LANGUAGE, COGNITION AND NEUROSCIENCE
2022, VOL. 37, NO. 4, 457–468
https://doi.org/10.1080/23273798.2021.1985536

REGULAR ARTICLE

How are words with diacritical vowels represented in the mental lexicon? Evidence from Spanish and German

Manuel Perea[a], Melanie Labusch[a] and Ana Marcet[b]

[a]Departamento de Metodología and ERI-Lectura, Universitat de València, Valencia, Spain; [b]Departamento de Didáctica de la Lengua y la Literatura, Universitat de València, Valencia, Spain

ABSTRACT
Recent research has shown that the omission of diacritics in words does not affect the initial contact with the lexical entries, as measured by masked priming. In the present study, we directly examined whether diacritics' omission slows down lexical access using a single-presentation semantic categorisation task ("is the word an animal name?"). We did so in a language in which diacritics reflect lexical stress but not vowel quality (Spanish; e.g. ratón [mouse] vs. raton; Experiment 1) and in a language in which diacritics reflect vowel quality but not lexical stress (German; e.g. Kröte vs. Krote; Experiment 2). In Spanish, word response times were similar for words with diacritics that were either present or omitted. In contrast, German words were responded more slowly when the words' diacritics were omitted. Thus, the function of diacritics in each language determines how words with diacritics are represented in the mental lexicon.

ARTICLE HISTORY
Received 11 April 2021
Accepted 19 September 2021

KEYWORDS
Word recognition; lexical access; modelling; cross-language differences

Introduction

Except for English, the rest of the European languages using Latin script have diacritics (also called accent marks) added to some of the letters. Regarding vowel letters, which are the focus of the present paper, the role of diacritics differs widely across languages (see Wells, 2000). The diacritics on a vowel may: i) specify the stressed syllable (e.g. in Spanish, ámbar /ˈam.bar/ [amber]); ii) indicate a long vowel (e.g. in Czech, á→/aː/ but a→/a/); iii) indicate a hiatus (e.g. in Dutch, ruïne /ryˈinə/), or iv) represent a phoneme that was absent in Latin (e.g. in German, ä→/ɛ/ but a→/a/). Furthermore, in Romance languages, diacritics are also used to distinguish among otherwise homonyms (e.g. mí [me] vs. mi [my] in Spanish).

Somewhat surprisingly, none of the most influential neurally-inspired models of visual word recognition (e.g. Dehaene et al., 2005; Grainger et al., 2008) discuss how diacritical vowels are encoded and represented in the letter/word recognition system. Most computational models of visual word recognition (e.g. LTRS model: Adelman, 2011; spatial coding model: Davis, 2010; multiple read-out model: Grainger & Jacobs, 1996; CDP+ model: Perry et al., 2007) are also silent regarding the representation of accented vowels. The implementation of the above models only contains the 26 letters of English. Indeed, when simulating data from languages other than English, scholars usually remove the word's diacritics. For instance, Conrad et al. (2010) conducted simulations in Spanish and German using Grainger and Jacobs (1996) multiple read-out model and removed the accent marks from diacritical words (e.g. camión [truck] was encoded as camion). However, Conrad et al. (2010) did not provide a rationale for why it was adequate to remove the diacritics—note that their choice implicitly assumed that diacritical and non-diacritical vowels shared their letter representations. Notwithstanding, two computational models include diacritical vowels in the letter/word levels: 1) the multitrace memory model of reading (Ans et al., 1998) contains all 40 alphabetic characters in French (e.g. the characters e, é, è, ê, or ë would be treated as separate letter units); and 2) the connectionist model developed by Hutzler et al. (2004) contains the three diacritical vowels in German (i.e. ä, ö, and ü). However, neither Ans et al. (1998) nor Hutzler et al. (2004) provided a rationale for why diacritical and non-diacritical vowels should be treated as separate letters in the letter/word recognition system. [1] Thus, in previous research, words with diacritical vowels were simulated in two different ways, but neither of them provided evidence for their appropriateness. The present experiments were designed to fill this gap in the literature.

CONTACT Manuel Perea mperea@uv.es; mperea@valencia.edu Departamento de Metodología and ERI-Lectura, Av. Blasco Ibáñez, 21, Valencia 46010, Spain
The scripts and data of the two experiments are available at the following OSF link: https://osf.io/t8fnu/?view_only=06c7e54a578f4d269af3e485a490156f

Recent masked priming experiments have shown that non-accented vowels can initially activate their accented counterparts to a large degree. This evidence has been obtained in languages in which diacritics indicate lexical stress (e.g. Spanish; e.g. á→/a/ and a→/a/; ámbar /'am.bar/[amber] Perea et al., 2020b) and languages where diacritics indicate vowel quality (e.g. Finnish: ä→/æ/ but a→/ɑ/; Perea et al., 2021). In Spanish, Perea et al. (2020b) found that lexical decision times to a target word like FÁCIL (easy) were equivalent when preceded by the identity prime fácil (/' fa.θil/) and when the prime omitted the accent mark (facil /fa'θil/). These two priming conditions were more effective than the control condition fecil (/fe'θil/). The same pattern also occurs in lexical decision and naming experiments in Finnish: for a target word like TÄYTE (filling) the prime tayte (/'tɑyte/) was as effective as the identity prime täyte (/'tæyte/), and more effective than the control prime toyte (/'toyte/) (Perea et al., 2021). These findings suggest that a non-diacritical vowel (e.g. a) initially activates its accented counterparts (e.g. á, ä). Furthermore, this effect does not appear to be modulated by the role of the diacritics in the language (i.e. a similar pattern occurred in Spanish and Finnish), thus suggesting that the effect has a perceptual rather than a linguistic nature.

However, while providing valuable information, the findings from masked priming experiments present two interpretive difficulties. First, these effects could be due to perceptual noise regarding letter identity in the first moments of word processing (see Norris & Kinoshita, 2012). Keep in mind that non-accented vowels closely resemble their accented counterparts (i.e. a is very similar to á; see Simpson et al., 2013, for a visual letter similarity database). Indeed, Marcet et al. (2020) found that muneca-MUÑECA [doll] was responded as fast as muñeca-MUÑECA and faster than the control museca–MUÑECA – note that n and ñ are different letters in Spanish (see also Gutiérrez-Sigut et al., 2019; Marcet & Perea, 2017, 2018, for evidence of visual similarity effects with non-diacritical letters). Thus, the effectiveness of the masked prime facil for the target FÁCIL (Perea et al., 2020b) or the masked prime tayte for the target TÄYTE (Perea et al., 2021) are demonstrations of a visual similarity effect during the first moments of processing. Second, in masked priming, the research question is how a prime affects the first moments of processing a target word (i.e. the relationship between prime and target). Therefore, the masked priming technique does not directly test whether a prime with omitted diacritics (e.g. facil) activates the lexical-semantic representations of fácil (see Gómez et al., 2021, for discussion). To test whether a word with omitted diacritics such as facil produces as much activation at a lexical-semantic level as fácil, one would need to present these items in an unprimed paradigm.

The main goal of the present experiments was to examine whether the omission of a word's diacritics has a deleterious effect on lexical access when compared to its accented baseline. To that end, we selected a set of diacritical words (e.g. vagón [wagon]). These words were presented either with their diacritics or not (e.g. vagón vs. vagon). A second goal was to test whether diacritics' function in the language modulates the processing cost of omitting the word's diacritics. For that reason, we conducted parallel experiments in Spanish and German, two languages in which diacritics' role is very different. In Spanish, diacritics do not affect the grapheme-to-phoneme mapping but only designate the stressed syllable. In contrast, in German, diacritics only indicate vowel quality (ä /ɛ/, ö /ø/, ü /y/ vs. a /a/, o /o/, u /u/). While accented vowels in languages like Spanish are considered mere variants of the non-accented vowels, they are considered different letters in German. For instance, when children learn the alphabet, Spanish children only learn the non-diacritical vowels, whereas German children learn both the non-diacritical and the diacritical vowels. This dissociation can be seen in dictionaries (e.g. ä is a letter in German, but á is not a letter in Spanish) and computer keyboards (e.g. ä has a key in German keyboards, whereas á does not form part of Spanish keyboards).

Before introducing the experiments, we first review the very scarce literature examining this issue via unprimed experiments. In a go/no-go lexical decision experiment (i.e. respond to "words" but refrain to responding to "pseudowords"), Schwab (2015) tested the cost of omitting the diacritics in Spanish. The stimuli were presented in a block composed of diacritical items (e.g. the word cárcel [jail] or the pseudoword méstul) or in a block consisting of items where the diacritics were omitted (e.g. carcel, mestul). Participants were instructed not to pay attention to whether the stimulus required an accent mark. Schwab (2015) found similar response times for cárcel and carcel and concluded that, at least in Spanish, á and a should be considered as variants of the same abstract letter representation. However, the blocked presentations have one interpretive issue: accented vs. non-accented items were not compared directly in the same block. Furthermore, using a go/no-go lexical decision could alter some core processes relative to the standard yes/no lexical decision (see Vergara-Martínez et al., 2020, for electrophysiological evidence).

More recently, Marcet and Perea (2021) examined whether the pattern obtained by Schwab (2015) in

Spanish could be generalised to a more ecological scenario: silent sentence reading. They inserted a target word with diacritics (e.g. vagón [wagon]) in sentences that participants read silently while their eye movements were recorded. Critically, the target words were written with their diacritics present (vagón) or omitted (vagon). All other words in the sentence were presented with diacritics when needed (e.g. "Es mejor estar en el último vagón para salir antes que los demás." [It is better to be in the last car to leave before the others.]). Marcet and Perea (2021) found remarkably similar first-pass eye fixation times on items like vagón and vagon, measured by first-fixation durations and gaze durations (sum of first-pass fixations). They also found a small reading cost for the target words with the diacritics omitted (around 21 ms) in the total times (i.e. when including re-reading [second/third-pass fixations]). Specifically, participants re-read the target word more often when the accent mark was omitted. Marcet and Perea (2021) interpreted this latter finding as being due to post-access integration processing (e.g. participants could have re-checked the word's spelling with the omitted diacritic) – of note, the overall cost was less than 3% in sentences reading time.

Thus, Schwab (2015), Marcet and Perea (2021) found no (or minimal) processing cost due to omitting the diacritics during word recognition and reading. However, these experiments were conducted in a language, Spanish, where the diacritics do not modify the phonemic value of the vowels (e.g. á and a have the same phonological representation [/a/]). Indeed, many researchers assume that non-accented and accented vowels share the letter representations in Spanish (see Chetail & Boursain, 2019; Perea et al., 2020b; Schwab, 2015). Thus, the above findings may well be specific to those languages where diacritics do not affect the grapheme-to-phoneme mapping but only designate the stressed syllable. To reach broader conclusions, it is critical to run parallel experiments in a language where diacritics indicate vowel quality (e.g. German).

Thus, the present study examined whether omitting a word's diacritics slows down lexical access in Spanish (Experiment 1) and German (Experiment 2). We selected a set of words, all of them with diacritics. Each item was presented to the readers either with their diacritics present or omitted. All the selected words had an unambiguous spelling in the sense that omitting the diacritic did not form another word (e.g. the Spanish word revólver [gun] would not have been selected because revolver [to stir] is also a word). The participant's task was to decide whether the word was an animal name or not (i.e. a semantic categorisation task) (e.g. Spanish: non-animal vagón [wagon] vs. vagon; animal águila [eagle] vs. aguila; German: non-animal Porträt [portrait] vs. Portrat; animal Kröte [toad] vs. Krote). [2] Participants were told to ignore the diacritics' omission and were shown several trial examples for illustration before the experiment.

We preferred a semantic categorisation task over the lexical decision task for two reasons. First, the semantic categorisation task requires access to meaning (Forster & Shen, 1996), whereas lexical decision can be performed without unique word identification (see Grainger & Jacobs, 1996). Indeed, non-human species can perform reasonably well in lexical decision experiments (e.g. baboons: Grainger et al., 2012; pigeons: Scarf et al., 2016). Second, the omission of accent marks in German (e.g. Kröte→Krote) would produce a visually unfamiliar stimulus (i.e. Krote is not a word). As a result, participants might be biased to respond "nonword" to non-accented words in a lexical decision task. (Indeed, this was one reason to use blocked presentations in the Schwab, 2015, experiment in Spanish.) Perea et al. (2020a) found an effect of visual familiarity in lexical decision (e.g. hOuSe slower and more error-prone than HOUSE). This interpretive issue is minimised in semantic categorisation (hOuSe and HOUSE produce similar response times and error rates; Perea et al., 2020a).

The predictions are the following. First, let's assume diacritics are an intrinsic part of a word's orthographic representation regardless of language. In that case, we would expect slower word identification times when omitting the diacritics in *both* Spanish and German. This outcome would require reinterpreting the accounts that claim that accented, and non-accented vowels are only variants of the same abstract representation in Spanish (Chetail & Boursain, 2019; Perea et al., 2020b; Schwab, 2015). Furthermore, this outcome would require modellers to add diacritical letters in the letter level of future implementations of visual-word recognition models. Second, diacritics may be an integral part of a word's abstract representations but only when they indicate parameters such as vowel quality (i.e. when accented vowels map onto different phonological representations). In this case, we would find slower word identification times when omitting the diacritics in German but not in Spanish. This outcome would reinforce the view that each language's idiosyncrasies shape how letters and words are represented (see Frost, 2012, for discussion): modellers should include diacritical letters (e.g. ä, ö, and ü) in models of visual word recognition in languages where diacritics indicate vowel quality. And third, if diacritics – regardless of their function in language – are not an integral part of a

Table 1. Mean response times (in ms) and error rates (in percentages) for non-animal words and animal words written with their diacritics present or omitted in Experiment 1 (Spanish).

	Diacritic Present		Diacritic Omitted	
	Response time	Accuracy	Response time	Accuracy
Non-Animals	626	1.6	633	1.9
Animals	636	6.2	637	5.7

word's abstract representations, we would expect equivalent word identification times when the word's diacritics are present and when they are omitted in both Spanish and German. This outcome would suggest that diacritics' role in vowels during lexical access in Western-European languages is less important than previously thought – particularly in languages like German where they signal vowel quality.

Experiment 1 (Diacritics in Spanish)

Method

Participants

We tested 46 individuals (20 women; mean age = 24.1 years [SD = 6.13]), using Prolific Academic (http://prolific.ac). As we had 60 trials in each condition, this sample size ensured 2760 observations in each format (accent present vs. accent omitted), which is in line with Brysbaert and Stevens (2018) suggestions. Prolific Academic's recruitment filter was used such that only native Spanish university students with no reading problems and with normal or corrected-to-normal vision could participate. All participants gave informed consent before the experiment and received monetary compensation according to Prolific's participant policy. This research followed the Helsinki convention requirements, and we obtained ethical approval from the Research Ethics Committee of the University of Valencia.

Materials

We selected a set of 80 words with diacritics from the 120-word list of Marcet and Perea's (2021) sentence reading experiment. None of these 80 words were animal names or referred to concepts related to animal names or parts of body or plants (e.g. we excluded items like `fósil` [fossil], `árbol` [fossil], or `biología` [biology]). The diacritics could appear on the last syllable (e.g. `vagón` [wagon]), the last-but-one syllable (e.g. `cárcel` [jail]), or the third-from-last syllable (e.g. `lógica` [logic]). The mean number of letters was 6.5 (range: 5-10) and the mean word-frequency per million was 27.0 (range: 0.07-264.64) in the EsPal database (Duchon et al., 2013). We also selected 40 words with diacritics that were animal names (e.g. `ratón`). [3] The number of letters of animal names was matched with the number of letters of non-animal names (M = 6.8, range: 4-10; $p > .19$) and the mean Zipf frequency was 3.10 (range: 0.64-4.49). Each item (either animal name or not) was presented with its corresponding diacritic (e.g. `vagón`, `ratón`) or with the diacritic omitted (e.g. `vagon`, `raton`). We created two counterbalanced lists consisting of 60 words that kept the diacritics (40 non-animal names; 20 animal names) and 60 words in which the diacritic was omitted (60 non-animal names; 20 animal names). Those words in List 1 with diacritics were presented without diacritics in List 2 and vice versa. Participants were randomly assigned to List 1 or List 2. The list of words is presented in the Appendix.

Procedure

The experiment was created with Psychopy 3 software (Peirce & MacAskill, 2018) and was conducted in an online setting using the online server Pavlovia (www.pavlovia.org). Before starting the experiment, participants filled out a questionnaire with demographic data (age, gender, education level) on LimeSurvey (www.limesurvey.org). Participants were instructed to do the experiment in a quiet room without any distractions. Before starting the actual experiment, all participants went through sixteen practice trials to get familiarised with the task. The experiment consisted of a semantic categorisation task ("Does this word refer to an animal?"), in which participants had to press the button "M" on their keyboard for answering "yes" and the button "Z" on their keyboard for answering "no" as fast and accurately as possible. A trial started presenting a fixation cross in the centre of the screen for 500 ms. Afterward, the target item occurred in the same location until a response was made (or until a deadline of 2000ms). The trials were presented in a randomised order for each participant. Altogether, the experiment took about 10 min, including a short break every 60 trials.

Results and discussion

Incorrect responses (3.83%) and response times shorter than 250 ms (less than 0.01% [1 observation] were omitted from the response time (RT) analyses. The mean RT (in ms) and error rate (in percentage) in each experimental condition are displayed in Table 1.

We created Bayesian linear mixed-effects models for each dependent variable with the brms package (Bürkner, 2017) in the R environment (R Core Team, 2021). The two fixed factors in the model were Format (accented vs. non-accented; encoded as −0.5 and 0.5) and Type of word (non-animal vs. animal; encoded as −0.5 and 0.5). We chose the models with the maximal random effect structure (see Barr et al., 2013):

$$\text{RT [accuracy]} \sim \text{Format} * \text{TypeWord} + (1 + \text{Format} * \text{TypeWord}|\text{subject}) + (1 + \text{Format} \mid \text{item})$$

Due to the positive skew of latency data, we modelled them with the exgaussian distribution (`family = exgaussian()`). Given the binary nature of the accuracy data (correct [1], incorrect [0]), we modelled these data with the Bernoulli distribution (`family = bernoulli()`). For each model, we conducted 5,000 iterations (1,000 as a warm-up) with four chains. The models converged successfully (all $\hat{R}$s = 1.00). The output of Bayesian linear mixed-effects models offers the estimate of each parameter, its standard deviation, and the 95% Credible Intervals (95% CrI). If the 95% Credible Interval of a given parameter does not contain zero, it is interpreted as evidence of an effect (see Cutter et al., 2021, for a similar reasoning).

Response Time data. Our Bayesian linear mixed-effects model revealed clear effects of Type of word, indicating that participants responded more rapidly to the non-animal words than to the animal words, $b = -18.78$, $SE = 6.95$, 95% CrI (−32.51, −5.42). More important, responses to the words were remarkably similar when they kept their accent mark and when it was omitted, $b = -1.56$, $SE = 3.83$, 95% CrI (−9.11, 5.93). The 95% Credible Interval of this small effect (centered at −1.56 ms) contained zero, thus revealing a null/negligible effect of Format (accented vs. non-accented) in Spanish. There were no signs of an interaction between the two factors, $b = 4.04$, $SE = 4.74$, 95% CrI (−5.21, 13.27) (see Figure 1 for the posterior distributions).

Accuracy data. Participants committed fewer errors for trials involving non-animal words than for trials involving animal words, $b = 1.54$, $SE = 0.32$, 95% CrI (0.93, 2.20). Besides, error rates to words were very similar when the accent mark was present and when it was omitted, $b = 0.31$, $SE = 0.30$, 95% CrI (−0.24, 0.93). There were no signs of an interaction between the two factors, $b = -0.25$, $SE = 0.39$, 95% CrI (−1.01, 0.52) (see Figure 1 for the posterior distributions).

The present data in Spanish showed remarkably similar response times and accuracy for those words that kept the accent mark (e.g. `ratón`, `vagón`) and those words in which the accent mark was omitted (e.g. `raton`, `vagon`). This pattern is consistent with the null effects reported by Schwab (2015) with a go/no-go lexical decision task and by Marcet and Perea (2021) in the first-pass fixations on the target words during sentence reading.

To complement the above linear-mixed effects models with a quantile-based analysis, we created the delta plots (see Ridderinkhof, 2002) for the effect of Format. Specifically, we plotted the difference in response times between the words with diacritics present and the words with diacritics omitted for the .1, .3, .5, .7, and .9 quantiles in each participant along the y-axis, whereas the x-axis indicated the means per quantile. As shown in Figure 2, there were essentially no differences between those words with the accent present and those words with the accent omitted. The only slight divergence was in the slower responses (i.e..9 quantile), where the words with the omitted accent mark yield slower response times than the words with the accent mark. An ad hoc analysis in the .9 quantile showed a 18-ms cost, $t(45) = 2.29$, $p = 0.026$. While one should be cautious of ad hoc analyses, this difference in the higher quantiles aligns well with the view that diacritics in Spanish may play some role during a post-access integration stage (see Marcet & Perea, 2021, for evidence during sentence reading).

The question now is whether this same pattern holds when the diacritics represent vowel quality, as occurs in German. This was the main aim of Experiment 2.

Experiment 2 (Diacritics in German)

Methods

Participants

We tested a sample of 50 participants (23 women; mean age = 25.7 years [SD = 4.44]) using Prolific Academic for participant recruitment with the same profile as in Experiment 1. The only difference in the filters was that the native language should be German instead of Spanish. This sample size ensured 2,700 observations in each format (i.e. parallel to that of Experiment 1). All participants gave informed consent before the experiment and received monetary compensation for their participation.

Materials

For the non-animal names, we selected a set of 72 German words with a diacritical vowel (e.g. `Porträt`). The mean number of letters was 7.4 (range: 4-14), and the mean word-frequency per million was 2.9 (range: 0.04-16.69) in the German database created by Oganian et al. (2016) (We did not select any items that could be words with and without diacritics (e.g. `Bar` [bar] vs. `Bär` [bear]). For the animal names, we selected 36 German words with a diacritical vowel (e.g. `Kröte`).[4]

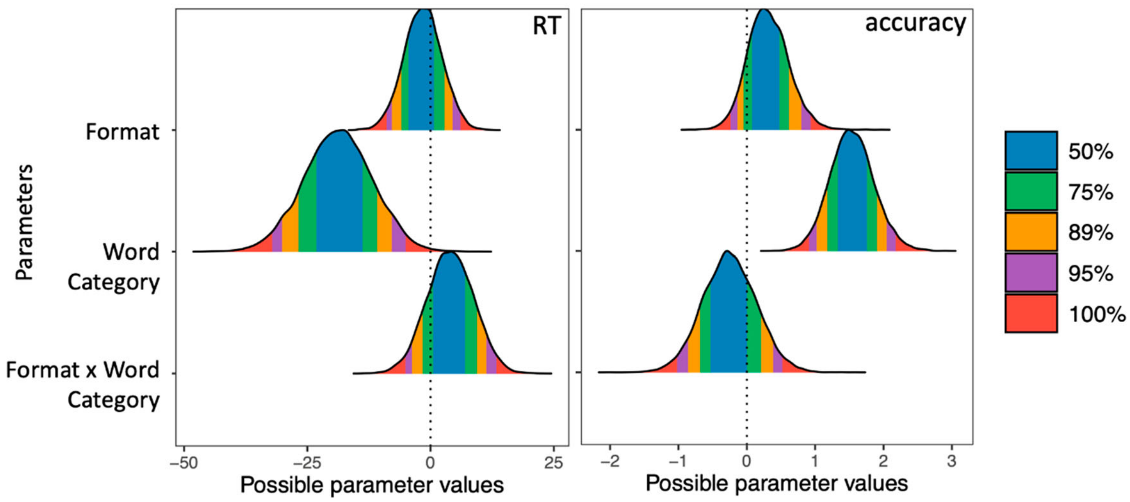

Figure 1. Highest Density Intervals with the 50%, 75%, 89%, 95%, and 100% Credible Intervals for each of the estimates of the Bayesian Linear Mixed-Effects models on response time (left panel) and accuracy (right panel) for Spanish words with and without diacritics (Experiment 1).

The average number of letters was also 7.4 (range: 4-14), and the mean word-frequency per million was 2.9 (range: 0.04-20.16). The items could be presented with their diacritical vowel (e.g. `Porträt`, `Kröte`) or not (e.g. `Portrat`, `Krote`). For both types of words, we used approximately the same ratio of plural words and the same ratio of diacritical vowels (`ä`, `ö`, `ü`). As in Experiment 1, we created two counterbalanced lists in a Latin Square manner, each composed 54 intact diacritical words (36 non-animal names; 18 animal names) and 54 words in which the diacritical vowel was omitted (36 non-animal names; 18 animal names). The list of words is presented in the Appendix.

Procedure

It was the same as in Experiment 1, except that the task was done in German instead of Spanish. Furthermore, due to the different architecture of German keyboards, for a "no" response, participants pressed the "X" button on their keyboard instead of the "Z" button.

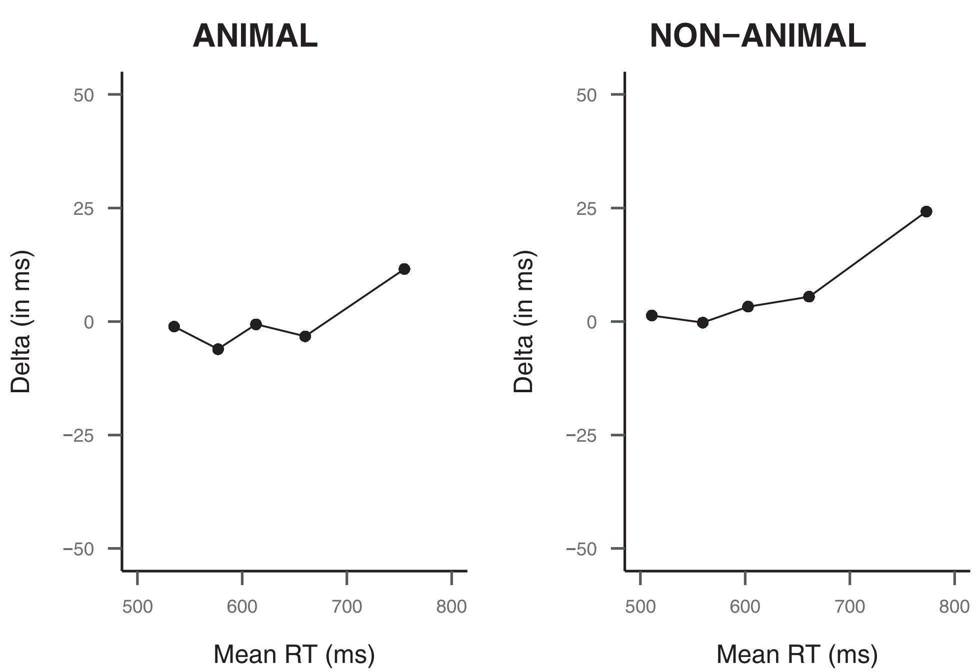

Figure 2. Delta Plots showing the difference between Spanish words with and without diacritics for the .1, .3,.5, .7, and .9 quantiles.

Results and discussion

As in the previous experiment, we removed error responses (4.2%) and very short responses (<250 ms; 0 observations) from the RT analyses. Table 2 displays the average RTs and error rates in each condition. The inferential analyses were parallel to those of Experiment 1 – again, all models converged adequately ($\hat{R} = 1.00$ in all estimates).

Response Time data. Responses were faster when the words kept their diacritic than when it was omitted, $b = 16.16$, $SE = 3.76$, 95% CrI (8.76, 23.52). In this case, the 95% Credible Interval of this parameter did not contain zero, thus revealing an effect of Format (accented vs. non-accented) in German. There were no signs of a difference in response times between animal and non-animal words, $b = -1.47$, $SE = 7.27$, 95% CrI (−15.94, 12.80) or an interaction between the two factors, $b = -3.94$, $SE = 5.04$, 95% CrI (−13.94, 5.81) (see Figure 3 for the posterior distributions).

Accuracy data. As occurred in Experiment 2, participants made fewer errors to non-animal words than to animal words, $b = 1.04$, $SE = 0.33$, 95% CrI (0.40, 1.71). In addition, accuracy was remarkably similar for words with the accent present vs. omitted, $b = 0.03$, $SE = 0.26$, 95% CrI (−0.47, 0.54). The two factors did not interact, $b = 0.36$, $SE = 0.39$, 95% CrI (−0.38, 1.13) (see Figure 3 for the posterior distributions).

The present experiment has shown that, unlike in Spanish (Experiment 1), the omission of the diacritics in German words entails a processing cost: word identification times were shorter to `Porträt` [portrait] or `Kröte` [toad] than to `Portrat` or `Krote`. Importantly, this processing cost did not entail a lower accuracy: readers could successfully encode `Portrat` as a non-animal and `Krote` as an animal despite that, in German, the pairs ä/a, ö/o, and ü/u correspond to different letters (each with a different sound).

To complement these findings, we computed a delta plot of the effect of format for animal and non-animal words in a parallel manner as in Experiment 1 (Spanish). As shown in Figure 4, the delta plot shows a robust effect of format even in the distribution's leading edge (.1 quantile). This effect increased progressively in successive quantiles – this pattern is consistent with the diacritics' influence on central word identification processes (see Gomez, 2012).

Table 2. Mean response times (in ms) and error rates (in percentages) for non-animal words and animal words written with their diacritics present or omitted in Experiment 2 (German).

	Diacritic Present		Diacritic Omitted	
	Response time	Accuracy	Response time	Accuracy
Non-Animals	615	2.6	648	2.1
Animals	608	6.1	636	6.0

Combined analyses of experiments 1 and 2

The findings of Experiments 1 and 2 strongly suggest that the function of diacritics in each language determines how diacritical words are represented in the mental lexicon. As suggested by a Reviewer, to empirically support this claim, it may be important to run, for the RT data, a combined analysis of Experiments 1 and 2 (i.e. including Language as a factor) and show an interaction between Format (accented vs. non-accented) and Language (Spanish vs. German).

The combined analyses of Experiments 1 and 2 were parallel to the individual experiments except that we added Language (Spanish [Experiment 1] vs. German [Experiment 2], encoded as −0.5 and 0.5 respectively) as a fixed factor. Thus, the resulting maximal factor-structure model was:

$$\begin{aligned} RT \sim\ & \text{Format} * \text{TypeWord} * \text{Language} \\ & + (1 + \text{Format} * \text{TypeWord} * \text{Language} \mid \text{subject}) \\ & + (1 + \text{Format} \mid \text{item}) \end{aligned}$$

As in the individual analyses, we conducted 5,000 iterations in the model (1,000 as sampling) and the value of $\hat{R}$ was 1.00 for all estimates. As expected, the model provided evidence favouring an interaction between Format and Language, $b = 18.35$, $SE = 9.14$, 95% CrI (0.55, 36.37) – note that this interaction occurred similarly for animal and non-animal names: three-way interaction, $b = -8.87$, $SE = 11.37$, 95% CrI (−31.11, 13.10). (The full output is presented in the OSF.) Thus, the present analysis has revealed that the effect of Format in accented and non-accented words is different for the two languages, Spanish (where the effect of Format is negligible; see Figures 1 and 2) and German (where the effect of Format is present; see Figures 3 and 4).

General discussion

We conducted two semantic categorisation experiments that tested whether omitting the diacritics in a word delays lexical access in Spanish and German. While both Spanish and German have transparent orthographies with a variable stress assignment, the diacritics' role in vowels is very different. Diacritical vowels in Spanish indicate the stressed syllable with no changes in vowel quality – note that Spanish keeps the five vowel sounds of Latin (i.e. `a` /a/, `e` /e/, `i` /i/, `o` /o/, and `u` /u/). In contrast, diacritical vowels in German do not

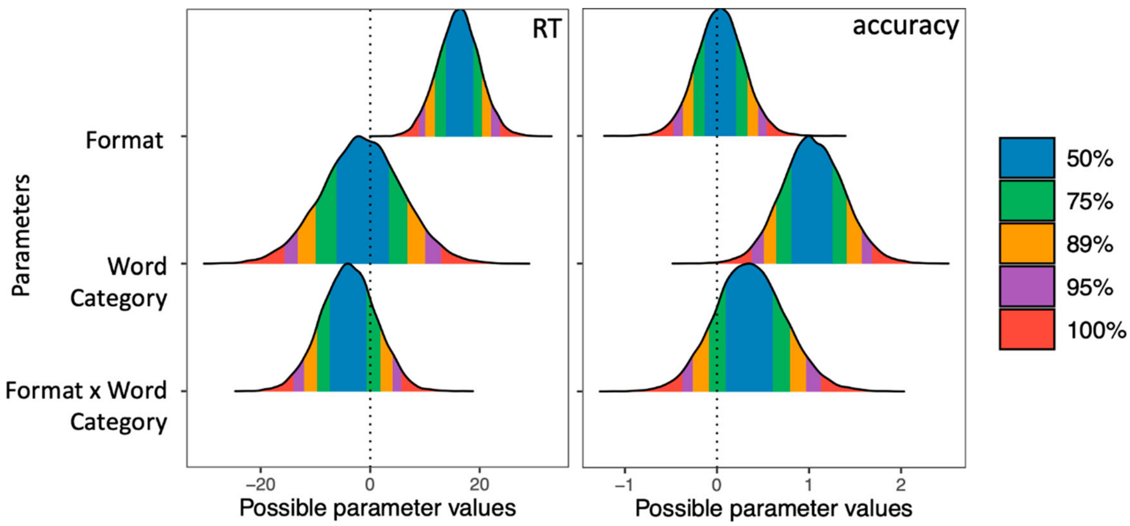

Figure 3. Highest Density Intervals with the 50%, 75%, 89%, 95%, and 100% Credible Intervals for each of the estimates of the Bayesian Linear Mixed-Effects models on response time (left panel) and accuracy (right panel) for German words with and without diacritics (Experiment 2).

have a prosodic function; instead, they specify the three vowel sounds of German that did not exist in the Latin alphabet (i.e. ä /ɛ/, ö /ø/, ü /y/). Results were straightforward: In Spanish, omitting a vowel's diacritics did not hamper (or only minimally) lexical access (Experiment 1), whereas, in German, the omission of a vowel's diacritics delays lexical access (Experiment 2). We now examine the theoretical and applied implications of these findings.

First, regarding Spanish, the similar word identification times for items like `vagón` and `vagon` in a task requiring access to meaning extends recent research with other tasks (lexical decision: Schwab, 2015; sentence reading: Marcet & Perea, 2021). That is, the omission of diacritics in a word's vowel has a minimal impact in (silent) reading tasks. These findings favour the idea that accented vowels in Spanish should be considered variants of their non-accented counterparts at the letter level in visual word recognition models (see Chetail & Boursain, 2019; Perea et al., 2020b; Schwab, 2015). Thus, current word recognition models (i.e. using the English orthography) can be employed to simulate experimental data

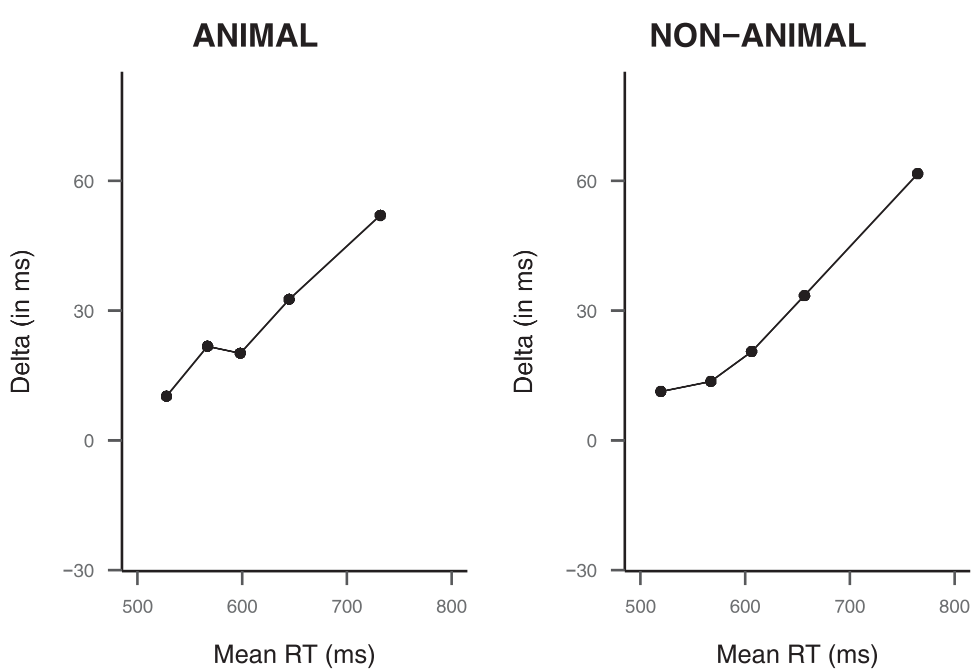

Figure 4. Delta Plots showing the difference between German words with and without diacritics for the .1, .3, .5, .7, and .9 quantiles.

from Spanish stimuli. We acknowledge, however, that some abstract orthographic markers of the accented vowels could be added to these models – this would be in the spirit of Peressotti et al. (2003) orthographic cue hypothesis. These markers for diacritical vowels could help the interplay between orthography and phonology at the whole-word level under some circumstances, especially for developing readers (or for learners of Spanish as L2). To tackle this issue, future research should examine in detail the role of diacritical marks in Spanish comparing the findings between developing vs. adult readers using tasks that require an explicit grapheme-to-phoneme mapping (e.g. naming task) and tasks that only require access to lexico-semantic representations (e.g. semantic categorisation task).

At an applied level, our findings fit well with the idea that diacritics in Spanish are an unnecessary element during reading – with the possible exception of learners of Spanish or of novel words. After all, nearly 80% of Spanish words have their stress of the last-but-one syllable, and only less than 3% of words have their stress in the second-to-half syllable or earlier (Quilis, 1993). Indeed, a repeated claim from Spanish scholars (including winners of the Nobel Prize for literature such as García Márquez) is that the rules of accentuation in Spanish should be more straightforward. The norms of accentuation in Spanish, which the Real Academia Española regulates, are rather complex (and with many exceptions). A noticeable trend is that the number of diacritical words in Spanish has decreased in the progressive reforms since the creation of the Real Academia in 1715. As an example, the initial spelling of the Real Academia Española was "Real Académia Españóla" – currently, less than 13% of the 500 most frequent words in Spanish require accent marks (Real Academia Española, 2010). Indeed, Romance languages such as Italian only have a very sparse use of diacritics (see Colombo & Sulpizio, 2021): when the stress is on the final syllable ending in a vowel (e.g. `libertà` [freedom]), and to distinguish potential homographs. Similarly, for a short period (1981-2003), the Real Acadèmia de Cultura of the Valencian language (i.e. the variety of Catalan spoken in the Valencian Community in Spain) only recommended the use of diacritics to distinguish otherwise homographs (e.g. `déu` [god] vs. `deu` [ten]) (Calpe, 2004). We believe that the present findings in Spanish, together with Schwab (2015) and Marcet and Perea (2021), offer relevant data so that the academies responsible for dictating the norms of accentuation can make informed decisions to simplify them.

Second, the scenario is quite different for German words. In this case, we found a reading cost of around 31 ms when omitting the vowel's diacritic on lexical access (e.g. `Portrat` was responded more slowly than `Porträt`). Notably, this reading cost occurred in the latency data but not in the accuracy data. That is, participants could effectively reconstruct the meaning of each item in German. We believe that this pattern occurs because non-diacritical vowels can initially activate their diacritical counterparts (e.g. `a→ä`) even when they refer to different phonemes (via visual similarity effects; see Perea et al., 2020c, for evidence with masked priming in Finnish). These findings strongly suggest that, in German, diacritical vowels (e.g. `a` and `ä`) should be treated as different letter units than their non-diacritical counterparts. Indeed, as stated in the Introduction, Hutzler et al. (2004) added the three diacritical vowels as different units when simulating an experiment in German. In the future, modellers of visual word recognition using German words should take the lead of Hutzler et al. (2004) and add the vowels `ä`, `ö`, and `ü` to the letter level. Another modification when modelling visual word recognition in German is to consider the initial capitalisation of common nouns (see Jacobs et al., 2008; Wimmer et al., 2016).

In sum, we have found a strong dissociation of the diacritics' role in vowel letters between Spanish and German in two semantic categorisation experiments. The omission of diacritics does not affect (or only minimally) lexical access to Spanish words (e.g. `vagón = vagon`), whereas it produces a sizeable reading cost in German (e.g. `Porträt < Portrat`). Thus, while `a` vs. `á` can be parsimoniously treated as variants of the same letter in Spanish, `ä` and `a` should be considered separate letter units in German. Further research is necessary to determine the diacritic's role in languages where only a subset of accented vowels indicate vowel quality (e.g. French: `à` and `a` have the same pronunciation, but `è` and `é` have different pronunciations; see Chetail & Boursain, 2019; Massol & Grainger, 2021) or vowel length (e.g. Czech). At a theoretical level, our findings offer cues on how to model the letter level in current computational modelling software designed to simulate visual word recognition experiments in Spanish and German (e.g. easyNet: Adelman et al., 2018). At the practical level, our findings in Spanish suggest that accent marks do not play an essential role during lexical access, and the Real Academia should consider further decreasing their use.

Notes

1. In a recent study on the role of repeated letters using the lexical decision data of French Lexicon Project, Trifonova and Adelman (2019) used two scenarios

depending on whether the diacritics on the vowels were disregarded or not.
2. The initial letter was capitalized in the German experiment because all words were nouns.
3. We could only select 40 familiar animal words in Spanish containing diacritics. Thus, the ratio of animal/non-animal responses was 2:1 (see Labusch et al., 2021; Perea et al., 2020a, for a similar ratio).
4. We could only find 36 familiar diacritical words in German that corresponded to animal names. As in Experiment 1, the ratio of animal/non-animal responses was 2:1.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Spanish Ministry of Science and Innovation [grant PSI2017-862120-P] and the Department of Innovation, Universities, Science and Digital Society of the Generalitat Valenciana [grant GV/2020/074].

References

Adelman, J. S. (2011). Letters in time and retinotopic space. *Psychological Review*, *118*(4), 570–582. https://doi.org/10.1037/a0024811

Adelman, J. S., Gubian, M., & Davis, C. J. (2018). *easyNet: A computational modeling software package for cognitive science, bridging the gap between novices and experts* [Computer software]. http://adelmanlab.org/easyNet/

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review*, *105*(4), 678–723. https://doi.org/10.1037/0033-295x.105.4.678-723

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1)), https://doi.org/10.5334/joc.10

Bürkner, P. C. (2017). Brms: An R package for Bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Calpe, À. V. (2004). *La producció editorial en les normes de la Real Acadèmia de Cultura Valenciana, 1979-2004. Una aproximació bibliomètrica* [The publishing production with the norms of the Real Academy of Valencian Culture, 1979-2004. A bibliometric approach]. Associació d'Escritors en Llengua Valenciana.

Chetail, F., & Boursain, E. (2019). Shared or separated representations for letters with diacritics? *Psychonomic Bulletin & Review*, *26*(1), 347–352. https://doi.org/10.3758/s13423-018-1503-0

Colombo, L., & Sulpizio, S. (2021). The role of orthographic cues to stress in Italian visual word recognition. *Quarterly Journal of Experimental Psychology*, https://doi.org/10.1177/17470218211006062

Conrad, M., Tamm, S., Carreiras, M., & Jacobs, A. M. (2010). Simulating syllable frequency effects within an interactive activation framework. *European Journal of Cognitive Psychology*, *22*(5), 861–893. https://doi.org/10.1080/09541440903356777

Cutter, M. G., Paterson, K. B., & Filik, R. (2021). Online representations of non-canonical sentences are more than good-enough. *Quarterly Journal of Experimental Psychology*, https://doi.org/10.1177/17470218211032043

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*(3), 713–758. https://doi.org/10.1037/a0019738

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences*, *9*(7), 335–341. https://doi.org/10.1016/j.tics.2005.05.004

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). Espal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, *45*(4), 1246–1258. https://doi.org/10.3758/s13428-013-0326-1

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(3), 696–713. https://doi.org/10.1037/0278-7393.22.3.696

Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences*, *35*(5), 263–279. https://doi.org/10.1017/s0140525x11001841

Gomez, P. (2012). Mathematical models of the lexical decision task. In J. Adelman (Ed.), *Visual word recognition (volume 1): models and methods, orthography and phonology* (pp. 70–89). Psychology Press.

Gómez, P., Marcet, A., & Perea, M. (2021). Are better young readers more likely to confuse their mother with their mohter? *Quarterly Journal of Experimental Psychology*.

Grainger, J., Dufau, S., Montant, M., Ziegler, J. C., & Fagot, J. (2012). Orthographic processing in baboons (papio papio). *Science*, *336*(6078), 245–248. https://doi.org/10.1126/science.1218152

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, *103*(3), 518–565. https://doi.org/10.1037/0033-295x.103.3.518

Grainger, J., Rey, A., & Dufau, S. (2008). Letter perception: From pixels to pandemonium. *Trends in Cognitive Sciences*, *12*(10), 381–387. https://doi.org/10.1016/j.tics.2008.06.006

Gutiérrez-Sigut, E., Marcet, A., & Perea, M. (2019). Tracking the time course of letter visual-similarity effects during word recognition: A masked priming ERP investigation. *Cognitive, Affective, and Behavioral Neuroscience*, *19*(4), 966–984. https://doi.org/10.3758/s13415-019-00696-1

Hutzler, F., Ziegler, J. C., Perry, C., Wimmer, H., & Zorzi, M. (2004). Do current connectionist learning models account for reading development in different languages? *Cognition*, *91*(3), 273–296. https://doi.org/10.1016/j.cognition.2003.09.006

Jacobs, A. M., Nuerk, H. C., Graf, R., Braun, M., & Nazir, T. A. (2008). The initial capitalization superiority effect in German: Evidence for a perceptual frequency variant of the orthographic cue hypothesis of visual word recognition.

Psychological Research, *72*(6), 657–665. https://doi.org/10.1007/s00426-008-0168-0
Labusch, M., Kotz, S. A., & Perea, M. (2021). The impact of capitalized German words on lexical access. *Psychological Research*, https://doi.org/10.1007/s00426-021-01540-3
Marcet, A., Ghukasyan, H., Fernández-López, M., & Perea, M. (2020). Jalapeno or Jalapeño: Do diacritics in consonant letters modulate visual similarity effects during word recognition? *Applied Psycholinguistics*, *41*(3), 579–593. https://doi.org/10.1017/S0142716420000090
Marcet, A., & Perea, M. (2017). Is nevtral NEUTRAL? Visual similarity effects in the early phases of written-word recognition. *Psychonomic Bulletin and Review*, *24*(4), 1180–1185. https://doi.org/10.3758/s13423-016-1180-9
Marcet, A., & Perea, M. (2018). Visual letter similarity effects during sentence reading: Evidence from the boundary technique. *Acta Psychologica*, *190*, 142–149. https://doi.org/10.1016/j.actpsy.2018.08.007
Marcet, A., & Perea, M. (2021). Does omitting the accent mark in a word affect sentence reading? Evidence from Spanish. *Quarterly Journal of Experimental Psychology*, https://doi.org/10.1177/17470218211044694
Massol, S., & Grainger, J. (2021). The sentence superiority effect in young readers. *Developmental Science*, *24*(2), https://doi.org/10.1111/desc.13033
Norris, D., & Kinoshita, S. (2012). Reading through a noisy channel: Why there's nothing special about the perception of orthography. *Psychological Review*, *119*(3), 517–545. https://doi.org/10.1037/a0028450
Oganian, Y., Conrad, M., Aryani, A., Heekeren, H. R., & Spalek, K. (2016). Interplay of bigram frequency and orthographic neighborhood statistics in language membership decision. *Bilingualism: Language and Cognition*, *19*(3), 578–596. https://doi.org/10.1017/s1366728915000292
Peirce, J. W., & Macaskill, M. R. (2018). *Building experiments in PsychoPy*. Sage.
Perea, M., Fernández-López, M., & Marcet, A. (2020a). Does CaSe-MiXinG disrupt the access to lexico-semantic information? *Psychological Research*, *84*(4), 981–989. https://doi.org/10.1007/s00426-018-1111-7
Perea, M., Fernández-López, M., & Marcet, A. (2020b). What is the letter é? *Scientific Studies of Reading*, *24*(5), 434–443. https://doi.org/10.1080/10888438.2019.1689570
Perea, M., Hyönä, J., & Marcet, A. (2021). Does vowel harmony affect visual word recognition? Evidence from Finnish. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0000907
Peressotti, F., Cubelli, R., & Job, R. (2003). On recognizing proper names: The orthographic cue hypothesis. *Cognitive Psychology*, *47*(1), 87–116. https://doi.org/10.1016/s0010-0285(03)00004-5
Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, *114*(2), 273–315. https://doi.org/10.1037/0033-295x.114.2.273
Quilis, A. (1993). *Tratado de fonología y fonética españolas*. Gredos.
R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
Real Academia Española. (2010). *Ortografía de la lengua española*. Espasa.
Ridderinkhof, K. R. (2002). Activation and suppression in conflict tasks: Empirical clarification through distributional analyses. In W. Prinz, & B. Hommel (Eds.), *Common mechanisms in perception and action. Attention and performance XIX* (pp. 494–519). Oxford University Press.
Scarf, D., Boy, K., Uber Reinert, A., Devine, J., Güntürkün, O., & Colombo, M. (2016). Orthographic processing in pigeons (Columba livia). *Proceedings of the National Academy of Sciences*, *113*(40), 11272–11276. https://doi.org/10.1073/pnas.1607870113
Schwab, S. (2015). Accent mark and visual word recognition in Spanish. *Loquens*, *2*(1), e018. https://doi.org/10.3989/loquens.2015.018
Simpson, I. C., Mousikou, P., Montoya, J. M., & Defior, S. (2013). A letter visual-similarity matrix for latin-based alphabets. *Behavior Research Methods*, *45*(2), 431–439. https://doi.org/10.3758/s13428-012-0271-4
Trifonova, I. V., & Adelman, J. S. (2019). A delay in processing for repeated letters: Evidence from megastudies. *Cognition*, *189*, 227–241. https://doi.org/10.1016/j.cognition.2019.04.005
Vergara-Martínez, M., Gomez, P., & Perea, M. (2020). Should I stay or should I go? An ERP analysis of two-choice versus go/no-go response procedures in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(11), 2034–2048. https://doi.org/10.1037/xlm0000942
Wells, J. C. (2000). Orthographic diacritics and multilingual computing. *Language Problems and Language Planning*, *24*(3), 249–272. https://doi.org/10.1075/lplp.24.3.04wel
Wimmer, H., Ludersdorfer, P., Richlan, F., & Kronbichler, M. (2016). Visual experience shapes orthographic representations in the visual word form area. *Psychological Science*, *27*(9), 1240–1248. https://doi.org/10.1177/0956797616657319

Appendix. Materials of the experiments

Experiment 1 (Spanish)

Non-animals: lápiz, dólar, volátil, difícil, seísmo, geología, psicología, ortografía, automóvil, egoísta, frágil, túnel, ecología, ámbar, bahía, caída, mástil, cárcel, tráiler, astronomía, filosofía, energía, poesía, álbum, decaído, mártir, móvil, versátil, flúor, cráter, mármol, encía, alegría, distraído, débil, hábil, portátil, cáliz, fácil, azúcar, éxito, látigo, vértigo, típico, brújula, cámara, sílaba, música, común, solución, pirámide, sábado, misión, película, vagón, cólera, vocación, formación, ábaco, cómico, cínico, ración, lógica, canción, sillín, ética, nítido, acordeón, física, cajón, rincón, almacén, buzón, túnica, plástico, champú, cerámica, único, meditación, bisturí

Animals: atún, alacrán, milpiés, antílope, chimpancé, camaleón, tórtola, león, búho, halcón, mejillón, búfalo, luciérnaga, orangután, salmón, colibrí, libélula, hurón, tarántula, víbora, lémur, tucán, delfín, pitón, jabalí, pelícano, hipopótamo, caimán, murciélago, ciempiés, boquerón, escorpión, faisán, pájaro, hámster, tiburón, águila, cacatúa, gorrión, ratón

Experiment 2 (German)

Non-animals: Bürgersteig, Bücherei, Heizöl, Behörde, Fußbälle, Käsebrot, Äste, Bäckerin, Töpferei, Töne, Paläste, Körbe, Küsten, Knöpfe, Wände, Flügel, Eisverkäufer, Köstlichkeiten, Geländer, Tonhöhe, Gemüse, Käfige, Obstbäume, Münze, Möbelhaus, Anträge, Jubiläum, Gewürz, Hutgröße, Getränk, Zäune, Südpol, Molekül, Baulärm, Glühwein, Erdnüsse, Kostüme, Eintöpfe, Löffel, Umzüge, Gemälde, Diät, Zahnärzte, Reiseführer, Perücken, Fertighäuser, Strände, Porträt, Möhren, Röcke, Knödel, Absätze, Schneestürme, Bahnhöfe, Sägemehl, Säcke, Speiseöl, Süßigkeit, Bustür, Hügel, Fremdkörper, Hochzeitspläne, Gefäß, Dünen, Hörbuch, Frisör, Säulen, Hütten, Windmühlen, Räume, Unglück, Gästeklo

Animals: Siebenschläfer, Mäuse, Löwe, Känguru, Lämmer, Walrösser, Schwäne, Regenwürmer, Läuse, Maulwürfe, Maikäfer, Hündin, Schildkröte, Kröte, Kühe, Hyäne, Wühlmaus, Chamäleon, Seelöwe, Eichhörnchen, Büffel, Gänse, Frösche, Hühner, Marienkäfer, Küken, Gürteltier, Möwe, Nashörner, Erdmännchen, Fledermäuse, Kälber, Füchse, Krähe, Wölfe, Würmer