjects. Therefore, as inquired by our research question, it can be argued that the process of converting the visual input into its phono-articulatory representations for production, which is mediated by lexical selection, involves the activation of orthographic codes, corroborating the hypothesis that for second language learners, orthography acts as a compensatory mechanisms that assists lexical selection in speech production. By calling it a compensatory mechanism, we argue that it compensates for lack of skill in computing the graphophonic combinations used in the inconsistent lexicon in the present study.

In this vein, it is important to note that these orthographic effects might be due to a frequency effect caused by the graphophonic combinations used in the stimuli. Once that a new graphophonic combination was encoded by the subject, the orthographic information of this combination would be recruited in tandem with the phonological

information as a way of “assisting” lexical access for recently established lexical categories that might still be unstable. This echoes previous research that claimed that inconsistent mappings would affect subjects’ performance in phonological tasks (ESCUADERO *et al.*, 2008; ESCUADERO *et al.*, 2014; HAYES-HARB *et al.*, 2010). However, the reason for this effect might not rely on the inconsistency of the graphophonic combination, considering that this sample of subjects is highly literate, but on the infrequency of such a combination. This could be regarded as an effect of lack of skill for computing such associations in the sublexical route. The degree of activation of orthography in this particular case is rendered higher because of the low graphophonic frequency, thus motivating an orthographic effect.

Research has argued that idiosyncrasies between languages orthographic depths can highly impact auditory and visual processing (FROST, 1992, 1998, 2005; FROST; ZIEGLER, 2007; ZIEGLER; FERRAND, 1998; ZIEGLER *et al.*, 2004; ZIEGLER; MUNEUX, 2007). Thus, given that Brazilian Portuguese is considered a relatively transparent language if compared to English opacity, graphophonic frequency, that is, the frequency to which a grapheme maps onto a phoneme, affected subjects’ phonological processing. It is paramount to note here that the opaque mappings in this experiment were based on multiple conversion rules with which these participants did not share the same level of experience in operating grapho-phonetic conversions when compared to the same skill in their first language.

In order to analyze the incorrect responses of this task, the number of incorrect oral responses was calculated for each word by using the function “crosstabs” on SPSS. By conducting a visual inspection of the results displayed in Table 2, the pseudowords “geop” (99), “doup” (87) and “pood” (65), all of which present with inconsistent orthography, accounted for the highest number of incorrect responses.

TABLE 2 – Words along with number of errors and neighborhood sizes

Nonword	<i>N</i> of errors in production	Orthographic <i>N</i> -size	Orthographic <i>N</i> -frequency	Phonological <i>N</i> -size	Phonological <i>N</i> -frequency
geesh	28	2	3,000	9	3,000
keet	18	8	3,375	34	3,088
aegg	25	7	2,857	17	3,000
deit	49	5	3,600	26	3,000
geib	52	1	4,000	4	3,000
meip	36	1	1,980	20	3,000
geop	99	1	3,000	15	3,000
teog	72	0	0	11	3,000
teob	43	0	0	11	3,000
bup	48	16	2,938	21	3,048
nup	32	9	0	13	3,000
sud	38	17	3,176	32	3,125
doup	87	2	3,000	22	3,045
soug	36	8	3,375	25	3,120
toud	49	6	3,000	28	3,036
dood	55	10	2,889	32	3,045
pood	65	16	2,938	22	3,045
loob	33	7	2,857	20	3,050

Source: Gonçalves (2017, p. 117-118)

Table 3 also displays orthographic and phonological neighborhood sizes for each tested word. To observe if neighborhood size, both orthographic and phonological, motivated any sort of lexical competition that could result in errors in word retrieval for spoken production, the number of incorrect responses of each word was correlated to all neighborhood measures. Spearman correlations did not achieve significance for any of the correlated variables (orthographic size:  $\rho$ : -.253,  $p$ : .312; orthographic frequency:  $\rho$ : .074,  $p$ : .770; phonological size:  $\rho$ : -.023,  $p$ : .927; phonological frequency:  $\rho$ : -.126,  $p$ : .618).

Next, in another attempt to observe which variable motivated the errors in production the number of errors in production and the neighborhood sizes were recoded into a total, according to the type of digraph that the lexicon encompassed. Spearman correlations were run again to observe whether the type of digraph in each word could be correlated to the neighborhood sizes and to the number of errors in the naming task. Such correlations did not achieve significance (orthographic

neighborhood size:  $\rho$ : -.600,  $p$ : .208; phonological neighborhood size:  $\rho$ : -.029,  $p$ : .967). However, it can be interestingly noted that the graphophonic combination with the most number of errors in oral production is the one with the smallest orthographic neighborhood (<eo>, which accounts for 214 incorrect responses and only one orthographic neighbor). This shows that the lack of subjects' familiarity with this specific graphemic string resulted in errors in the conversion of such a combination to its phonological components preceding production, which can be interpreted as evidence to the fact that orthographic information actively influenced subjects' generation of spoken responses to this specific item and throughout the task, in general. This also demonstrates that the orthographic effect was resulting of lack of skill in the sublexical route to compute such visual information onto its phonological components.

As concerns subjects' incorrect answers provided for the task, they consist of similar-sounding words or of items that shared orthographic components in the syllable with the target pseudowords. To take a case in point, "calm" is a frequent word that was provided in many answers instead of the target pseudoword "galm", which was used as a distractor in the dataset. Moreover, other frequent words that shared the same onset and coda with the target pseudowords of the study were also provided: "dude" and "dad" instead of the target "dood"; "nap" instead of "nup", "sad" for "sud", "gum" for "galm". Taken together, these results illustrate evidence for the fact that subjects were able to encode the orthographic form of these new lexical items presented over training. However, due to lack of vast experience with the trained items that resulted in unstable lexical categories, or perhaps interference in processing due to lexical competition among the trained items and real words, subjects ended up providing these frequent words as responses for they were more strongly activated for naming.

## 6 Final remarks

The Timed Picture Naming task indicated that orthographic consistency influenced subjects' latencies in naming the trained lexicon. We argued that lexical selection involved the activation of orthographic codes as if orthography were a compensatory mechanism to assist lexical selection in speech production, at least for recently learned words in

experimental conditions. In such enterprise, we pointed out that this orthographic effect might have been due to a frequency effect because of the infrequency of the graphophonetic combinations used in the lexicon so that phonological and orthographic information would be recruited in tandem as a way of assisting lexical access for recently established lexical categories.

To warrant further research, psycholinguists still need to unveil whether the paradigm investigated here reflects a strategic, problem-solving operation or if such a mechanism belongs to long-term knowledge and is invariable to language activation (TAYLOR; DAVIS; RASTLE, 2017). A different study design could be able to track whether the recruitment of orthography is committed to learning conditions, or if it engages in everyday, more naturalistic tasks of language use, in which attention is not so stimulus-driven (such as listening to music, watching TV etc.). Moreover, a control group tested only within the L1 would be able to indicate whether such an effect is language-specific (for example, in picture naming in Brazilian-Portuguese), or motivated by the idiosyncrasies of the two orthographic systems in contact with bilingual speakers. In this vein, an ampler statistical model could factor in participants' individual differences regarding proficiency level and age of acquisition in the L2 to observe whether these are intervening factors in the acquisition of the artificial lexicon and in orthographic activation as a strategic mechanism that aids phonological processing in this type of experimental setting.

### **Acknowledgements**

We would like to thank all participants who took part in the different study phases, since the piloting of experiments to the final data collection sessions. We would also like to express our gratitude to Giovanni Luigi Piazza for creating all the visual stimuli, the FONAPLI staff for providing the equipment for recording the auditory stimuli, and Professor Hanna Kivistö-de Souza for inspecting the pseudowords created and the DMDX scripts. Finally, we thank Professor Mailce Borges Mota for all suggestions with the background literature and the study design as the study was being developed. We also feel indebted to Professor Markus Damian for the guidance with DMDX programming during his stay in Brazil. We are also grateful to CNPq for the research grants awarded to both authors.

### Authorship statement

This study reports on data from Alison Roberto Gonçalves' Doctoral dissertation, which was supervised by Professor Rosane Silveira. Both authors worked together to design the experiment and complete the data analysis, including the statistical analysis. The first author was in charge of gathering data, transferring the data to spreadsheets for data analysis, and writing the first draft of the article. Both authors collaborated on interpreting results and revising the article.

### References

BAPTISTA, B. O. Adult Phonetic Learning of a Second Language Vowel System. In: BAPTISTA, B. O.; WATKINS, M. A. (org.). *English with a Latin Beat: Studies in Portuguese/Spanish – English Interphonology*. Amsterdam: John Benjamins, 2006. p. 19-40. DOI: <https://doi.org/10.1075/sibil.31.03bap>

BARTOLOTTI, J.; MARIAN, V. Bilinguals' Existing Languages Benefit Vocabulary Learning in a Third Language. *Language Learning*, v. 67, n. 1, p. 110-140, 2017. DOI: <https://doi.org/10.1111/lang.12200>

BASSETTI, B.; ATKINSON, N. Effects of Orthographic Forms on Pronunciation in Experienced Instructed Second Language Learners. Orthographic Effects in Second Language Phonology. *Applied Psycholinguistics*, Cambridge, v. 36, n. 1, p. 67-91, 2015. DOI: <https://doi.org/10.1017/S0142716414000435>

BASSETTI, B.; ESCUDERO, P.; HAYES-HARB, R. Second Language Phonology at the Interface between Acoustic and Orthographic Input. *Applied Psycholinguistics*, Cambridge, v. 36, n. 1, p. 1-6, 2015. DOI: <https://doi.org/10.1017/S0142716414000393>

BAUER, L. English Phonotactics. *English Language and Linguistics*, Cambridge, v. 19, n. 3, p. 437-475, 2015. DOI: <https://doi.org/10.1017/S1360674315000179>

BERTELSON, P.; VROOMEN, J.; GELDER, B. Visual Recalibration of Auditory Speech Identification: A McGurk Aftereffect. *Psychological Science*, [S.l.], v. 14, n. 6, p. 592-597, 2003. DOI: [https://doi.org/10.1046/j.0956-7976.2003.psci\\_1470.x](https://doi.org/10.1046/j.0956-7976.2003.psci_1470.x)

CORTESE, M. J.; SIMPSON, G. B. Regularity Effects in Word Naming: What Are They? *Memory & Cognition*, [S.l.], v. 28, p. 1269-1276, 2000. DOI: <https://doi.org/10.3758/BF03211827>

CUTLER, A. Representation of Second Language Phonology. *Applied Psycholinguistics*, Cambridge, v. 36, n. 1, p. 115-128, 2015. DOI: <https://doi.org/10.1017/S0142716414000459>

CUTLER, A. The Abstract Representations in Speech Processing. *Quarterly Journal of Experimental Psychology*, [S.l.], v. 61, n. 11, p. 1601-1619, 2008. DOI: <https://doi.org/10.1080/13803390802218542>

CUTLER, A.; DAVIS, C. An Orthographic Effect in Phoneme Processing, and Its Limitations. *Frontiers in Psychology*, Brussels, Belgium, v. 3, n. 18, p. 1-7, 2012. DOI: <https://doi.org/10.3389/fpsyg.2012.00018>

DAMIAN, M. F. Effects of Orthography on Speech Production in a Form-Preparation Paradigm. *Journal of Memory & Language*, Egham, UK, v. 49, p. 119-132, 2003. DOI: [https://doi.org/10.1016/S0749-596X\(03\)00008-1](https://doi.org/10.1016/S0749-596X(03)00008-1)

DAMIAN, M. F. BOWERS, J. S.; STADTHAGEN-GONZALEZ, H.; SPALEK, K. Does Word Length Constrain the Spoken Production of Single Words? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, [S.l.], v. 36, p. 892-905, 2010. DOI: <https://doi.org/10.1037/a0019446>

DAMIAN, M. F.; BOWERS, J. S. Effects of Orthography on Speech Production in a Form-Preparation Paradigm. *Journal of Memory & Language*, Egham, UK, v. 49, p. 119-132, 2003. DOI: [https://doi.org/10.1016/S0749-596X\(03\)00008-1](https://doi.org/10.1016/S0749-596X(03)00008-1)

DAMIAN, M. F.; BOWERS, J. S. Assessing the Role of Orthography in Speech Perception and Production: Evidence from Picture-Word Interference Tasks. *European Journal of Cognitive Psychology*, [S.l.], v. 21, p. 581-598, 2009. DOI: <https://doi.org/10.1080/09541440801896007>

DZULKIFLI, M. A.; MUSTAFAR, M. F. The Influence of Colour on Memory Performance: A Review. *The Malaysian Journal of Medical Sciences*, Pulau Pinang, Malaysia, v. 20, n. 2, p. 3-9, 2013.

EDERNER, V. D.; BURNHAM, D. K. The Role of Audiovisual Speech and Orthographic Information in Nonnative Speech Production.

*Language Learning*, [S.l.], v. 55, n. 2, p. 191-228, 2005. DOI: <https://doi.org/10.1111/j.0023-8333.2005.00303.x>

ESCUADERO, P. Orthography Plays a Limited Role when Learning the Phonological Forms of New Words: The Case of Spanish and English Learners of Novel Dutch Words. *Applied Psycholinguistics*, Cambridge, v. 36, n. 1, p. 7-22, 2015. DOI: <https://doi.org/10.1017/S014271641400040X>

ESCUADERO, P. Speech Processing in Bilingual and Multilingual Listeners: The Interrelation between Sound Perception and Word Recognition. In: COHN, A. C.; FOUGERON, C.; HUFFMAN, M. K. (org.). *The Handbook of Laboratory Phonology*. New York: Oxford University Press, 2011. p. 407-417.

ESCUADERO, P.; HAYES-HARB, R.; MITTERER, H. Novel L2 Words and Assymetric Lexical Access. *Journal of Phonetics*, [S.l.], v. 36, n. 2, p. 345-360, 2008. DOI: <https://doi.org/10.1016/j.wocn.2007.11.002>

ESCUADERO, P.; SIMON, E.; MULAK, K. Learning Words in a New Language: Orthography Doesn't always Help. *Bilingualism, Language and Cognition*, Cambridge, v. 17, n. 2, p. 384-395, 2014. DOI: <https://doi.org/10.1017/S1366728913000436>

ESCUADERO, P.; WANROOIJ, K. The Effect of L1 Orthography on L2 Vowel Perception. *Language and Speech*, [S.l.], v. 53, n. 3, p. 343-365, 2010. DOI: <https://doi.org/10.1177/0023830910371447>

FERNÁNDEZ, E. M.; CAIRNS, H. S. *Fundamentals of Psycholinguistics*. West Sussex: Wiley-Blackwell, 2011.

FORSTER, K. I.; FORSTER, J. C. DMDX: A Windows Display Program with Millisecond Accuracy. *Behavior Research Methods, Instruments, & Computers*, [S.l.], v. 35, p. 116-124, 2003. DOI: <https://doi.org/10.3758/BF03195503>

FROST, R. Orthography and Phonology: The Psychological Reality of Orthographic Depth. In: NOONAN, M.; DOWNING, P.; LIMA, S. (org.). *The Linguistics of Literacy*. Amsterdam/Philadelphia: John Benjamins, 1992. p. 255-274. DOI: <https://doi.org/10.1075/tsl.21.19fro>

FROST, R. Towards a Strong Phonological Theory of Visual Word Recognition: True Issues and False Trails. *Psychological Bulletin*,

Washington, v. 123, p. 71-99, 1998. DOI: <https://doi.org/10.1037/0033-2909.123.1.71>

FROST, R. Orthographic Systems and Skilled Word Recognition Processes in Reading. In: HULME, C.; SNOWLING, M. (org.). *The Science of Reading: A Handbook*. Malden, Mass: Blackwell, 2005. p. 272-295. DOI: <https://doi.org/10.1002/9780470757642.ch15>

FROST, R.; KATZ, L. Orthographic Depth and the Interaction of Visual and Auditory Processing in Word Recognition. *Memory & Cognition*, [S.l.], v. 17, p. 302-311, 1989. DOI: <https://doi.org/10.3758/BF03198468>

FROST, R.; ZIEGLER, J. C. Speech and Spelling Interaction: The Interdependence of Visual and Auditory Word Recognition. In: GASKELL, M. G. (org.). *The Oxford Handbook of Psycholinguistics*. Oxford: Oxford University Press, 2007. p. 107-118. DOI: <https://doi.org/10.1093/oxfordhb/9780198568971.013.0007>

GONÇALVES, A. R. *The Orthographic Signature in Second Language Speech Acquisition and Processing*. 2017. 156f. Tese (Doutorado em Inglês: Estudos Linguísticos) – Programa de Pós-Graduação em Inglês: Estudos Linguísticos e Literários, Universidade Federal de Santa Catarina, 2017.

HAN, J. I.; KIM, J. Y. The Influence of Orthography on the Production of Alphabetic, Second-Language Allophones by Speakers of a Non-Alphabetic Language. *Psycholinguist Research*, [S.l.], v. 46, n. 4, p. 963-982, 2017. DOI: <https://doi.org/10.1007/s10936-016-9474-7>

HAYES-HARB, R.; NICOL, J.; BARKER, J. Learning the Phonological Forms of New Words: Effects of Orthographic and Auditory Input. *Language and Speech*, [S.l.], v. 53, n. 3, p. 367-381, 2010. DOI: <https://doi.org/10.1177/0023830910371460>

JIANG, N. *Conducting Reaction Time Research in Second Language Studies*. New York: Routledge, 2012. DOI: <https://doi.org/10.4324/9780203146255>

KATZ, L.; FROST, R. The Reading Process is Different for Different Orthographies: The Orthographic Depth Hypothesis. In: FROST, R.; KATZ, L. (org.). *Orthography, Phonology, Morphology, and Meaning*. Amsterdam: Elsevier North Holland Press, 1992. p. 67-84. DOI: [https://doi.org/10.1016/S0166-4115\(08\)62789-2](https://doi.org/10.1016/S0166-4115(08)62789-2)

KATZ, L.; FROST, S. Phonology Constrains the Internal Orthographic Representation. *Reading and Writing: An Interdisciplinary Journal*, [S.l.], v. 14, n. 3/4, p. 297-332, 2001. DOI: <https://doi.org/10.1023/A:1011165407770>

KOLINSKY, R. How Learning to Read Influences Language and Cognition. In: POLLATSEK, A.; TREIMAN, R. (org.). *The Oxford Handbook of Reading*. Oxford: Oxford University Press, 2015. p. 377-393.

KOLINSKY, R.; PATTAMADILOK, C.; MORAIS, J. The Impact of Orthographic Knowledge on Speech Processing. *Ilha do Desterro*, Florianópolis, v. 63, p. 161-186, 2012. DOI: <https://doi.org/10.5007/2175-8026.2012n63p161>

LACHAUD C. M.; RENAUD, O. A Tutorial for Analysing Human Reaction Times: How to Filter Data, Manage Missing Values, and Choose a Statistical Model. *Applied Psycholinguistics*, Cambridge, v. 32, p. 389-416, 2011. DOI: <https://doi.org/10.1017/S0142716410000457>

MARIAN, V.; BARTOLOTTI, J.; CHABAL, S.; SHOOK, A. CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLoS ONE*, [S.l.], v. 7, n. 8, [s.p.], 2012. DOI: <https://doi.org/10.1371/journal.pone.0043230>

PROTOPAPAS, A. CheckVocal: A Program to Facilitate Checking the Accuracy and Response Time of Vocal Responses from DMDX. *Behavior Research Methods*, [S.l.], v. 39, n. 4, p. 859-862, 2007. DOI: <https://doi.org/10.3758/BF03192979>

PYTLYK, C. Shared Orthography: Do Shared Written Symbols Influence the Perception of L2 Sounds? *The Modern Language Journal*, [S.l.], v. 95, n. 4, p. 541-557, 2011. DOI: <https://doi.org/10.1111/j.1540-4781.2011.01244.x>

QU, Q.; DAMIAN, M. Orthographic Effects in Mandarin Spoken Language Production. *Memory and Cognition*, [S.l.], v. 47, p. 326-334, 2019a. DOI: <https://doi.org/10.3758/s13421-018-0868-7>

QU, Q.; DAMIAN, M. The Role of Orthography in Second-Language Spoken Word Production: Evidence from Tibetan Chinese Bilinguals. *Quarterly Journal of Experimental Psychology*, [S.l.], v. 72, p. 2597-2604, 2019b. DOI: <https://doi.org/10.1177/1747021819850382>

RASTLE, K.; MCCORMICK, S. F.; BAYLISS, L.; DAVIS, C. J. Orthography Influences the Perception and Production of Speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, [S.l.], v. 37, n. 6, p. 1588-1594, 2011. DOI: <https://doi.org/10.1037/a0024833>

RAUBER, A. S. *Perception and Production of English Vowels by Brazilian EFL Speakers*. 2006. 218f. Tese (Doutorado em Letras) – Programa de Pós-Graduação em Inglês: Estudos Linguísticos e Literários, Universidade Federal de Santa Catarina, 2006.

SALETTA, M.; GOFFMAN, L.; BRENTARI, D. Reading Skill and Exposure to Orthography Influence Speech Production. *Applied Psycholinguistics*, Cambridge, v. 37, n. 2, p. 411-434, 2016. DOI: <https://doi.org/10.1017/S0142716415000053>

SCHMALZ, X. *et al.* Unpredictability and Complexity of Print-to-Speech Correspondences Increase Reliance on Lexical Processes: More Evidence for the Orthographic Depth Hypothesis. *Journal of Cognitive Psychology*, [S.l.], v. 28, n. 6, p. 658-672, 2016. DOI: <https://doi.org/10.1080/20445911.2016.1182172>

SILVEIRA, R. L2 Production of English Word-Final Consonants: The Role of Orthography and Learner Profile Variables. *Trabalhos em Linguística Aplicada*, Campinas, v. 51, n. 1, p. 15-28, 2012. DOI: <https://doi.org/10.1590/S0103-18132012000100002>

SIMON, E.; CHAMBLESS, D.; ALVES, U. K. Understanding the Role of Orthography in the Acquisition of a Non-Native Vowel Contrast. *Language Sciences*, [S.l.], v. 32, n. 3, p. 380-394, 2010. DOI: <https://doi.org/10.1016/j.langsci.2009.07.001>

SMILJANIC, R. Second-Language Speech Learning. In: COHN, A. C.; FOUGERON, C.; HUFFMAN, M. K. (org.). *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press, 2011. p. 417-425.

TAMMINEN, J.; DAVIS, M. H.; RASTLE, K. From Specific Examples to General Knowledge in Language Learning. *Cognitive Psychology*, [S.l.], v. 79, p. 1-39, 2015. DOI: <https://doi.org/10.1016/j.cogpsych.2015.03.003>

TAYLOR J. S. H.; DAVIS, M. H.; RASTLE, K. Comparing and Validating Methods of Reading Instruction Using Behavioural and Neural Findings in an Artificial Orthography. *Journal of Experimental*

*Psychology*, Bethesda, MD, v. 146, n. 6, p. 826-858, 2017. DOI: <https://doi.org/10.1037/xge0000301>

VAN HEUVEN, W. J. B.; DIJKSTRA, A.; GRAINGER, J. Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, [S.l.], v. 39, p. 458-483, 1998. DOI: <https://doi.org/10.1006/jmla.1998.2584>

VEIVO, O.; JÄRVIKIVI, J. Proficiency Modulates Early Orthographic and Phonological Processing in L2 Spoken Word Recognition. *Bilingualism: Language and Cognition*, Cambridge, v. 16, n. 4, p. 864-883, 2013. DOI: <https://doi.org/10.1017/S1366728912000600>

VEIVO, O.; PORRETTA, V.; HYÖNÄ, J.; JÄRVIKIVI, J. Spoken Second Language Words Activate Native Language Orthographic Information in Late Second Language Learners. *Applied Psycholinguistics*, Cambridge, v. 39, n. 5, p. 1011-1032, 2018. DOI: <https://doi.org/10.1017/S0142716418000103>

VENTURA, P.; KOLINSKY, R.; BRITO-MENDES, C.; MORAIS, J. Mental Representations of the Syllable Internal Structure Are Influenced by Orthography. *Language and Cognitive Processes*, [S.l.], v. 16, n. 4, p. 393-418, 2001. DOI: <https://doi.org/10.1080/01690960042000184>

YOUNG-SCHOLTEN, M.; LANGER, M. Orthographic Input in Second Language German: Evidence from Naturalistic Adult Learners' Production. *Applied Psycholinguistics*, Cambridge, v. 36, n. 1, p. 93-114, 2015. DOI: <https://doi.org/10.1017/S0142716414000447>

ZIEGLER, J. C.; FERRAND, L. Orthography Shapes the Perception of Speech: The Consistency Effect in Auditory Word Recognition. *Psychonomic Bulletin & Review*, [S.l.], v. 5, p. 683-689, 1998. DOI: <https://doi.org/10.3758/BF03208845>

ZIEGLER, J. C.; FERRAND, L.; MONTANT, M. Visual Phonology: The Effects of Orthographic Consistency on Different Auditory Word Recognition Tasks. *Memory & Cognition*, [S.l.], v. 32, p. 732-741, 2004. DOI: <https://doi.org/10.3758/BF03195863>

ZIEGLER, J. C.; MUNEUX, M. Orthographic Facilitation and Phonological Inhibition in Spoken Word Recognition: A Developmental Study. *Psychonomic Bulletin & Review*, [S.l.], v. 14, p. 75-80, 2007. DOI: <https://doi.org/10.3758/BF03194031>

## APPENDICES

### APPENDIX A – Words that encompass the artificial lexicon

	Targets	Distractors
	/i/	/ɨ/
Control	geesh / keet / seeg	bup / nup/ sud
Experimental	deit / geib / meip geop / teog/ teob	doup / soug / toud dood/ pood / loob

### APPENDIX B – Participants’ general background information

Code	Sex	Age	Occupation	Proficiency	Age of acquisition	Other additional languages
1	F	19	English major	H	6	French, Spanish
2	F	20	Accounting major	I	5	Spanish
3	F	21	English major	H	16	
4	F	18	English major	H	6	
5	M	26	English major	H	13	Italian
6	M	18	English major	m	8	
7	M	32	MA candidate	I	10	
8	F	28	English major	H	22	
9	F	32	English major	I	13	Japanese, Korean, Spanish
10	M	28	English major	H	18	
11	F	25	MA Candidate	I	10	French, Italian
12	M	19	Engineering major	H	10	Italian
13	M	19	English major	H	15	
14	M	24	English major	H	18	French
15	F	46	PhD candidate	H	18	Spanish

16	F	20	English major	H	12	Spanish
17	F	40	MA candidate	H	16	German, Spanish
18	F	27	English major	H	15	
19	M	27	PhD candidate	H	8	French
20	M	27	English major	H	18	
21	F	29	English instructor	H	18	
22	M	22	MA candidate	H	11	
23	F	28	PhD candidate	H	14	
24	F	46	Japanese and Portuguese instructor	H	15	Japanese
25	F	25	English major	H	16	
26	F	23	English major	H	5	
27	M	22	English major	H	17	Spanish
28	F	29	MA candidate	H	15	German
29	F	18	English major	H	11	
30	F	21	English major	H	16	
31	F	26	Philosophy major	H	13	Spanish
32	M	30	PhD candidate	H	10	French
33	F	20	English major	H	17	
34	F	19	English major	H	15	
35	F	47	PhD candidate	H	23	Ukrainian
36	M	20	English major	H	8	Spanish