A striking element for native speakers of English is the abundance of diacritics in other Latin-based orthographies (e.g., French: “Éléanor a reçu un diplôme de l’école la plus chère de Genève.” [Eleanor received an award from the most expensive school in Geneva.]). These diacritics were introduced several centuries ago to adapt the letters from the Latin alphabet to the nuances of each language, including, among other functions, those phonemes that did not exist in Latin (e.g., ñ/ in Spanish). A less frequent option is to use a letter from another alphabet (e.g., ch/ in English). Despite the presence of diacritical letters in nearly all Latin-based orthographies, current neurally-inspired accounts of visual-word recognition (e.g., Dehaene et al., 2005; Grainger, 2008), being focused on the representation of diacritical vowels: Do diacritical and nondiacritical vowels share their abstract letter representations? Recent research suggests that the answer is “yes” in languages where diacritics indicate suprasegmental information (e.g., lexical stress, as in cámara [‘ka.máɾa] camera; Spanish), but “no” in languages where diacritics indicate segmental information such as a different phoneme (e.g., the German vowels ò/ and ò/). Here we examined this issue in French, a language that contains a complex set of diacritical vowels (e.g., for the letter: è, ë, ë, and ë). In Experiment 1, using a semantic categorization task, we compared the word identification times to intact diacritical words (e.g., chèvre, goat in English) with a condition with omitted diacritics (chev). Results showed that the two conditions behaved similarly. In Experiments 2–4, we compared the intact diacritical words with a condition containing a mismatching diacritic, either existing in French (e.g., chèvre, chèvre) or not (the macron sign, as in chèvre). We only found a reading cost when replacing the diacritic with an existing one. In Experiments 5–6, we compared the semantic categorization times to intact nondiacritical words (e.g., cheval, horse in English) versus a condition with an added diacritic, either existing (chéval) or not (chéval). We found a reading cost for the words with the added diacritical mark in both cases. We discuss how models of visual-word recognition can be modified to represent diacritical vowels.

Keywords: cross-language differences, diacritics, lexical-semantic access, word recognition

This research was supported by Grants from the Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital, Generalitat Valenciana (CIAICO/2021/172 to Manuel Perea; GV/2020/074 to Ana Marcet) and from the Ministerio de Ciencia, Innovación y Universidades (PID2020-116740GB-I00 funded by the MCIN/AEI/10.13039/501100011033) to Manuel Perea.

Correspondence concerning this article should be addressed to Manuel Perea, Departamento de Metodología and ERI-Lectura, Universitat de València, Avenida Blasco Ibáñez, 21, 46010-Valencia, Spain. Email: mperea@uv.es

This article was published Online First December 22, 2022.
Manuel Perea  https://orcid.org/0000-0002-3291-1365
The stimuli, data, scripts, and output of the experiments are available at https://osf.io/9kup/.

An often overlooked but fundamental issue for any comprehensive model of visual-word recognition is whether diacritical and non-diacritical vowels share their abstract letter representations? Recent research suggests that the answer is “yes” in languages where diacritics indicate segmental information (e.g., lexical stress, as in cámara [‘ka.máɾa] camera; Spanish), but “no” in languages where diacritics indicate segmental information such as a different phoneme (e.g., the German vowels ö/ and ò/). Here we examined this issue in French, a language that contains a complex set of diacritical vowels (e.g., for the letter: è, ë, ë, and ë). In Experiment 1, using a semantic categorization task, we compared the word identification times to intact diacritical words (e.g., chèvre, goat in English) with a condition with omitted diacritics (chev). Results showed that the two conditions behaved similarly. In Experiments 2–4, we compared the intact diacritical words with a condition containing a mismatching diacritic, either existing in French (e.g., chèvre, chèvre) or not (the macron sign, as in chèvre). We only found a reading cost when replacing the diacritic with an existing one. In Experiments 5–6, we compared the semantic categorization times to intact nondiacritical words (e.g., cheval, horse in English) versus a condition with an added diacritic, either existing (chéval) or not (chéval). We found a reading cost for the words with the added diacritical mark in both cases. We discuss how models of visual-word recognition can be modified to represent diacritical vowels.

Keywords: cross-language differences, diacritics, lexical-semantic access, word recognition

This research was supported by Grants from the Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital, Generalitat Valenciana (CIAICO/2021/172 to Manuel Perea; GV/2020/074 to Ana Marcet) and from the Ministerio de Ciencia, Innovación y Universidades (PID2020-116740GB-I00 funded by the MCIN/AEI/10.13039/501100011033) to Manuel Perea.

Correspondence concerning this article should be addressed to Manuel Perea, Departamento de Metodología and ERI-Lectura, Universitat de València, Avenida Blasco Ibáñez, 21, 46010-Valencia, Spain. Email: mperea@uv.es

An often overlooked but fundamental issue for any comprehensive model of visual-word recognition is whether diacritical and non-diacritical vowels share their abstract letter representations? Recent research suggests that the answer is “yes” in languages where diacritics indicate segmental information (e.g., lexical stress, as in cámara [‘ka.máɾa] camera; Spanish), but “no” in languages where diacritics indicate segmental information such as a different phoneme (e.g., the German vowels ö/ and ò/). Here we examined this issue in French, a language that contains a complex set of diacritical vowels (e.g., for the letter: è, ë, ë, and ë). In Experiment 1, using a semantic categorization task, we compared the word identification times to intact diacritical words (e.g., chèvre, goat in English) with a condition with omitted diacritics (chev). Results showed that the two conditions behaved similarly. In Experiments 2–4, we compared the intact diacritical words with a condition containing a mismatching diacritic, either existing in French (e.g., chèvre, chèvre) or not (the macron sign, as in chèvre). We only found a reading cost when replacing the diacritic with an existing one. In Experiments 5–6, we compared the semantic categorization times to intact nondiacritical words (e.g., cheval, horse in English) versus a condition with an added diacritic, either existing (chéval) or not (chéval). We found a reading cost for the words with the added diacritical mark in both cases. We discuss how models of visual-word recognition can be modified to represent diacritical vowels.

Keywords: cross-language differences, diacritics, lexical-semantic access, word recognition

This research was supported by Grants from the Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital, Generalitat Valenciana (CIAICO/2021/172 to Manuel Perea; GV/2020/074 to Ana Marcet) and from the Ministerio de Ciencia, Innovación y Universidades (PID2020-116740GB-I00 funded by the MCIN/AEI/10.13039/501100011033) to Manuel Perea.
system has already extracted the perceptual features and activated the abstract letter units.

A General Framework for Representing Diacritical Letters

Concerning those diacritical consonants that correspond to phonemes that did not exist in Latin (e.g., ā/ā in Spanish, ŗ/ŗ in Czech, etc.), one might reasonably assume that they activate different letter units than their nondiacritical counterparts. Indeed, the letters ŗ and ţ represent speech-specific sounds and form part of the Spanish and Czech alphabets, respectively (e.g., see Marcet et al., 2020, for discussion). Although these diacritical letters are visually similar to their nondiacritical counterparts, this is not a unique case in the Latin alphabet: other pairs of letters also share most of their visual features (e.g., t/f, i/j, C/G, E/F, among others).

Thus, when modeling experiments of visual-word recognition in Latin-based languages with diacritical consonants using platforms such as easyNet (Adelman et al., 2018), one would need to add these consonants to the letter level. This modification would require a more refined letter feature level to capture the extra visual features of the diacritical marks. Keep in mind that the font used by the family of interactive activation models (i.e., the Rumelhart & Siple, 1974; font) cannot be used to represent diacritical marks.

More interesting is the case of diacritical vowels, which is the central issue of the present article. The reason is that the function of diacritics in vowels differs greatly across languages (see Wells, 2000; for review). Diacritics may indicate: (a) another phoneme (i.e., vowel quality; e.g., German: ä/ä vs. a/a); (b) lexical stress (i.e., the stressed syllable under some rules; e.g., Spanish: pājarro ‘[pa.xa.ro] bird’); (c) both vowel quality and lexical stress (e.g., Catalan: desprès [das′pret] later); (d) vowel length (e.g., Czech: vagūn [va.gu:n] wagon); (e) vowel length and vowel quality (e.g., Hungarian: compare ez [ez] this vs. él [el] alive); (f) tone information (Vietnamese: e→i/l [mid falling tone]); or (g) distinguishing otherwise homonyms (e.g., Spanish: él [el] he vs. el [el] the (masculine singular)).

A sensible working hypothesis is that whether or not diacritical and nondiacritical vowels have separate abstract letter units depends on their function in the language (see Marcet et al., 2022; Perea, Labusch, & Marcet, 2022). Before introducing the rationale of the experiments, which examine the role of diacritical vowels in French, we first examine two extreme scenarios: German vs. Spanish.

First, in a language like German, diacritical vowels explicitly refer to a different phoneme than their nondiacritical counterparts. Specifically, German has three diacritical vowels, always with an umlaut (ā, ŏ, ŭ; ā/ā vs. a/a; ō/ō vs. o/o; ŭ/ü vs. u/u). Following the above reasoning on diacritical consonants, one might assume that these diacritical and nondiacritical vowels activate separate abstract letter units (e.g., a ≠ ā). It is just that, as also occurs with C/G or O/Q, they are visually similar. Notably, diacritical vowels in German are considered separate letters from their nondiacritical counterparts at all levels (e.g., when learning to read; in the alphabet; in dictionaries; on computer keyboards, etc.). Consistent with this hypothesis, Hutzler et al. (2004) implemented a connectionist model of word recognition in German containing separate letter units for diacritical and nondiacritical vowels.

To test this account, Perea, Labusch, and Marcet (2022) compared, in a semantic categorization task, the word recognition times to diacritical German words (e.g., Kröte [toad]) when presented intact and when the diacritical mark was omitted (e.g., Krote). Participants were instructed to categorize words as animals or nonanimals, regardless of whether the diacritic was present or omitted (e.g., both Kröte and Krote would be categorized as animals). The logic of the Perea, Labusch, and Marcet (2022) experiment was that if the German vowels õ and ð share their abstract letter units, then word identification times would be virtually similar for the intact word Kröte and its counterpart with the omitted diacritics Krote. Alternatively, if the vowels õ and ð activate separate abstract letter units, word identification times would be slowed down by the diacritics’ omission. Results showed that, although participants were extremely accurate at categorizing the words with the omitted diacritics (i.e., participants could successfully reconstruct the base words), word identification times were substantially longer (around 28–33 ms) when the diacritics were omitted than when they were present (e.g., Kröte > KrOTE). Thus, these findings support Hutzler et al.’s (2004) assumptions concerning separate letter units for diacritical and nondiacritical vowels in German.

Second, in a language like Spanish, diacritical vowels do not modify individual phonemes (e.g., a and á are always pronounced/a/ and the same occurs for e-é/é, i-í/i/, o-ó/o/), and u-ú/u/). Instead, they indicate lexical stress following some accentuation rules (see Appendix A in Labusch, Gómez, & Perea, 2022, for a detailed description1). In this scenario, there would be no reason why diacritical and nondiacritical vowels would be represented as separate units in the mental lexicon (see Chetai & Boursain, 2019; Schwab, 2015). Indeed, when simulating word recognition experiments in Spanish, researchers typically omit the accent marks from diacritical words (e.g., pájaro [pa.xa.ro] [bird] being encoded as pajaro; see Conrad et al., 2010; Perea & Rosa, 2000; for simulations with Grainger & Jacobs, 1996; Multiple Read-Out model).

Consistent with the above idea, in a sentence reading task in Spanish, Marpec and Perea (2022) found remarkably similar first-pass eye fixation times on diacritical words when they were presented intact (e.g., ratón [mousel]) and when the diacritics were omitted (e.g., ratón = raton; the difference was only 3 ms in the first-pass durations on the target word). Parallel evidence has been reported in a semantic categorization task in which participants had to indicate whether the target item was an animal or not. Perea, Labusch, and Marcet (2022) found similar word response times for diacritical Spanish words when presented intact and when the diacritical mark was omitted (e.g., ratón = raton).

Thus, the above findings suggest that the function of diacritical vowels in a language shapes their representations in the word recognition system. In languages where diacritics unambiguously modify segmental information, such as the grapheme-phoneme mapping (i.e., vowel quality; e.g., a/a/ vs. á/á/), as in German, they would be represented as separate letters units. As a result, omitting the diacritical marks in a diacritical word would convey a sizable reading cost. Alternatively, in languages where diacritics only modify suprasegmental information such as lexical stress, as in Spanish, diacritical and nondiacritical vowels would share their abstract representations. In this latter scenario, omitting a word’s diacritics would not entail an appreciable reading cost (see Figure 1 for illustration).

---

1 Spanish also has the very infrequent diacritical vowel ü, which indicates that the vowel u must be pronounced in the sequence gué/gui, as in pingüino [pin gwi.no] penguin.
The Case of French Diacritics

An important remaining question is whether the above ideas can be generalized to a language where the function of diacritical vowels is less straightforward than German (ä, ñ, ü, denoting a different phoneme) or Spanish (á, é, í, ó, ñ, marking the stressed syllable). We chose French because, as pointed out earlier, there is an influential computational model of word recognition in French that assumes separate letter units for diacritical vowels (Multiple Trace-Memory model; Ans et al., 1998).

French contains 12 diacritical vowels, including acute accents (é), grave accents (à, è, ë), circumflexes (â, ê, ë, ÿ, and ù), and diereses (ë, ï, and ü). These diacritics often have a phonological function (e.g., see Le Petit Robert, 2001; Peereman et al., 2007; for a more detailed overview). This is particularly prominent for the vowel e, which is consistently pronounced /e/ when presented with an acute accent (i.e., é), and it is pronounced /ɛ/ when presented with a grave accent (i.e., è). The vowel ê is usually pronounced /ɛ̃/ or /ɛ̃/, but it can be pronounced /ɛ/ in many words (e.g., bêtise [bé.tíz] foolishness; bâlémir [blè.mir] to pale). In addition, the nondiacritical vowel e can be pronounced, depending on some rules, /ɛ/ or /ɛ̃/ (or mute at the end of a word). For instance, compare ëlève [el.lév] pupil vs. ëléve [el.lév] elevated. Instead, the diacritics do not typically alter the sound of the vowels i (i.e., i and ï are pronounced /i/) and u (i.e., û and ü are pronounced /y/)—note that there are exceptions (e.g., compare jeûne [ʒe.ĵn] young vs. jeûne [ʒe.ĵn] fasting). The vowels a and å are typically pronounced /a/ (e.g., achèter [a.chɛ.te] to buy). In contrast, the vowel ë is typically pronounced with a long /a/ (e.g., bâton [ba.tɔ̃] stick)—however, the /a/-/a/ distinction tends to disappear in European French. Depending on some rules, the vowel ë can be pronounced /ɛ/ or /ɛ̃/, whereas the vowel ê is consistently pronounced /ɛ̃/.

In addition, diacritics in French may have other functions: (a) they serve to distinguish otherwise homonym words (e.g., la [feminine “the”] vs. lâ [there]; (b) they reflect the etymology of a word (e.g., the circumflexes in a word like hôpital [hospital] indicate that its Latin ancestor contained the letter “s”); and (c) the diereses indicate that the vowel is pronounced differently from the preceding vowel, creating a hiatus (e.g., compare maïs [mæ] but vs. maïs [mæis] com).

Some empirical support for the special role of diacritical vowels in French, as posited by Ans et al.’s (1998) Multiple-Trace Memory model, came recently from two masked priming experiments conducted by Chetail and Boursain (2019). The logic of these experiments was that if diacritical and nondiacritical vowels in French (e.g., â, a) shared their abstract letter units, the prime â and the prime a would be equally effective for the letter A. However, in a masked priming alphabetic decision task, they found faster response times to â-A than â-Â. To generalize these findings to a word recognition scenario, Chetail and Boursain (2019) conducted a masked priming lexical decision experiment. They found that word identification times to a nondiacritical French target word (e.g., tâper [to type]) were faster when preceded by a lowercase prime that could be the same (e.g., tâper) than when the prime had an extra diacritical mark (täper). Furthermore, the diacritical prime täper was only slightly more effective (i.e., a nonsignificant 7-ms difference) than a control replacement-letter prime such as tuper (e.g., taper-TAPER < tâper-TAPER ≈ tuper-TAPER).3 Chetail and Boursain (2019) concluded that “base letters and their diacritic counterparts activate separated letter representations in scripts such as French” (p. 351).

Chetail and Boursain’s (2019) findings offer relevant information on processing diacritical marks in French. However, one might argue that, for a nondiacritical target word (e.g., TAPER), the diacritical mark in a prime like täper adds some perceptual noise that could have slowed down target processing relative to the identity prime täper. Indeed, in masked priming lexical decision experiments in English (i.e., a language with no diacritics), the pair clâck-CLOCK produces longer response times than the pair clâck-CLOCK (Perea, Gómez, & Baciero, 2022). Thus, at least part of the difference between täper-TAPER vs. tâper-TAPER in the Chetail and Boursain (2019) experiment could have been attributable to perceptual, nonorthographic elements (see also Perea et al., 2020a, for a similar observation in Spanish).

Furthermore, we must bear in mind that, in general terms, priming paradigms do not inform us of the direct activation of a target word. Instead, they inform us about the degree to which a target word is affected by an explicitly presented prime (see

---

2 Of note, prior research has reported small, but reliable visual similarity effects in masked priming (e.g., object-OBJECT faster than object-OBJECT; see Marcet & Perea, 2017, 2018; see also Perea et al., 2020b; Perea, Hyönä, & Marcet, 2022).
The use of an unprimed procedure is a more direct approach to examine whether the abstract letter representations of diacritical vowels are shared with nondiacritical vowels during access to lexical-semantic information in the mental lexicon (see Andrews, 1997; for a similar observation regarding the effects of orthographic neighborhood in word recognition). In the present experiments, we chose an unprimed semantic categorization task ("is the word an animal?"). The semantic categorization task requires participants to access lexical-semantic knowledge, and, unlike lexical decision, it is not easily influenced by visual format (e.g., in lexical decision, houSe produces slower response times than HOUSE, but this difference does not occur in semantic categorization; see Perea et al., 2020b). Another interpretive issue related to the use of the lexical-decision task when comparing diacritical versus nondiacritical items is that participants may treat the diacritical item as more wordlike. For instance, Marcet et al. (2021) found a dissonative pattern for words and nonwords in the lexical-decision task: participants responded faster to intact diacritical Spanish words than to those with an omitted diacritic (e.g., ratōn [mouse] faster than ratōn) but they responded to slowed to diacritical than nondiacritical pseudowords (e.g., bugōn slower than bugōn; see Perea et al., 2020b, for a similar dissociation when comparing same-case versus mixed-case stimuli in the lexical-decision task). Critically, this difference vanished when the diacritical items were presented in one block, and the nondiacritical items were presented in another block. Marcet et al. (2021) interpreted the differences for word stimuli in their first experiment as being due to the characteristics of the task rather than an inherent cost of diacritical processing—it is also a word [young]).

If diacritical and nondiacritical vowels in French are processed as separate letter entries when accessing the mental lexicon, one would expect a substantial reading cost when omitting the diacritics relative to the intact words (e.g., chêvre substantially slower than châ\-vre). This outcome, parallel to that reported in German (see Perea, Labusch, & Marcet, 2022), would support the idea that in French, diacritical and nondiacritical vowels are represented separately in the mental lexicon (Multiple-Trace Memory model, Ans et al., 1998; see also Chetail & Boursain, 2019). Alternatively, if diacritical and nondiacritical vowels in French activate the same letter units when accessing lexical-semantic memory, one would expect very similar word identification times to words like chêvre and chevre (i.e., the same pattern as in Spanish; see Perea, Labusch, & Marcet, 2022). This last outcome would suggest that, in French, despite their ample variety of functions, diacritical vowels share their abstract letter units with their nondiacritical counterparts. Thus, a direct implication of this data pattern is that the letter level in the Multiple-Trace Memory model (Ans et al., 1998) would need to be simplified.

To anticipate the findings of Experiment 1, we found no signs of a reading cost when omitting a word’s diacritic (i.e., the pattern of data was similar to the parallel Spanish experiment conducted by Perea, Labusch, & Marcet, 2022). Although this finding alone has straightforward theoretical implications, we then tested whether there was a reading cost attributable to placing a mismatching diacritic on French words in the following five experiments. Specifically, we examined the impact of replacing the correct diacritic with an incorrect diacritical mark (e.g., chëvre, chëvre, or chëvre for the word chê-vrë; Experiments 2–4) and the impact of adding a diacritical mark on nondiacritical words (e.g., chëval or chäval for the word cheval [horse]; Experiments 4–6). We leave a more detailed explanation of these experiments for brevity’s sake until later.

**Experiment 1 (Intact Versus Omitted Diacritic)**

**Method**

**Participants**

We recruited 50 participants (23 women; mean age = 27.8 years [SD = 5.27]) via the online crowd-working platform Prolific Academic (http://prolific.ac). This sample size ensured at least 1,800 observations in each condition (intact vs. omitted), following Brysbaert and Stevens’s (2018) guidelines for small-sized effects. In this and all subsequent experiments: (a) we used Prolific Academic’s recruitment filters to only include native French speakers with no reading problems and with normal or corrected-to-normal vision; and (b) participants gave informed consent before the experiment, receiving monetary compensation according to Prolific’s average participant salary. Ethical approval for this research was obtained from the Research Ethics Committee of the University of Valencia, and the study followed the requirements of the Helsinki convention.

3 We also preferred the semantic categorization task over the naming task (i.e., another widely used word recognition task) because the latter does not necessarily reflect access to lexical-semantic representations (e.g., the French word élève /ɛl.ɛv/ can be pronounced following a grapheme-to-phoneme route).
Materials

We selected a set of 72 French words with diacritics (e.g., étui [case]) from the LEXIQUE 3 database (www.lexique.org; New et al., 2004, 2007). None of these words were animal names or referred to concepts related to animals (e.g., plants or body parts). All words were common nouns, and according to French spelling rules, they normatively required diacritics. In addition, we selected a set of 36 French words with diacritics that were animal names (e.g., chèvre [goat]) from the LEXIQUE database (New et al., 2004, 2007). The number of letters, the word frequency, OLD-20, and the amount and type of diacritics according to the LEXIQUE-database (New et al., 2004, 2007) were matched between animal names and nonanimal words (see Table 1). The ratio of animals versus nonanimals was the same as in Perea et al.’s (2020a) and Labusch, Gómez, and Perea’s (2022) semantic categorization experiments on the role of diacritics during word recognition. We chose this ratio because of the limited amount of diacritical animal words in French—note that prior research using this same ratio has shown the same pattern of findings for animal and nonanimal words using other manipulations (e.g., letter rotation: Fernández-López et al., 2022; case alternation; Perea et al., 2020a). Each item, always in lowercase, was presented intact (i.e., with its corresponding diacritic, e.g., chèvre, etui) or without the diacritic (e.g., chevre, etui). We created two counterbalanced lists in which 54 words were presented with the diacritic (18 animal nouns, 36 common nouns), and 54 words were presented without the diacritic (18 animal nouns, 36 common nouns). Those words that kept the diacritic in List 1 were presented without it in List 2 and vice versa. Each participant was assigned randomly to one of the two lists. The list of stimuli 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 servers Pavlovia (www.pavlovia.org) and LimeSurvey (www.limesurvey.org). Before the beginning of the experiment, participants filled out a questionnaire with demographic data (age, gender, level of education). They also went through sixteen practice trials to be familiarized with the task. They were instructed to do the experiment in a quiet room without any distractions. During the experiment, participants had to complete a semantic categorization task by answering whether the presented word referred to an animal name or not. As each word was presented individually, they had to press the button “L” on their keyboard for “yes” and the button “S” on their keyboard for “no” as fast and accurately as possible. Participants were instructed to classify a word as “animal” irrespective of whether it was written with the correct diacritic or not. Thus, both chèvre and chevre ought to be classified as “animals.” Before presenting each word, a fixation cross appeared for 500 ms in the center of the screen. Afterward, the word appeared at the same location until a response was made (or until a time limit of 2,000 ms). The trials were presented in a randomized order for each participant. Altogether, the experiment took 7–8 minutes, including a short break after 56 trials.

Data Analysis

This study was not preregistered. We created Bayesian linear mixed-effects models for each dependent variable using the brms package (Bürkner, 2016) in the R environment (R Core Team, 2021). The two fixed factors of the models were Format (diacritical vs. nondiacritical; encoded as −.5 and .5) and Type of Word (animal vs. nonanimal; encoded as −.5 and .5)—note that the critical factor was Format. Following the suggestion of Barr et al. (2013), we fit the models with the maximal random-effect structure in the design:

\[
\text{Response Time [accuracy] \sim Format \times TypeWord } + (1 + \text{Format} \times \text{TypeWord} \mid \text{subject}) + (1 + \text{Format} \mid \text{item})
\]

Because latency data typically show a positive skew, we chose to model these data with the Gaussian distribution (family = Gaussian()) after the −1000/RT transformation. This transformation equals the number of words per second, and the negative sign was to keep the same direction of the effects as the untransformed data. As accuracy data occurs in a binary manner (correct = 1; incorrect = 0), we modeled these data with the Bernoulli distribution (family = bernoulli()). Each model was executed with four chains, including 5,000 iterations with a warm-up of 1,000 iterations in each chain. The model’s output provided an estimate of each parameter (i.e., the mean of its posterior distribution in brms), its standard error, and the 95% credible intervals (95% CrI). We considered evidence of an effect when the 95% CrI of its estimate did not contain zero. We also presented the posterior distributions of each estimate in both RT and accuracy models. The data, scripts, and outputs of this and the rest of the experiments are available at https://osf.io/e9kup/ (Perea, Marcet, et al., 2022).

Results and Discussion

In this and the following experiments, we excluded incorrect responses and response times shorter than 250 ms from the response time (RT) analyses. Due to the 2,000 ms deadline, there were no latencies above that duration. The mean RTs and error rates in each condition are presented in Table 2. The Bayesian linear mixed-effects models on the RTs and the accuracy data converged successfully, and all Rs were less than 1.01. The focus of the analysis was to

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison of the Mean (SD) Characteristics of the Word Items in Experiment 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Words</th>
<th>Word frequency (per million)</th>
<th>Word length</th>
<th>OLD-20</th>
<th>Types of diacritics used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animal nouns</td>
<td>2.49 (3.31)</td>
<td>6.81 (1.55)</td>
<td>2.46 (0.67)</td>
<td>é, ê, î, â, ï</td>
</tr>
<tr>
<td>Common nouns</td>
<td>2.46 (2.62)</td>
<td>7.11 (2.03)</td>
<td>2.31 (0.64)</td>
<td>é, ê, î, â, ï</td>
</tr>
<tr>
<td>p value (t-test)</td>
<td>0.95</td>
<td>0.42</td>
<td>0.25</td>
<td></td>
</tr>
</tbody>
</table>

This study was not preregistered. We created Bayesian linear mixed-effects models for each dependent variable using the brms package (Bürkner, 2016) in the R environment (R Core Team, 2021). The two fixed factors of the models were Format (diacritical vs. nondiacritical; encoded as −.5 and .5) and Type of Word (animal vs. nonanimal; encoded as −.5 and .5)—note that the critical factor was Format. Following the suggestion of Barr et al. (2013), we fit the models with the maximal random-effect structure in the design:
Table 2
Mean Response Times (in ms) and Error Rates (in Percentages) for Nonanimal Words and Animal Words Written With Their Diacritics Present or Omitted in Experiment 1

<table>
<thead>
<tr>
<th>Word</th>
<th>Diacritic present</th>
<th></th>
<th>Diacritic omitted</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Response time</td>
<td>Error rate</td>
<td>Response time</td>
<td>Error rate</td>
</tr>
<tr>
<td>Nonanimals</td>
<td>631</td>
<td>3.5</td>
<td>640</td>
<td>3.9</td>
</tr>
<tr>
<td>Animals</td>
<td>626</td>
<td>6.3</td>
<td>630</td>
<td>5.7</td>
</tr>
</tbody>
</table>

compare the intact words and the words with the omitted diacritic (or replaced/added diacritic in other experiments). For completeness, as in prior research (see Labusch, Gómez, & Perea, 2022; Labusch, Kotz, & Perea, 2022; Perea, Labusch, & Marcet, 2022), we also reported the effect of the category (animal vs. Nonanimal) and the two-way interaction between the two factors—note that “animals” and “nonanimals” require different responses. Thus, we prefer to analyze and report all outcomes.

Response Time Analysis

Response times were only 7 ms faster to the intact words than to the words with the omitted diacritic (629 vs. 636 ms, respectively)—the 95% CrI of this difference crossed zero: \( b = .01, \ SE = .01, 95\% \text{ CrI} [−.01, .04] \).

In addition, response times were only slightly faster to animal words than to nonanimal words (628 vs. 636 ms, respectively; \( b = −.01, \ SE = .03, 95\% \text{ CrI} [−.05, .05] \)). Finally, there were no signs of an interaction between the two factors (\( b = .01, \ SE = .02, 95\% \text{ CrI} [−.03, .04] \)). For the posterior distributions, see the left panel of Figure 2.

Accuracy Analysis

There were similar error rates for words that were written with their diacritic and words without diacritic (4.9 vs. 4.8%, respectively; \( b = .01, \ SE = .27, 95\% \text{ CrI} [−.53, .53] \)). Participants made fewer errors to nonanimal than to animal words (3.7% vs. 6.0%, respectively; \( b = 1.00, \ SE = .37, 95\% \text{ CrI} [.30, 1.76] \)). No signs of an interaction between the two factors were observed (\( b = −.26, \ SE = .34, 95\% \text{ CrI} [−.93, .40] \)). For the posterior distributions, see the right panel of Figure 2.

The present experiment showed a minimal advantage for those words presented intact than when the diacritical mark was omitted (629 vs. 636 ms, respectively). This minor difference resembles more that reported in a parallel experiment in Spanish (a 4-ms difference) than that reported in German (a 30-ms difference; Perea, Labusch, & Marcet, 2022).

For consistency with the Perea, Labusch, and Marcet (2022) experiments, we conducted a descriptive quantile-based analysis of the reading cost (omitted condition minus intact condition) using delta plots (see Riddervik, 2002; for a description of the utility of these plots). Specifically, we obtained the average reading cost across participants at the .1, .3, .5, .7, and .9 quantiles for animal and nonanimal words. Then, we computed the subtraction of these average RTs (i.e., RTs in the condition with the omitted diacritics—RTs in the intact condition). A flat line centered in zero would reflect no reading cost (i.e., a null cost across the various quantiles of the RT distribution). As shown in Figure 3, there were no apparent signs of a reading cost across the RT distribution in French. Indeed, this plot resembles much more that of the parallel Spanish experiment than the German one (see Figures 2 and 4, respectively, in Perea, Labusch, & Marcet, 2022). As in the Spanish experiment, there was a slight cost for the higher quantiles; however, this effect was negligible at the .5 quantile or lower.

Therefore, omitting the diacritical mark only has a minimal effect using a task that measures access to lexical-semantic information. This finding suggests that diacritical vowels may not have separate letter representations in French. If they were, one would have expected substantially faster word identification times to \( \text{chévre} \) than \( \text{Kröte} \), as recently reported in German (e.g., \( \text{Kröte} \) [toad] produces much more fast word identification times than \( \text{chévre} \); Perea, Labusch, & Marcet, 2022).

This outcome aligns with the observations of the daily usage of diacritics in native speakers of computer-mediated French (see Anis,
In Experiment 3, using the same set of intact words as Experiment 2, we replaced the diacritical mark of the letter e because this modification involves unambiguously different phonemes (i.e., é/l vs. è/ê; chêvre [jɛvʁ] vs. chèvre [jɛvʁ]). Thus, this scenario is closer to the German experiments conducted by Perea, Labusch, and Marcit (2022) in the sense that, as occurs with the vowels a and ò in German, the French vowels é and ë correspond to different phonemes. In Experiment 3, using the same set of intact words as Experiment 2, we replaced the diacritical mark of the letter e with another diacritic that did not necessarily change the word’s phonology (i.e., a circumflex, e.g., chêvre). Finally, Experiment 4 tested whether there was a reading cost when replacing the word’s diacritic with a mismatching diacritic that had no orthographic/phonological value (i.e., the macron, a nonexisting diacritic in French, e.g., chèvre). As in Experiment 1, participants were told that the word’s diacritical mark could be altered in some of the words (e.g., chèvre→chévre) but that both forms would correspond to the same answer (e.g., “yes” for chèvre and chêvre).

Experiment 2 used the most extreme manipulation: for example, chèvre→chévre, as it switched an explicit marker of vowel quality (é/l vs. è/ê). If an incorrect unambiguous marker of vowel quality of one of the letters (é for the word chèvre, as in chêvre) has an impact on the access to lexical-semantic information, word recognition times would be faster for the intact words than for the words with the modified diacritic (e.g., chèvre < chêvre). Alternatively, if the word recognition system in French is virtually insensitive to the type of diacritics, even those explicitly marking vowel quality, then one would expect similar response times to the intact words and the words with the inverted diacritical mark (e.g., chèvre > chêvre).

**Experiment 2 (Correct Versus French Diacritic That Changes Vowel Quality)**

**Method**

**Participants**

Using Prolific Academic, we tested 50 additional participants (23 women; mean age = 26.7 years [SD = 5.21]) to obtain the same number of observations as Experiment 1. Participants had to fulfill the same recruitment criteria as in the previous experiment. They all gave informed consent before the experiment and received monetary compensation for their participation.
Materials

We selected the same set of words as in Experiment 1, except that we replaced 16 words that did not contain the diacritical vowels ê or è with 16 words containing one of these diacritics. The resulting 72 French common nouns and 36 French animal names were matched in word frequency, word length, OLD-20, amount, and type of diacritics (see Table 3, for an overview) and were all written either with the diacritic ê (e.g., guêpard, étui) or with the diacritic è (e.g., chèvre, cuillère). Thus, each item was presented with its corresponding diacritic (e.g., étui, chèvre) or with the inverted form of its diacritic (e.g., âtui, chêvre).

As in Experiment 1, we created two counterbalanced lists and assigned participants to them in the same manner than in the previous experiment. For the full list of words, see Appendix.

Procedure

The experimental procedure was parallel to that in Experiment 1. Participants were instructed to classify a word as “animal” irrespective of whether it was written with the correct diacritic or not (e.g., both chèvre and chêvre would correspond to “animal”).

Data Analysis

We followed the same steps for data analysis as in Experiment 1.

Results and Discussion

Table 4 shows the average RT and error rate for each experimental condition. Again, the Bayesian linear mixed-effects models on the latency and accuracy data converged successfully (R = 1.00 in all estimates).

Response Time Analysis

Participants responded, on average, 20 ms faster to words with the correct diacritic than to words with an inverted diacritic (669 ms vs. 689 ms, respectively; b = .04, SE = .01, 95% CrI [.01, .07]).

In addition, the differences in response times between animal words and nonanimal words were minimal (678 ms vs. 682 ms, respectively; b = −.01, SE = .03, 95% CrI [−.07, .04]). There were also no signs of interaction between the two factors (b = .00, SE = .02, 95% CrI [−.03, .03]; see the left panel of Figure 4 for the posterior distributions).

Accuracy Analysis

There was no effect of the type of diacritics on accuracy ratings, b = .38, SE = .25, 95% CrI [−.09, .91].

In addition, participants made fewer errors to nonanimal words than animal words (3.9% vs. 8.3%, respectively; b = 1.19, SE = .41, 95% CrI [.37, 2.00]), but this effect did not interact with the type of diacritics, b = −.25, SE = .33, 95% CrI [−.88, .40]. The right panel of Figure 4 presents the posterior distributions.

The present experiment showed that intact words were identified faster than those with an inverted diacritical mark. This difference occurred similarly for animal words (666 ms vs. 689 ms, respectively, e.g., chèvre < chêvre) and nonanimal words (673 ms vs. 690 ms, respectively, flèche [arrow] < fléche). Therefore, it is possible to obtain a reading cost in French when replacing a word’s diacritical mark with another diacritical mark, at least when it involves mismatching orthographic and phonological information (e.g., chèvre [jœv] < chêvre [jœvi]).

Experiment 3 examined whether access to lexical-semantic information in diacritical French words is slowed down when replacing a word’s diacritical mark with another diacritical mark that does not necessarily involve a change in phonology: the circumflexed vowel ê (e.g., as in chèvre). This diacritical mark, which exists in French orthography, is often pronounced /e/. However, the letter ê can also be pronounced /ɛ/ depending on the words and the speakers. For instance, as stated in the Introduction, there are a number of French words in which the letter ê is consistently pronounced /ɛ/ (e.g., blêmir [ble.mi] blush) or in which the letter ê may be pronounced /ɛ/ or /ɛ/ depending on the speaker (e.g., crépière [kre pijœ] or [kre.pijœ] pancake maker).

Thus, in Experiment 3, the replaced diacritical mark did not involve an unambiguous change in the word’s pronunciation. Therefore, if the reading cost found in Experiment 2 were entirely attributable to the activation of incorrect phonological codes (e.g., /ɛ/ for chèvre vs. /ɛ/ for chêvre), one would expect similar response times for the intact words and for those words with a circumflex replacing the correct diacritical mark (e.g., chêvre ≈ chêvre). Conversely, if the reading cost resulting from the replaced diacritic is mainly attributable to the mere presence of a mismatching diacritical mark (i.e., the letter ê in chèvre instead of the correct vocal ê), one would expect longer word response times for the words with the circumflex mark when compared with the intact words (e.g., chèvre > chêvre). Of note, we conducted some exploratory analyses to study whether the reading cost differed for those circumflexed words in which native speakers of

Table 4
Mean Response Times (in ms) and Error Rates (in Percentages) for Nonanimal Words and Animal Words Written With Their Correct Diacritics or Inverted Diacritics in Experiment 2

<table>
<thead>
<tr>
<th>Word</th>
<th>Correct diacritic</th>
<th>Inverted diacritic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Response time</td>
<td>Error rate</td>
</tr>
<tr>
<td>Nonanimals</td>
<td>673</td>
<td>3.7</td>
</tr>
<tr>
<td>Animals</td>
<td>666</td>
<td>9.0</td>
</tr>
</tbody>
</table>

Table 3
Comparison of the Mean (SD) Characteristics of the Word Items of Experiment 2, 3, and 4

<table>
<thead>
<tr>
<th>Words</th>
<th>Word frequency (per million)</th>
<th>Word length</th>
<th>OLD-20</th>
<th>Types of diacritics used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animal nouns</td>
<td>2.02 (2.91)</td>
<td>7.08 (1.45)</td>
<td>2.57 (0.69)</td>
<td>é, ê (Exp. 2) é (Exp. 3)</td>
</tr>
<tr>
<td>Common nouns</td>
<td>2.16 (2.42)</td>
<td>7.19 (1.95)</td>
<td>2.31 (0.62)</td>
<td>é, ê (Exp. 2) é (Exp. 3)</td>
</tr>
</tbody>
</table>

p value (t test) 0.81 0.77 0.051
French detected a change in phonology and for those circumflexed words that would be pronounced as the intact word.

**Experiment 3 (Correct Versus French Diacritic That Does Not Imply a Change in Pronunciation)**

**Method**

**Participants**

We recruited an additional 50 participants (21 women, mean age = 27.5 years [SD = 5.32]) via Prolific Academia. We used the same recruitment filters as in the previous experiments, presented a consent form, and got the same number of observations per condition.

**Materials**

We used the same materials as in Experiment 2 (see Table 3). The only difference was that we replaced the diacritical marks é or è with a circumflex (ê; e.g., chèvre was compared with chêvre).

**Data Analysis and Procedure**

They were the same as in the previous experiments.

**Results and Discussion**

Table 5 shows the average RT and error rate for each experimental condition. Again, the Bayesian linear mixed-effects models on the latency and accuracy data converged successfully (R^2 = 1.00 in all estimates).

**Response Time Analysis**

Participants responded around 10 ms faster to words with the correct diacritic than to words with an incorrect diacritic (624 vs. 634 ms, respectively; b = .04, SE = .01, 95% CrI [.01,.06]).

In addition, we found no differences in response times between animal words and nonanimal words (631 vs. 627 ms, respectively; b = -.02, SE = .03, 95% CrI [-.07, .04]). There were no signs of an interaction between the two factors (b = -.01, SE = .02, 95% CrI [-.05, .02]; see the left panel of Figure 5 for the posterior distributions).

**Accuracy Analysis**

We found no signs of an effect of the type of diacritics, b = -.13, SE = .24, 95% CrI [-.59, .31].

In addition, participants made fewer errors to nonanimal words than animal words (4.3% vs. 6.4%, respectively; b = .82, SE = .37, 95% CrI [.10, 1.54]), but this effect did not interact with the type of diacritics (b = -.12, SE = .31, 95% CrI [-.74, .47]). The right panel of Figure 5 presents the posterior distributions.

The present experiments revealed a reading cost of the circumflexed words (chèvre) relative to the intact condition (chêvre). The posterior distributions showed that the evidence was clear (see left panel of Figure 5); however, the size of the effect was numerically smaller (10 ms) than in Experiment 2, where the replaced diacritical mark necessarily involved a change in vowel quality (20 ms; chèvre < chêvre; i.e., é is always pronounced/e/ whereas è is always pronounced/ɛ/).

To examine whether the diminished reading cost in the present experiment was affected by phonology, we conducted an exploratory post hoc analysis. Eight additional native speakers of French indicated, on an item basis for all the words, whether the circumflexed words
caused a change in the pronunciation relative to the intact words (e.g., “how would ê be pronounced in ‘pédicure’?”). For 73% of the words, participants indicated that the pronunciation of the circumflexed word was identical to that of the intact word (e.g., pédicure; i.e., [pe.di.ky]). About 26% of the words were classified as ambiguous (e.g., the word caféine [base word: caféine [ka.fe.in] could be pronounced as [ka.fe.in] or [ka.fe.in]) and, finally, for only one word, participants unanimously decided to pronounce the circumflexed ê differently than in the original word (e.g., chronomètre would be pronounced as chronomètre [kro.nɔ.mɛtʁ], but the original word was chronomètre [kro.nɔ.mɛtʁ]). When averaging the reading cost across items, we found approximately similar costs for those circumflexed words with ambiguous pronunciation (12.2 ms) and those with a pronunciation that coincides with the original word (9.7 ms). Although we prefer to remain cautious about these exploratory analyses, they suggest that the observed reading cost when replacing the acute/grave diacritic with a circumflex may not have been primarily affected by pronunciation.

In sum, Experiments 2–3 revealed a reading cost when replacing the intact diacritical marks with other diacritical marks that are naturally occurring in French (i.e., grave, acute, circumflex). A sensible question to ask is whether this reading cost was mainly attributable to: (a) the replaced diacritic being familiar to French speakers, thus producing some interference at an orthographic/phonological level; or (b) a general effect of perceptual noise generated by any mismatching mark that replaced the original diacritic.

To examine this issue, in Experiment 4, we replaced the original diacritical mark with a diacritical mark that does not exist in French (i.e., the macron sign; e.g., ê, as in chêvre) and compared them to the intact words (chêvre). We used the same set of words as in Experiments 2–3.

If the observed reading cost for the words with a mismatching diacritical mark found in Experiments 2–3 was attributable to the activation of mismatching orthographic/phonological codes, we would expect this cost to vanish when a macron replaced the correct diacritical mark (e.g., chêvre ≈ chêvre). Conversely, if the reading cost of the replaced diacritical marks in Experiments 2–3 was mainly attributable to the added perceptual noise caused by the presence of a mismatching diacritic (i.e., regardless of its function), one would expect longer word response times for the words with the macron-replaced mark when compared with the intact words (e.g., chêvre > chêvre).

Experiment 4 (Correct Versus Nonexisting Diacritic)

Method

Participants

An additional sample of 50 participants (20 women; mean age = 27.2 years [SD = 5.37]) with the same profile as in the previous experiments was recruited via Prolific Academic, thus resulting in the same number of observations. Before the experiment, all participants signed a consent form and received a small monetary compensation.

Materials

We used the same words as in Experiments 2–3 (see Table 3). The only difference was that each word was presented with its corresponding diacritic (e.g., étui, chêvre) or with a macron (i.e., a neutral diacritic that is never used in French; e.g., étui, chêvre). The two counterbalanced lists were created in the same way as in the previous experiments.

Procedure and Data Analysis

They were parallel to the prior experiments.

Results and Discussion

The Bayesian linear mixed-effects models converged successfully and R^2 = 1.00 in all estimates. Table 6 shows the average RT and error rate for each condition.
Response Time Analysis

Participants responded, on average, 5 ms faster to the intact words than to the words with a neutral diacritic (653 vs. 658 ms, respectively)—note that the estimate of this difference crossed zero: $b = .02, SE = .01, 95\% CI [-.01, .05]$.

There was also a negligible difference in response times between animal words and nonanimal words (657 vs. 655 ms, respectively; $b = -.02, SE = .03, 95\% CI [-.08, .04]$), and no signs of an interaction between the two factors ($b = -.02, SE = .02, 95\% CI [-.05, .02]$). For the posterior distributions, see the left panel of Figure 6.

Accuracy Analysis

There were no signs of an effect of the type of diacritics ($b = .04, SE = .24, 95\% CI [-.41, .51]$).

In addition, Participants made fewer errors in trials with nonanimal words (3.6% vs. 7.1%, respectively; $b = 1.01, SE = .34, 95\% CI [1.37, 1.69]$). We found no signs of an interaction between the two factors ($b = -.10, SE = .31, 95\% CI [-.71, .52]$). For the posterior distributions, see the right panel of Figure 6.

The present experiment showed that word identification times were only minimally shorter when the words were presented intact than when a macron replaced the diacritical mark (653 vs. 658 ms, respectively), thus resembling the pattern observed in Experiment 1. In other words, replacing the original diacritic with a nonexistent diacritical mark only entails a negligible reading cost. This outcome suggests that the reading cost observed in Experiments 2–3 (with an existing mismatching diacritic replacing the original diacritic) was primarily attributable to the mismatching information from a familiar diacritic.

Summary of Experiments 1–4

We have shown that the access to lexical-semantic information of diacritical French words (e.g., chèvre) is only minimally delayed when the diacritical mark is omitted (e.g., chevre; Experiment 1) or replaced with a nonexistent diacritical mark (Experiment 4; see Figures 2 and 6 for the posterior distributions of the effects). These data can be parsimoniously interpreted in terms of diacritical and nondiacritical vowels sharing their abstract letter representations in French, as has been claimed in Spanish (e.g., Marcet & Perea, 2022; Perea et al., 2020b; Perea, Labusch, & Marcet, 2022). At the same time, the findings of Experiments 2 and 3, in which the word’s diacritical mark was replaced by an existing diacritic (e.g., chèvre, chêvre), revealed a reading cost (see Figures 3 and 4 for the posterior distributions of the effects). Thus, replacing the correct diacritic with a mismatching existing diacritic may slow down access to lexical-semantic representations in French.

How can a model of word recognition account for this dissociative pattern? Cubelli and Beschin (2005) proposed a hypothesis that can explain the observed patterns in a conclusive way. They suggested that the word recognition system processes a word’s diacritical mark in parallel to letter identity. More precisely, during visual word recognition, information on the diacritical marks of a word would be stored and processed as a visual cue for lexical access. A mismatch between the perceptual representation of the
word and its stored representation would induce a reading cost. Thus, the Cubelli and Beschin (2005) model can readily account for the reading cost with mismatching diacritics in Experiment 2–3: the diacritical marks ê in chèvre or ô in chêvre would induce some orthographic/phonological mismatch relative to a stored representation of the word chèvre, thus delaying the access to lexical-semantic information. Notably, this interpretation can also explain the minor but consistent reading cost for the words with the omitted diacritical mark (or with nonexisting diacritics) found in semantic categorization tasks in French (6 ms in Experiment 1; 5 ms in Experiment 4) and Spanish (4 ms slower for ratón when compared with rataón [mouse], Perea, Labusch, & Marcet, 2022; see also Marcet & Perea, 2022).

**Rationale of Experiments 5–6: Addition of Diacritics**

To obtain a complete picture of the role of diacritics during the access to lexical-semantic information in French, it is important to also examine the potential cost of adding diacritics to nondiacritical words (i.e., not just replacing the correct diacritic with a mismatching one). Prior word recognition experiments on this issue are scarce. In French, Chetail and Boursain (2019) used masked priming lexical decision to compare the impact of an intact identity prime (e.g., tajo) versus a prime with an additional diacritic (e.g., tajo̞) on the processing of nondiacritical target words (e.g., TAPER). They found longer response times in the condition with the additional diacritic than in the identity condition. This pattern was interpreted in terms of diacratical and nondiacritical French vowels activating separate letter units. However, as we indicated in the Introduction, the diacritical vowel may add some perceptual noise when identifying a nondiacritical target. Indeed, Perea, Gómez, and Baciero (2022) found this same pattern in a masked priming lexical decision experiment in English (e.g., clóck-CLOCK produced longer latencies than clock-CLOCK)—note that, as English orthography lacks diacritics, native English speakers do not have orthographic representations for diacritical vowels.

Notably, in a recent semantic categorization experiment in Spanish (i.e., a language in which diacratical and nondiacritical vowels presumably share their abstract letter units), Labusch, Gómez, and Perea (2022) found a small, but reliable cost of adding a diacritical mark in the unstressed syllable in Spanish (around 7–9 ms). For instance, the intact word cebra ['θebra] (zebra) was responded to faster than cebra [θebra]. In addition, Labusch, Gómez, and Perea (2022) found that this cost was negligible when the added diacritic occurred on the stressed syllable (e.g., cebra for cebra). Labusch, Gómez, and Perea (2022) interpreted their findings in terms of the Cubelli and Beschin (2005) proposal mentioned earlier: cebra would involve some mismatching of orthographic and phonological information with the stored entry (i.e., cebra), slowing down its recognition.

Thus, in this block of experiments, we examined whether adding a diacritical mark to a nondiacritical French word (e.g., cheval [horse]) delays access to lexical-semantic information. Similar to the experiments from the previous block, we added a diacritical mark that existed in French (Experiment 5) or a diacritic that did not exist in French (Experiment 6). Parallel to the previous experiments, participants were instructed to respond to “animal” vs. “nonanimal” regardless of whether a diacritical mark was added to the word.

Thus, in Experiment 5, we examined whether the access to lexical-semantic information is hindered when adding an existing diacritical mark to a nondiacritical French word (e.g., whether chèval produced longer response times than cheval). If adding an existing diacritical mark to a nondiacritical French word hinders the access to lexical-semantic information, one would expect a reading cost relative to the intact words (e.g., chêval > cheval). This outcome would favor the proposal that the information from the diacritical mark is processed in parallel with abstract letter identities (Cubelli & Beschin, 2005), extending to French the recent findings reported in Spanish (see Perea, Gómez, & Baciero, 2022). Of note, we conducted an exploratory analysis to study whether the reading cost differed for those words where the diacritic was added to the vowel é—note that ô or ê are French vowels with unambiguous phonological codes—or to the other vowels.

**Experiment 5 (Intact Versus Added [Existing] Diacritic)**

**Method**

**Participants**

We recruited 50 additional individuals (22 women; mean age = 31.28 years [SD = 9.7]) via Prolific Academia with the same recruitment filters as in the previous experiments. All participants gave informed consent to participate in the study.

**Materials**

We selected a set of 108 French words from the LEXIQUE 3 database (www.lexique.org; New et al., 2004, 2007), of which 36 were animal names and 72 were nonanimal common nouns that did not refer to animals or concepts related to animals. Importantly, all the words were common nouns and were written without any diacritical mark according to the French orthographic rules (for a full list of the words, see the Appendix). The number of letters, the word frequency, and the OLD-20 according to the LEXIQUE database (New et al., 2004, 2007) were matched between animal names and nonanimal words (see Table 7). Each item was presented in its usual form without a diacritical mark (e.g., cheval) or with a diacritical mark that was added to one of the vowels within the word (e.g., chêval). We chose only viable combinations of diacritics in French, and the amount and type of diacritics were counterbalanced for animal names and common (nonanimal) nouns. The stimulus lists were created in the same way as in Experiments 1–4.

**Procedure**

The experimental procedure was the same as in Experiments 1–4. In this case, participants were told to classify a word as “animal” irrespective of whether it was written intact or with an added diacritic (e.g., cheval and chêval were ought to be classified as “animal”).

**Data Analysis**

We followed the same steps for data analysis as in Experiments 1–4. The fixed factors were Format (added diacritics vs. [intact] nondiacritics; encoded as −.5 and .5) and type of Word (animal vs. nonanimal; encoded as −.5 and .5).
Results and Discussion

The analyses were similar to those in previous experiments. Both Bayesian linear mixed-effects models converged successfully (all $R^2_\text{p} = 1.00$). The mean RTs and the error rates for each condition are given in Table 8.

Response Time Analysis

Responses to intact words were, on average, 19 ms shorter than to the words with an added diacritic (606 vs. 625 ms, respectively; $b = .04$, $SE = .01$, 95% CI [$0.01, .07$]).

We also found a small difference in response times between animal words and nonanimal words (621 vs. 611 ms, respectively), but its estimate crossed zero; $b = -.01$, $SE = .03$, 95% CI [$-.07, .04$]. There were no signs of a two-way interaction, $b = -.00$, $SE = .02$, 95% CI [$-.03, .03$]. The left panel of Figure 7 displays the posterior distributions.

Accuracy Analysis

Responses to the intact words were only minimally more accurate than to words with an additional diacritic (error rates: 2.2% vs. 2.8%, respectively; $b = .45$, $SE = .33$, 95% CI [$-.19, 1.14$]).

In addition, participants made fewer errors in trials with nonanimal words than in trials with animal words (.8% error rate for nonanimal words vs. 4.2% error rate for animal words; $b = 2.21$, $SE = .45$, 95% CI [1.43, 3.20]). We did not find an interaction between the two factors ($b = -.63$, $SE = .56$, 95% CI [$-.74, .48$]). The right panel of Figure 7 displays the posterior distributions.

The current experiment showed faster responses to nondiacritical French words when presented intact than when adding an existing diacritical mark to one of the vowels (e.g., cheval [606 ms] < chêval [625 ms]; see Figure 7).

To further scrutinize this reading cost, we conducted an exploratory post hoc analysis that examined whether the effect differed for those words in which an acute/grave accent was added to the vowel e (e.g., $e$ was changed to $ê$ or $é$; 48 words) and for those words in which the diacritical mark was added to another vowel (i.e., as in $ê$, $û$; 60 words). As we noted earlier, the additional acute/grave diacritics to the letter e (resulting in $ê$ or $é$) may lead to a different pronunciation of the letter, whereas this is not necessarily the case for other additional diacritical marks in French (compare chèval [ʃɛvalu] horse vs. chêval [ʃɛval] and loup [lu] wolf vs. loûp [lu]). When averaging the reading cost by items, for those words where the diacritical mark was added to the letter e (e.g., cheval [horse] vs. chêval), we found a reading cost of 30 ms relative to the intact condition. Instead, this reading cost was only 8 ms for those words when the diacritical mark was added to another vowel (e.g., loup [wolf] vs. loûp). Even though this post hoc analysis should be treated with caution, these findings seem to suggest that the bulk of the reading cost occurred mainly for those words in which the extra diacritical mark was placed on the letter e (i.e., there was a reading cost when reconstructing e from $ê$ in chêval). Instead, for those words in which the diacritical mark was placed on a letter other than e, the reading cost was small (8 ms; e.g., as in loup [lu] vs. loûp [lu]) and similar to that reported recently in Spanish (Labusch, Gómez, & Perea, 2022).

Thus, the present experiment showed that adding a diacritical mark to a nondiacritical French word may entail a reading cost. This finding contrasts with the minimal cost observed in Experiment 1, where the omission of the diacritics barely reflected a reading cost. Nonetheless, this dissociation can be interpreted as reflecting that omitting a feature from a given stimulus is not the same as adding an element to the stimulus (see Treisman & Gormican, 1988; Tversky, 1977; for models of perceptual asymmetries). Consistent with this idea, several masked priming experiments have shown a perceptual asymmetry in several languages containing diacritics: a $\rightarrow$ å, but å $\not\rightarrow$ a (see Perea et al., 2020b, 2021, for vowels; see Marcet et al., 2020, for consonants; see Kinoshita et al., 2021, for hiragana letters). Of note, the cost from the diacritical primes also occurs with native English speakers, in a language lacking diacritical vowels (e.g., month-month slower than monâł-monâł; see Perea, Labusch, & Marcet, 2022).

However, from the present experiment, it is uncertain whether the observed reading cost was attributable to adding a familiar, existing diacritical mark or whether the cost originated at a more perceptual level. Experiment 6 was designed to tease apart the familiarity vs. perceptual explanations. Specifically, in Experiment 6, we assessed whether adding a nonexistent diacritical mark (i.e., a macron, as in $ê$ [see Experiment 4 for the use of this same diacritical mark]) with no orthographic/phonological value to an otherwise nondiacritical word (e.g., chèval for cheval) hinders the access to lexical-semantic information. We used the same words as in Experiment 5.

If the sole addition of a diacritical mark to a nondiacritical word, even one that does not exist in the language, makes the percept less similar to the stored representation of the word in lexical memory, we would expect slower response times for those words with the added diacritical vowel than for the intact words (e.g.,

---

### Table 7
Comparison of the Mean (SD) Characteristics of the Word Items in Experiments 5 and 6

<table>
<thead>
<tr>
<th>Words</th>
<th>Word frequency (per million)</th>
<th>Word length</th>
<th>OLD-20</th>
<th>Types of diacritics used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animal nouns</td>
<td>18.89 (30.2)</td>
<td>6.17 (1.38)</td>
<td>1.88 (0.52)</td>
<td>é, è, ë, â, ï, û, ô</td>
</tr>
<tr>
<td>Common nouns</td>
<td>18.88 (24.3)</td>
<td>6.61 (1.36)</td>
<td>1.95 (0.44)</td>
<td>é, è, ë, â, ï, û, ô</td>
</tr>
<tr>
<td>$p$ value (t test)</td>
<td>0.99</td>
<td>0.11</td>
<td>0.42</td>
<td></td>
</tr>
</tbody>
</table>

---

### Table 8
Mean Response Times (in ms) and Error Rates (in Percentages) for Nonanimal Words and Animal Words Written Intact (i.e., Without Diacritics) With an Additional Diacritical in Experiment 5

<table>
<thead>
<tr>
<th>Word</th>
<th>Non diacritic</th>
<th>Additional diacritic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Response time</td>
<td>Error rate</td>
</tr>
<tr>
<td>Nonanimals</td>
<td>603</td>
<td>0.9</td>
</tr>
<tr>
<td>Animals</td>
<td>609</td>
<td>3.4</td>
</tr>
</tbody>
</table>
chêval slower than cheval). This outcome would favor a perceptual explanation of the reading cost due to the extra diacritical mark. Conversely, if adding a nonexisting diacritical mark to a nondiacritical word like cheval does not affect the access to lexical-semantic information, we expect similar response times to the intact words and those with an extra macron sign (e.g., cheval ≈ chêval). This latter outcome would suggest that the reading cost found in Experiment 5 was dependent on the familiarity of the added diacritical marks. Further, we conducted post hoc analyses comparing the reading cost in those words where the diacritical mark was added to the letter e of the word and those pairs in which the diacritical mark was added to a vowel other than e.

Experiment 6 (Intact Versus Added [Nonexisting] Diacritic)

Method

Participants

We recruited an additional sample of 50 participants (16 women, mean age = 31.18 years [SD = 9.51]), using Prolific Academia and the same recruitment filters as in the previous experiments. All participants gave informed consent before participating in the study.

Materials

We used the words from Experiment 5 (see Table 7 and the Appendix). Each word was presented either intact (e.g., cheval) or with an added nonexistent diacritic (e.g., chêval). The diacritical marks were added to the same locations of the words as in Experiment 5. The stimulus lists were created in the same way as in the previous experiments.

Procedure and Data Analysis

They were the same as in Experiment 5.

Results and Discussion

Both Bayesian linear mixed-effects models converged successfully (R^s = 1.00). The mean RTs and percentage of errors are given in Table 9.

Response Time Analysis

Responses to words were faster when presented intact than when they were presented with an extra diacritical mark (614 vs. 627 ms, respectively; b = .04, SE = .01, 95% CrI [−.07, −.02]).

There was virtually no difference in response times between animal words and nonanimal words (620 ms for animal words vs. 621 ms for nonanimal words; b = .02, SE = .02, 95% CrI [−.06, .02]) and no interaction between the two factors (b = .01, SE = .02, 95% CrI [−.02, .05]). For the posterior distributions, see Figure 8.

Accuracy Analysis

We found a similar accuracy for accented and nonaccented words (3.0% of errors for nonaccented words vs. 2.8% of errors for accented words; b = .10, SE = .30, 95% CrI [−.67, .49]).

Participants made fewer errors in the nonanimal nouns than in the animal nouns (1.6% of errors for nonanimal words vs. 4.2% of errors for animal words; b = .87, SE = .32, 95% CrI [.27, 1.52]). There was no interaction between format and type of word (b = .00, SE = .02, 95% CrI [−.01, .03]).

Table 9

<table>
<thead>
<tr>
<th>Word</th>
<th>No diacritic</th>
<th>Additional diacritic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Response time</td>
<td>Error rate</td>
</tr>
<tr>
<td>Nonanimals</td>
<td>616</td>
<td>1.4</td>
</tr>
<tr>
<td>Animals</td>
<td>612</td>
<td>4.6</td>
</tr>
</tbody>
</table>

Note. See the online article for the color version of this figure.

Figure 7

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 French Words Correctly Written Without a Diacritic and With an Added Existing Diacritic (Experiment 5)
Diacritical marks in French need to be done to examine the subtle—
that further experiments with a direct comparison of the various
mark had no orthographic/phonological value. We acknowledge
present experiment is that unlike in Experiment 5, the diacritical
planation for this reduced reading cost for the letter e in the
present experiment is that unlike in Experiment 5, the diacritical
mark had no orthographic/phonological value. We acknowledge
that further experiments with a direct comparison of the various
diacritical marks in French need to be done to examine the suble-
ties of this comparison.

General Discussion

An often unnoticed but essential issue for any universal model
of visual-word recognition and reading in alphabetic orthographies
is how the word recognition system represents diacritical letters.
Recent research has suggested that the representation of diacritical
vowels may depend on their function in the language (see Figure
1); when they only indicate a suprasegmental value such as lexical
stress—as occurs in Spanish—diacritical and non-diacritical vow-
els would share their letter units; in contrast, when diacritics desig-
nate a change in phonology, as in German, diacritical and
non-diacritical vowels would be represented as separate letter units
(see Perea, Labusch, & Marcet, 2022). The main aim of the pres-
et research was to examine this issue in French. We chose French
because (a) it has a more extensive variety of diacritical marks
than Spanish or German (e.g., è, è, ñ); (b) the function of
these diacritical marks are quite diverse, unlike Spanish or Ger-
man; and (c) there is an influential computational model of word
recognition that assumes that diacritical and non-diacritical French
vowels are processed as separate letter units (Ans et al.,
1998; see also Chetail & Boursain, 2019, for a similar view).

We used a semantic categorization task (i.e., animal vs. nonani-
mal) for comparison purposes with recent studies conducted in
Spanish and German. We conducted three blocks of experiments.
In the first two blocks (Experiments 1–4), we selected a set of dia-
critical words (e.g., chèvre). The most important finding was
that omitting the word’s diacritical mark (or its replacement with
a non-existing diacritic) only produced a minimal non-reliable reading
cost (5–6 ms). Critically, this pattern is similar to that reported
with the omission of diacritics in Spanish (Perea, Labusch, & Mar-
cet, 2022). At the same time, we found a deleterious effect, rela-
tive to the intact words, when replacing the correct diacritical
mark of the diacritical word with a mismatching existing diacritic
(10–20 ms; Experiments 2–3). In the third block (Experiments 5–
6), we selected a set of nondiacritical words and added an existing
or non-existing diacritical mark (e.g., cheval in Experiment 5;
chèval, in Experiment 6). We found a reading cost of adding a
diacritical mark in both cases (Experiment 5: 19 ms; Experiment
6: 13 ms). We now discuss the implications of these findings for
models of visual word recognition.
First, the present experiments showed that, for diacritical French words, neither the omission of a diacritical mark nor its replacement with a nonexisting diacritical mark has a substantial, deleterious effect on the access to lexical-semantic information: chèvre and chêvre produced only slightly longer (around 5–6 ms, respectively) word identification times than the intact word chèvre. These findings pose doubts on the necessity of having twelve abstract letter units for the diacritical vowels in computational models of visual-word recognition in French, as in the Ans et al. (1998) model.

We believe that the most straightforward account of the above-described findings is that diacritical and nondiacritical vowels in French are not represented as entirely separate abstract letter identities. Instead, as suggested in another Romance language like Spanish (Marcet et al., 2021; Schwab, 2015), abstract letter units may be shared for diacritical and nondiacritical vowels. These conclusions align with observations of the daily usage of diacritics in native French speakers. Notably, diacritics are often omitted in computer-mediated French (see Anis, 2007). For instance, van Compernolle (2011) found that the percentage of omission of acute, circumflex, and grave diacritics in a forum of an Internet dating site was relatively high (10.9%, 31.2%, and 32.6%, respectively). Furthermore, in a recent poll among 2,500 French employers, 76% indicated that they struggled with written diacritical French words on a daily basis (Laffont, 2021). Likewise, it is not uncommon to find signs of stores and restaurants in French without the prescriptive diacritical mark even when written in uppercase (e.g., ÉPICERIE instead of ÉPICERIE [grocery store]). Although a parallel scenario occurs in Spanish (e.g., PANADERÍA instead of the prescriptive spelling PANADERÍA [bakery]), this is not the case in German with umlauts, where uppercase words keep their diacritical marks (e.g., BüCHEREI [library]). Thus, the present findings coincide with the observations of the everyday use of diacritics in French and Spanish.

The above reasoning does not imply that diacritical marks are treated as visual characteristics such as font, color, or size during lexical access. Unlike these perceptual elements, diacritics convey useful information during lexical access and have several functions in French (e.g., distinguishing homonyms, indicating phonological information, demonstrating the difference between past/present tense, or displaying the origin of the words). A sensible account, first proposed by Cubelli and Beschin (2005) in Italian, is that diacritical marks serve as orthographic cues during visual word identification—note that this idea is similar to that put forward by Peressotti et al. (2003) for the coding of the initial capitalization of proper nouns (see Sulzpio & Job, 2018, for electrophysiological evidence). The logic of Cubelli and Beschin’s proposal is that in parallel to processing the word’s abstract letter units, the word recognition system also encodes the information from the diacritical vowels. As a result, the orthographic cues for the diacritical vowels in chèvre would activate incompatible features that could cause interference when compared with the correct lexical entry chèvre, thus slowing down the access to lexical-semantic information. This way, this “orthographic cue” hypothesis can easily capture the reading cost of the words with mismatching diacritics in Experiments 2–3. The reading cost would be smaller when the diacritical information is omitted, or a nonexisting diacritic replaces it. Thus, the “orthographic cue” account can also capture the minimal reading cost observed in Experiments 1 and 4 (around 5–6 ms overall). Notably, this account can also explain the small but consistent reading cost of the words with the omitted diacritics in parallel experiments in Spanish, where there is no principled reason why diacritical and nondiacritical vowels would activate separate abstract letter units (e.g., 4 ms, Perea, Gómez, & Baciero, 2022; see also Marcet et al., 2021; Marcet & Perea, 2022; for a similar pattern).

To fully delineate the role of diacritical vowels in French, we also conducted two experiments with nondiacritical French words (e.g., cheval). In these experiments, items were presented intact or with an added diacritical mark, either existing (e.g., chèval, Experiment 5) or nonexisting (e.g., chêval, Experiment 6). The idea is that adding incompatible features to nondiacritical words could cause extra interference (either via bottom-up or lateral inhibition) during word recognition. In both cases, we found slower word response times for the words with an extra diacritic than for the intact words (e.g., cheval < chèval; cheval < chêval). Again, these findings can also be accommodated within the “orthographic cue” hypothesis proposed above. The logic is that adding incompatible features (i.e., the added diacritics) would induce some perceptual noise, making the visual input less similar to the word’s representation stored in the mental lexicon and hence, producing a reading cost relative to the intact words. The reading cost was slightly higher for the case where the diacritical mark existed in the language (19 vs. 13 ms, respectively). Notably, the reading cost with the added existing diacritic occurred to a larger degree when it involved a marker of vowel quality like ê or ô (e.g., words like chêval produced a greater cost over its corresponding intact word than words like loup). Thus, adding a familiar diacritical mark with unambiguous spelling (e.g., as in chèval) could involve an extra cost above the perceptual cost of adding an extraneous diacritic (see Perea, Gómez, & Baciero, 2022, for a small cost when adding diacritics in English words). In line with this, future research could evaluate the special role of diacritical marks in cases where French diacritics may play a stronger role during processing, for instance when additional diacritical marks indicate a change from the French present to past tense (e.g., chanté [sing] vs. chanté [sang]). Note that the current experiments have been done with nouns, and we intentionally excluded these ambiguous cases. Altogether, the present findings suggest that in French, not only the word’s perceptual features play a role in the processing of the diacritical vowels but also their linguistic properties.

Another remarkable pattern in the present experiments is that adding a feature to a given letter harms word processing more than removing it. Specifically, adding a diacritical mark to a nondiacritical French word hindered lexical access, whereas removing it from a diacritical word only produced a minimal cost. This asymmetry is not new; it goes back to Tversky (1977) and Treisman and Gormican (1988). Critically, this dissociation can readily capture the advantage of taper-TAPER over tâper-TAPER in masked priming in French (Chetail & Boursain, 2019) without resourcing to assuming separate letter presentations: the added diacritical mark on the letter â makes the percept less similar to a than vice versa. Thus, the general idea is that incompatible features may cause interference while removing a feature does not cause interference (see Kinoshita et al., 2021, for a noisy-channel model of perceptual asymmetries during letter recognition).
How are diacritical vowels represented in neural and computational models of visual word recognition? As we indicated in the Introduction, in the case of French, the Multiple-Trace Memory model (Ans et al., 1998) assumes separate letter units for every diacritical vowel in French (12 extra letter units). Therefore, this model would have predicted a sizable reading cost for chevre compared with chèvre, presumably with an effect size close to that reported in German (see Perea, Labusch, & Marcut, 2022). A more parsimonious account is to assume that diacritical vowels in French share their representations with their base letters. We acknowledge that the French vowels é and ê (i.e., the ones in which the diacritics involve an unambiguous grapheme-phoneme mapping) could enjoy a special role, but its close examination would go beyond the scope of this article. This option does not exclude that diacritical marks, in general, may serve as an orthographic cue to speed up lexical processing and help with the orthography-to-phonology mapping (see Cubelli & Beschin, 2005). Future computational models of visual word recognition should consider adding an orthographic marker for diacritical vowels in French instead of encoding the diacritical vowels as completely separate units at the letter level.

A similar explanation applies to Spanish and probably other languages (e.g., Italian) in which diacritics also have a suprasegmental role (i.e., lexical stress; see Protopapas & Gerakaki, 2009, for a similar observation in Greek). Importantly, the scenario is quite different for languages like German or Finnish, where there is a one-to-one correspondence between the diacritical vowels and their associated phoneme (e.g., ä/é/ vs. a/a in German). In this latter case, diacritical vowels would be represented as different letter units; it just happens that vowels like a and ä are visually very similar (e.g., as ü/j, or 0/Q; see Marcet & Perea, 2017, 2018; see also Perea et al., 2018, for a similar observation in Arabic).

In sum, our findings favor the view that diacritical and nondiacritical vowels share their abstract letter units in French. Indeed, omitting a word’s diacritical mark has little effect on the access to lexical-semantic information (e.g., chèvre ≈ chevre). More research is necessary to delineate the role of diacritics in French and other languages using techniques that may be more sensitive to the time course of the effects (e.g., sentence reading, event-related potentials) and across various populations (e.g., L1 and L2 learners).

References


(Appendix follows)
Appendix

Materials of the Experiments

Experiment 1 (Correct Diacritic Versus Omitted Diacritic)

Nonanimals

dièse, étendue, mélodie, orfèvre, sortilège, récit, abréviation, écrou, éclair, évier, dégustation, arête, diadème, débat, régisseur, légume, mécanisme, sénat, sphère, étui, purée, manège, brèche, étal, mosaïque, pédale, emblème, virée, décorateur, hémisphère, bénévole, vêtement, absurdité, légion, cuillère, guéridon, palmarès, stabilité, tanie, ébauche, vidéothèque, glacière, démarrage, téléviseur, préau, gîte, îlot, abécédaire, croisière, régulateur, étnicelle, contrée, képi, économie, gérant, symétrie, dégât, déodorant, cavité, flèche, festivité, kilomètre, étape, flûte, pédicure, météore, démontage, démographie, héroïsme, dragée, dôme, fève.

Animals

murène, mésange, hérisson, têtard, cacatoès, huître, caméléon, lièvre, araignée, bélier, écrevisse, chimpancé, panthère, étourneau, scarabée, mérou, goéland, éléphant, écreuil, crustacé, vipère, chev, âne, léopard, guépard, guêpe, épervier, lévrier, zébu, hyène, pelican, lézard, zèbre, héron, cai, rhinocéros.

Experiments 2, 3, and 4 (Correct Diacritic Versus Existing Diacritic That Changes Pronunciation [Exp. 2], Existing Diacritic That Does Not Necessarily Change Pronunciation [Exp. 3], and Nonexisting Diacritic [Exp. 4])

Nonanimals

étape, écrou, mélodie, légion, météore, hémisphère, diadème, pédicure, sénat, bouffée, brèche, chronomètre, péniche, cafété, bénévole, dragée, stabilité, économie, réglage, dièse, étendue, diversité, guéridon, palmarès, sphère, virée, étalage, étal, résidu, préau, décorateur, démographie, récit, cuillère, crèneau, démontage, régulateur, rivalité, abréviation, orfèvre, comète, réclamation, gérant, débat, évier, purée, contrée, glacière, démarrage, vidéothèque, éclair, cavité, flèche, sortilège, abécédaire, absurdité, pédale, légume, mécano, fève, régisseur, festivité, déodorant, céleri, képi, hélise, ébauche, épice, dégustation, étui, emblème, symétrie.

Animals

cacatoès, émeu, pelican, chimpancé, lévrier, épervier, panthère, lézard, crustacé, scarabée, mérou, rhinocéros, mésange, zèbre, zébu, lémurien, héron, chèvre, murène, écrevisse, éléphant, flétan, léopar, guépar, coléoptère, caméléon, hérisson, bélier, vipère, araignée, étourneau, hyène, lièvre, écureuil, goéland.

Experiments 5 and 6 (No Diacritic Versus Additional Diacritic [Existing Diacritic in Experiment 5; Nonexisting Diacritic in Experiment 6])

Nonanimals

clavier, piscine, flamme, brume, libraire, galerie, bonheur, savon, poubelle, piquet, cloison, raquette, casino, bassin, manche, allergie, foulard, veste, gobelet, embalage, bus, remous, escale, colline, glaise, science, extrait, cante, di, magasin, carte, plage, tapis, bandeau, jupe, table, triage, chapitre, accent, magie, moniteur, demain, pollution, plateau, baguette, hoquet, four, chemin, pilotage, fromage, optimisme, tisane, carnet, cartouche, rideau, billet, bruit, disque, chapeau, grammaire, repos, cahier, matelas, futur, papier, camion, bille, couchette, pochette, espoir, ascenseur.

Animals

mouche, chien, canari, chat, cheval, renard, mouette, poule, lapin, poisson, sanglier, moineau, cochin, limace, loup, requin, singe, tortue, papillon, dauphin, chameau, aigle, cigogne, perroquet, aigle, canard, vache, phoque, girafe, souris, rat, baleine, moustique, crocodile, cygne.

Received January 5, 2022
Revision received November 7, 2022
Accepted November 7, 2022