Labusch, M., Perea, M., & Hyönä, J. (in press). The Costs of Adding versus Omitting Diacritics in Visual Word Recognition: Evidence from German and Finnish. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. https://doi.org/10.1037/xlm0001586

The Costs of Adding versus Omitting Diacritics in Visual Word Recognition: Evidence from German and Finnish

Melanie Labusch [1,2,3], Manuel Perea [1,2], and Jukka Hyönä [4]

[1] Department of Methodology and ERI-Lectura, Universitat de València, Valencia, Spain

[2] Centro de Investigación Nebrija en Cognición, Universidad Nebrija, Madrid, Spain

[3] Health Research Institute La Fe, Valencia, Spain

[4] Psychology Department, University of Turku, Turku, Finland

Short title: Addition vs. Omission of Diacritics

The data, scripts, and output of the three experiments are available at

https://osf.io/jtda8

Address for Correspondence:
Manuel Perea
Departamento de Metodología
Av. Blasco Ibáñez, 21
46010-Valencia (Spain)
mperea@uv.es

Abstract

In German, diacritical marks distinguish between vowel sounds in print (e.g., “o” pronounced /o/ versus “ö” pronounced /ø/). Unlike in Spanish, where diacritics primarily indicate lexical stress, omitting diacritics in German (e.g., *Kröte* [toad] → *Krote*) leads to longer word identification times compared to intact words. This suggests separate letter representations for diacritical and non-diacritical vowels in German. Current models of visual word recognition assume distinct letter representations for diacritical and non-diacritical vowels in German (Ziegler et al., 2000), but it remains unclear whether the reading cost differs when a diacritic is added versus omitted. We conducted three semantic categorization experiments to examine whether the presence of an added diacritic in a non-diacritical word (e.g., *Schwan* [swan] → *Schwän*) incurs a greater lexical-semantic cost than its omission (e.g., *Kröte* → *Krote*) in German and Finnish, another language where diacritical vowels signal distinct pronunciations. In noisy-channel models, adding a diacritic makes the percept less similar to the base word than omitting one, thus predicting a larger cost. In contrast, abstractionist models assume rapid activation of abstract letter representations, predicting a negligible asymmetry. Results were similar in German and Finnish. First, both types of misspellings showed a reading cost relative to the intact words. Second, the reading cost was larger for the addition than for the omission of diacritics, placing new constraints on the orthographic front-end of models of visual word recognition.

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

Introduction

Models of visual word recognition typically assume a hierarchy of detectors, from low-level visual features (e.g., lines, curves) to abstract letter identities and whole-word forms (e.g., Coltheart et al., 2001; Davis, 2010; Dehaene et al., 2005; see Grainger, 2018, for review). These models, often developed with English in mind, assume that each of the 26 Latin letters has a single case-invariant abstract representation (e.g., "a" and "A" would correspond to the same unit). However, they rarely consider diacritical letters, which are common in most Latin-based orthographies. This omission raises a fundamental question for models of reading: what counts as a letter? (Grainger, 2024; Hofstadter, 1985). As Hofstadter (1985) asked in the *Letter Spirit* project when questioning what makes an "a" an "a", we might now ask: is "ä" simply another form of "a", or a different letter altogether?

Diacritical marks serve different functions across Latin-based orthographies (see Labusch et al., in press; Wells, 2000, for review). In Spanish, they indicate lexical stress without changing phoneme identity (e.g., "a" and "á" are pronounced /a/) and are assumed to share a single abstract letter representation (Chetail & Boursain, 2019; Perea et al., 2020; see Conrad et al., 2010, for modeling). In contrast, in German, diacritics mark distinct phonemes (e.g., a /a/ vs. ä /ɛ/, o /o/ vs. ö /ø/, u /u/ vs. ü /y) and are modeled as separate abstract letter units (Hutzler et al., 2004; Ziegler et al., 2000). The same logic plausibly extends to other languages in which diacritics mark vowel contrasts (e.g., Finnish: a /ɑ/ vs. ä /æ/, o /o/ vs. ö /ø/). Supporting this view, Perea et al. (2022b), using semantic categorization tasks, found a sizeable reading cost when diacritics were omitted in German (*Kröte* [toad] → *Krote*), but only a minimal cost in Spanish (*ratón* [mouse] → *raton*). To our knowledge, no previous study has directly compared the cost of adding versus omitting a diacritic in words, nor examined whether these asymmetries generalize to another orthography in which diacritics mark phonemic distinctions.

The present experiments address this issue by testing whether adding a diacritic (*Schwan* [swan] → *Schwän*) is more disruptive than omitting one (*Kröte* [toad] → *Krote*) in a single-word recognition task requiring lexical-semantic access (i.e., semantic categorization: "does the word refer to an animal?"). This manipulation addresses a key issue in visual word recognition: how changes in letter features, instantiated here by the addition or omission of diacritics, map onto abstract representations during lexical access. While current models of visual word recognition assume that diacritical and non-diacritical vowels correspond to distinct representations in languages like German and Finnish (see Labusch et al., in press), it remains unclear whether the direction of the manipulation influences the reading cost. Given the functional similarity between the two languages, the key question is whether adding a diacritic incurs a larger cost than omitting one, or whether both manipulations yield comparable disruption.

To interpret this potential asymmetry, we contrast two families of models. Abstractionist models assume that perceptual variation is rapidly overridden by mappings onto abstract letter representations that fuel lexical access (Dehaene et al., 2005; Grainger et al., 2008). On this view, surface form is mapped onto identity-level codes, and variation due to case or font is neutralized early in processing (e.g., *chair*, `chair`, *CHAIR*). Supporting this, Chauncey et al. (2008) reported that masked repetition priming was equally strong regardless of font or size. ERP studies have similarly found that identity priming effects from the N250 onward are unaffected by letter case (Vergara-Martínez et al., 2015), and behavioral responses are likewise invariant (Jacobs et al., 1995; Perea et al., 2015; see also Laham & Leth-Steensen, 2023; Perea et al., 2020). Extending this logic to diacritics, if lexical access is governed primarily by abstract letter codes, adding and omitting a diacritic should incur comparable costs (e.g., *Schwan* vs. *Schwän*; *Kröte* vs. *Krote*), because early perceptual differences are rapidly regularized.

In contrast, noisy-channel models of word recognition (Norris & Kinoshita, 2012) propose that visual word recognition involves inference over a noisy perceptual input: features or letters may be lost, added, substituted, or transposed before abstract representations are activated. This account explains why pseudowords like *JUGDE* or *TELEPHOME* can be misread as *JUDGE* or *TELEPHONE* (O'Connor & Forster, 1981; Perea & Lupker, 2004). In its original formulation, the Bayesian Reader model assumed comparable costs for insertions and deletions of visual features of letters. [Footnote 1] However, Kinoshita et al. (2021) revised this model by proposing that insertions are more disruptive than deletions. The logic was that, during feature-matching in early letter perception, deletions create partial matches (no discrepant features), while additions introduce mismatches (Kinoshita et al., 2021). They justified this assumption by stating that "visual features are more likely lost than inserted", a characteristic they tied to the processes of "early vision" (p. 821). Indeed, in masked priming experiments with isolated letters, deleting a feature was less disruptive than adding one (e.g., in English letter pairs like E-F; Kinoshita et al., 2021; see also a similar asymmetry in visual search experiments, Treisman, 1986). Consequently, the revised noisy-channel model predicts greater disruption for diacritic addition (*Schwan* → *Schwän*) than omission (*Kröte* → *Krote*), reflecting the influence of perceptual features during lexical access.

Notably, a number of recent experiments have examined this asymmetry during the initial stages of word processing using the masked priming paradigm (Forster & Davis, 1984; see Grainger, 2008, for a review). Perea et al. (2020) found that in Spanish, for a diacritical target word like *FÁCIL* [easy], both *fácil* and *facil* primes were equally effective. In contrast, for a non-diacritical target like *FELIZ* [happy], the prime *féliz* was less effective than *feliz*, even though both a/á and e/é correspond to the same phonemes (/a/ and /e/, respectively; see also Perea et al., 2021). This asymmetry also emerges in languages where diacritical and non-diacritical vowels correspond to different phonemes. For instance, in Hungarian, the prime *néma* was as effective as *nema* for the target *NÉMA* [mute], but *mése* was less effective than *mese* for the target *MESE* [tale]—note that "e" corresponds to /ɛ/ and "é" to /eː/ (Benyhe et al., 2023; see also Perea et al., 2022a). A similar pattern is found for consonants marking phonemic contrasts. In Spanish, for example, *muneca–MUÑECA* [doll] primes as effectively as *muñeca–MUÑECA*, but *moñeda–MONEDA* [coin] primes less effectively than *moneda–MONEDA*, where "n" and "ñ" correspond to /n/ and /ɲ/, respectively (Marcet et al., 2020). [Footnote 2] While the masked priming technique provides valuable insights into early processing, its reliance on prime–target coactivation prevents a clear account of the dynamics that unfold upon presentation of a word (see Andrews, 1997; Gomez et al., 2021, for discussion). To address this issue, we used a single-presentation paradigm to examine the processing cost of omitting or adding diacritics in lexical access.

Empirical evidence for perceptual asymmetries between diacritical and non-diacritical vowels using single-presentation word recognition experiments remains scarce and somewhat inconclusive. As noted earlier, in Spanish, where diacritics do not alter phoneme identity, omitting a diacritical mark incurs only a minimal cost (approximately 4 ms; e.g., *ratón* [mouse] vs. *raton*; Perea et al., 2022b), while adding a diacritic produces a slightly larger cost (6–7 ms; e.g., *perro* [dog] vs. *perró*; Labusch et al., 2022). This asymmetry is consistent with the idea that prosodic diacritics carry limited perceptual weight during lexical access. A simple letter-frequency account cannot explain the present pattern; although diacritical characters are generally less frequent than their base forms (e.g., á vs a in Spanish), Spanish shows only minimal costs for both omission and addition. In contrast, French offers a more complex case: although some diacritical vowels do map onto distinct phonemes (e.g., é /e/ vs. è /ɛ/), the omission of diacritics still produces only a small cost (5–6 ms; e.g., *chèvre* [goat] vs. *chevre*), while the addition of a diacritic leads to a more substantial cost (19 ms; e.g., *chèval* vs. *cheval* [horse]; Labusch et al., 2023). However, the interpretation of these findings is complicated by the frequent omission of diacritics in informal French writing, meaning that readers are used to encountering words with omitted diacritics (e.g., *chevre*).

A stronger test of perceptual asymmetries in word recognition requires languages such as German or Finnish, where diacritical and non-diacritical vowels map onto different phonemes and are treated as distinct letters in both instruction and orthographic conventions (e.g., reading acquisition, keyboard layouts). Crucially, diacritics in these languages are spelled consistently, regardless of register (see Perea et al., 2022b, for a sizeable reading cost of omitting diacritics in German; Footnote 3). An additional methodological advance of the present experiments is the use of a within-subject design to directly compare the effects of the omission and addition of diacritics. Unlike prior studies in Spanish and French, which tested each manipulation in separate experiments, our approach allowed the same participants to receive both manipulations (omission: *Kröte* → *Krote*; addition: *Schwan* → *Schwän*). Our design therefore provides a direct test of whether adding diacritics incurs greater costs than omitting them.

In sum, we conducted three semantic categorization experiments to test whether adding a diacritic to a non-diacritical word (e.g., *Schwan* → *Schwän*) incurs greater reading costs than omitting a diacritic from a diacritical word (e.g., *Kröte* → *Krote*) in languages where diacritics encode phonemic contrasts. We focused on German and Finnish, two orthographies in which diacritical marks correspond to distinct phonemes and are treated as separate letters (Ziegler et al., 2000). Across all experiments, participants performed a semantic categorization task (i.e., deciding whether a word denoted an animal), ensuring comparability with prior work (e.g., Labusch et al., 2023; Perea et al., 2022b). Abstractionist models, which map surface variation onto abstract letter identities, predict comparable costs for addition and omission. In contrast, noisy-channel accounts predict larger costs for addition, since added features introduce mismatches whereas deletions yield partial matches. We first tested this prediction in German (Experiment 1) and then replicated it in Finnish with blocked (Experiment 2) and randomized (Experiment 3) presentations.

Experiment 1 (Semantic Categorization in German)

Methods

Participants

We recruited a sample of 76 native German participants (22 self-identifying as women; mean age = 29.7 years, SD = 11.5) through Prolific Academic (www.prolific.co). This sample size allowed us to have more than 2,000 observations for each critical condition (76 participants x 30 items = 2,280), thus in line with the suggestions of Brysbaert and Stevens (2018). We used Prolific's participant filter to ensure that the participants were native German speakers, had normal or corrected-to-normal vision, and had no reading disorders. Before the experiment, all participants provided informed consent and received monetary compensation following Prolific's standard hourly rate. The experiments were approved by the Research Ethics Committee of the University of Valencia and adhered to the guidelines set forth by the Declaration of Helsinki.

Materials

We selected 120 German words from the German lexical database of Oganian et al. (2016), of which 40 were animal names, and 80 did not refer to animals or concepts related to animals or other living beings. Half of the words contained at least one letter with a diacritic, and the other half contained none. For the words with diacritical letters, 20 were animal names (mean word length = 7.65; mean word frequency per million = 2.9; mean orthographic Levenshtein distance of the 20 closest neighbors [OLD20, see Yarkoni et al., 2008] = 2.8) and 40 were non-animal words (mean word length = 7.55; mean word frequency = 3.1; mean OLD20 = 2.9). For the words without diacritical letters, 20 were animal words (mean word length = 7.6; mean word frequency = 3.1; mean OLD20 = 2.8), and 40 were non-animal words (mean word length = 7.5; mean word frequency = 2.7; mean OLD20 = 2.7). The four groups were matched on word length, word frequency per million, and OLD20 (all *p*s > .20). The materials are presented in Appendix 1.

Each word was presented either intact or with an added/removed diacritical mark. For diacritical words, the diacritical mark was removed (e.g., *Kröte* [toad] → *Krote*). In contrast, for the non-diacritical words, a diacritical mark was added to a vowel (e.g., Schwan [swan] → Schwän). None of the altered forms created an existing word in German (e.g., the word *schon* [already] could not be selected because *schön* is a word [beautiful]). The words were presented in two blocks: one block contained the non-diacritical base words (in which half of them were presented with an additional diacritic), and the other block contained the diacritical base words (in which half of the words were presented with an omitted diacritical mark). The order of the blocks was counterbalanced across participants, and each participant saw each word only once, resulting in four different lists of stimuli. Participants were assigned randomly to a list. Each block contained a brief practice phase.

Please Insert Table 1 Around Here

Procedure

The experiment was programmed using PsychoPy 3 (Peirce & MacAskill, 2018) and hosted on their online server Pavlovia (www.pavlovia.org). Participants were first required to fill out a demographic questionnaire hosted on www.limesurvey.org. Participants were instructed to do the experiment in a quiet environment with minimal distractions. They were asked to categorize German words as “animal” or “non-animal” while ignoring diacritics, pressing the “m” key for animal names and “x” key otherwise, responding as quickly and accurately as possible. Each block began with 16 practice trials.

On each trial, a fixation cross appeared for 500 ms, followed by the stimulus for 2,000 ms or until response. Trials were presented in a randomized order for each participant. Reaction times (RTs) and accuracy were measured. The experiment lasted approximately 8-10 minutes, with a short break after every 80 trials.

Data analysis

Statistical analyses were conducted using Bayesian linear mixed-effects models implemented in brms (Bürkner, 2017) in R (R Core Team, 2023), with Stan as backend (Stan Development Team, 2023). [Footnote 4] Bayesian models provide full posterior estimation and robust handling of mixed-effects models (Schad et al., 2021; Vasishth et al., 2018). We specified maximal random-effects structures, as recommended by Barr et al. (2013). We modeled three factors using effect coding: word type (diacritical = -0.5, non-diacritical = 0.5), spelling (correct = -0.5, misspelled = 0.5), and semantic category (animal [-2/3] vs. non-animal [1/3]); centered to reflect the 1:3 ratio). Dependent variables (RT and accuracy) were analyzed separately. For RTs, we used an exGaussian distribution, as it captures well the shape of RT distributions (Balota & Yap, 2011; Ratcliff, 1979); for accuracy, a Bernoulli distribution. Each model ran four chains for 10,000 iterations (2,000 iterations were employed as a warm-up). The maximal random-effect structure allowed by the design (Barr et al., 2013) was the following:

```
Dependent Variable = word_type*spelling*category +
                     (1+word_type*spelling*category|subject) +
                     (1+spelling|item)
```

We used the default weakly informative priors from *brms* (Bürkner, 2017; Scholz & Bürkner, 2023). Following a reviewer's suggestion, we examined alternative priors for fixed effects (i.e., *N*(0,50), *N*(0,100), *N*(0,150)). Posterior estimates were nearly identical. Given the large dataset, prior choice had little influence on the posterior distributions (see OSF link for fits with these other priors; see also Scholz & Bürkner, 2023). Model outputs provide a coefficient for each parameter (the mean of the posterior distribution) posterior SDs, and 95% Credible Intervals (CrIs). We considered effects supported when the 95% CrI did not intersect zero.

The experiment was not pre-registered. Data, scripts, and full output of this and the following experiments are available at:

https://osf.io/jtda8

Results and Discussion

We excluded very short trials (<250 ms; 0 cases) and incorrect responses (3.5%) from the RT analyses. Trials with no response within the 2000-ms deadline were coded as errors. The mean response times and error rates are displayed in Table 2. The RT and accuracy models produced good fits. All chains converged (all $\hat{R}$ = 1.00), with bulk and tail effective sample sizes exceeding 2,000. The posterior distributions for all parameters, including interactions, are presented in Figure 1. The population-level estimates of the RT models for all experiments are presented in Appendix 2.

Please Insert Table 2 Around Here

Response Time Analysis. As expected, response times were longer for incorrectly spelled words than intact words (*b* = 16.93, *Est.Error* = 2.16, *95% CrI* [12.67, 21.17]). Critically, this reading cost was larger when a diacritical mark was added than when it was omitted (word type x spelling interaction: *b* = 8.89, *Est.Error* = 3.96, *95% CrI* [1.15, 16.71]). Finally, we observed a three-way interaction among word type, spelling, and semantic category (*b* = 24.57, *Est.Error* = 9.31, *95% CrI* [6.12, 42.63]). This interaction reflected that misspelling cost for addition vs. omission of diacritics was larger for non-animal words (39 ms vs. 25 ms, respectively) than for animal words (27 vs. 17 ms; see Figure 1, left panel, for the posterior distributions).

Accuracy Analysis. Accuracy was lower for animal than non-animal words (*b* = 0.93, *Est.Error* = 0.24, *95% CrI* [0.45, 1.40]). We also found evidence for a three-way interaction (*b* = -1.27, *Est.Error* = 0.65, *95% CrI* [-2.55, -0.02]), indicating that the lower accuracy for animal words occurred mainly in the misspelled condition (see Figure 1, right panel).

Please Insert Figure 1 Around Here

The present experiment demonstrated sizeable reading costs in RTs both when a diacritic was omitted from a German word (e.g., *Kröte → Krote*) and when a diacritic was added to a non-diacritical word (e.g., *Schwan → Schwän*). This replicates the omission cost reported by Perea et al. (2022b) with a different item set and further demonstrates that addition also incurs a cost. These findings reinforce the proposal that diacritical and non-diacritical vowels are represented as distinct abstract units in German (Ziegler et al., 2000).

More importantly, adding a diacritic was 10–14 ms more costly than omitting one, revealing a clear asymmetry. This result is consistent with the prediction of noisy-channel models (Kinoshita et al., 2021), which proposed that feature insertions are more disruptive than deletions. In contrast, the abstractionist view, which predicts comparable costs for addition and omission, was not supported. The asymmetry observed here parallels findings in French for the letter "e" (Labusch et al., 2023) but extends them using a within-subject design and a broader set of vowels.

Accuracy remained high, indicating that readers could readily recover base words despite minimal costs for the omission and addition of diacritics. This is consistent with earlier evidence that readers can reconstruct words from visually similar letter strings (Rayner & Kaiser, 1975). In addition, the three-way interaction in RTs suggested that the asymmetry was numerically larger for non-animal than animal words. However, given that this effect was small and not consistently found in prior experiments (e.g., Labusch et al., 2023), we prefer to interpret it cautiously.

Having established an asymmetry in German, where adding a diacritic was more disruptive than omitting one, we next tested whether this pattern generalizes to Finnish, another transparent orthography with distinct grapheme–phoneme mappings for diacritical vowels. Experiment 2 had two goals: (i) to assess whether Finnish also shows reading costs for diacritic addition and omission, and (ii) to determine whether the asymmetry observed in German generalizes to another language in which diacritics mark separate phonemes.

Importantly, Finnish differs from German in that diacritical vowels participate in vowel harmony, a phonological constraint whereby front vowels (ä, ö, y) and back vowels (a, o, u) typically do not co-occur within a morpheme (except when mediated by neutral vowels e and i). Vowel harmony plays a robust role in spoken word recognition (Suomi et al., 1997; Tuomainen, 2001) and in morphological parsing during reading (Bertram et al., 2004). Its role in visual word recognition, however, remains less clear. Although disharmony reduces wordlikeness in pseudowords (Perea et al., 2022a), its effects on words are less straightforward. Indeed, disharmonious misspellings may trigger two opposing mechanisms during lexical access. On the one hand, their anomalous pattern (i.e., violating Finnish vowel harmony) may slow down lexical processing due to the detection of an irregularity. On the other hand, because disharmonious misspellings produce a phonotactically illegal phonological representation, they may be more easily normalized to the base word (see Rumelhart, 1985, for the dynamics underlying this normalization process with orthotactically illegal strings). Notably, recent evidence from Turkish, another language with vowel harmony, suggests that vowel harmony exerts only weak effects on word processing (Özkan et al., 2025).

Experiment 2 (Semantic Categorization in Finnish)

Methods

Participants

We recruited 56 native Finnish participants (26 women, mean age = 32 years, SD = 8.9 years) through Prolific Academic with the same filters as in Experiment 1, restricted to native Finnish speakers. This sample size was chosen to yield a comparable number of observations as in Experiment 1 (i.e., 56 x 36 = 2,016 observations).

Materials

We selected 144 Finnish words from the Word-Mill database (Laine & Virtanen, 1999): 46 animal names and 98 words that did not denote animals or concepts related to living beings. Half of the words contained at least one letter with a diacritic, and half did not. The four sets: animal words with diacritics (mean length = 7.4; mean frequency = 4.8); animal words without diacritics (mean length = 7.6; mean frequency = 4.8); non-animal words with diacritics (mean length = 7.2; mean frequency = 4.7), and non-animal words without diacritics (mean length = 7.4; mean frequency = 4.7) were matched on length and frequency (all *p*s > .20). The complete list is in Appendix 1. The manipulation paralleled Experiment 1: diacritics could be removed from diacritical words (e.g., *lehmä* [cow] → *lehma*) or added to non-diacritical words (e.g., *varis* [crow] → *väris*).

Procedure

The procedure was the same as in Experiment 1, except that (i) all stimuli were in Finnish and (ii) participants pressed the “m” key for animal names and the “z” key for non-animal words.

Data Analysis

Data were analyzed as in Experiment 1 using Bayesian linear mixed-effects models in brms.

Results and Discussion

We excluded very short trials (<250 ms; 0 cases) and incorrect responses (5.0%) from RT analyses. The mean RTs and error rates appear in Table 3. All chains converged (all $\hat{R} = 1.00$), with bulk and tail effective sample sizes exceeding 1,800.

Please Insert Table 3 Around Here

Response Time Analysis. Misspelled words yielded longer response times than intact words (*b* = 15.32, *Est.Error* = 2.20, 95% *CrI* [11.04, 19.68]). Response times were also longer for non-diacritical than diacritical words (*b* = 14.59, *Est.Error* = 7.08, *95% CrI* [0.75, 28.48]) and for animal than non-animal words (*b* = -13.97, *Est.Error* = 7.07, *95% CrI* [-27.62, -0.03]). More importantly, as in Experiment 1, the misspelling cost was larger for the addition than for the omission of diacritics (word type x spelling: *b* = 12.59, *Est.Error* = 4.17, 95% *CrI* [4.38, 20.64]). This asymmetry was comparable for animal and non-animal words (three-way interaction: *b* = 7.98, *Est.Error* = 8.73, 95% *CrI* [-9.14, 25.15]).

Accuracy Analysis. Accuracy was lower for animal than non-animal words (*b* = 1.29, *Est.Error* = 0.27, *95% CrI* [0.76, 1.83]). No other reliable effects emerged (see Figure 2).

Please Insert Figure 2 Around Here

Experiment 2 replicated and extended the findings of Experiment 1 to Finnish. Both addition and omission of diacritics produced reading costs, and the addition cost (e.g., *varis* → *väris*) exceeded the omission cost (e.g., *lehmä* → *lehma*) by approximately 13–16 ms. This asymmetry supports noisy-channel accounts (Kinoshita et al., 2021), which predict insertions to be more disruptive than deletions, and challenges the abstractionist prediction of comparable costs.

In an exploratory analysis, we also tested whether vowel harmony influenced processing. Harmonious and disharmonious misspellings differed by less than 2 ms, indicating negligible impact. Thus, vowel harmony appears secondary in semantic categorization, consistent with findings from Turkish (Özkan et al., 2025). While vowel harmony may affect morphological parsing (Bertram et al., 2004), its contribution to single-word semantic access appears limited.

In summary, the present experiment in Finnish replicated the asymmetry obtained in the German experiment: adding diacritics was more disruptive than omitting them, reinforcing the generality of this pattern across languages where diacritical vowels mark phonemic contrasts.

One remaining issue is whether the asymmetry might be magnified in a randomized design. In Experiments 1 and 2, words were blocked by base type (whether with or without diacritics), which could have allowed participants to adapt their attention toward the omission or addition of diacritics. A fully randomized design would prevent anticipation of diacritic status and could amplify the asymmetry. Alternatively, if the asymmetry arises at early visual encoding, as suggested by masked priming and visual search studies (Benyhe et al., 2023; Kinoshita et al., 2021; Treisman, 1986), then blocking should not play a role, and a similar pattern should emerge. Experiment 3 tested these alternatives.

Experiment 3 (Semantic Categorization in Finnish without blocks)

Methods

Participants

We recruited a new sample of 56 native Finnish participants (25 women, mean age = 31.5 years, SD = 7.8 years) through Prolific Academic, using the same selection criteria as in

Experiment 2. This sample size was chosen to yield the same number of observations as in the previous experiment (i.e., 56 x 36 = 2,016 observations).

Materials

The materials were identical to those in Experiment 2.

Procedure

The procedure was the same as in Experiment 2, except that the order of trials was fully randomized rather than blocked by word type.

Data Analysis

Analyses were conducted as in Experiment 2.

Results and Discussion

As in the previous experiments, we excluded very short trials (<250 ms; 0 cases) and incorrect responses (4.6%) from the RT analyses. Mean RTs and error rates are displayed in Table 4. Model fits were very good (all $\hat{R}$s = 1.00; bulk and tail effective sample sizes exceeding 1,750).

Please Insert Table 4 Around Here

Response time analyses. Misspelled words produced slower responses than intact words (b = 23.90, Est.Error = 2.85, 95% CrI [18.38, 29.56]). Response times were also slower for animal than for non-animal words (b = -27.56, Est.Error = 8.87, 95% CrI [-44.93, -10.20]). As in Experiments 1–2, the cost of adding a diacritic was larger than the cost of omitting one (43 vs. 25 ms). The two-way interaction supported this pattern (b = 9.72, Est.Error = 5.06, 95% CrI [-0.16, 19.74]). Although this credible interval narrowly included zero, the vast majority of the posterior distribution was positive ($P(\beta > 0)$ = .97; see Footnote 5 for a complementary analysis). There was no clear evidence for a three-way interaction (*b* = 6.9, Est. Error = 10.8, 95% CrI [–14.2, 28.3]; see Figure 3, right panel).

Accuracy analyses. Accuracy was again lower for animal than non-animal words (*b* = –1.14, Est. Error = 0.33, 95% CrI [–1.80, –0.51]). No other effects were supported.

Please Insert Figure 3 Around Here

With a fully randomized presentation, Experiment 3 replicated the asymmetry from Experiments 1–2. Adding diacritics produced a larger reading cost than removing them (approximately 16–17 ms), closely matching the asymmetry in Experiment 2 (approximately 13–16 ms; see Figures 2 and 3). Thus, the asymmetry is robust to experimental design, occurring both in blocked and randomized presentations. This invariance suggests that the effect is unlikely to reflect strategic adaptation to the structure of the blocks. Instead, the pattern indicates that the asymmetry arises during early perceptual encoding, in line with noisy-channel accounts where insertions are more disruptive than deletions (Kinoshita et al., 2021).

General Discussion

We conducted three semantic categorization experiments to examine (i) whether the addition or omission of diacritics incurs a reading cost relative to intact words in two transparent orthographies from different families (German [Germanic] and Finnish [Uralic]) that use diacritics to mark phonemic contrasts, and (ii) whether this cost is asymmetrical. Across both languages, misspellings slowed responses relative to intact words, and, crucially, adding a diacritic was consistently more disruptive than omitting one. (footnote 6) These findings support the view that, in languages like German and Finnish, diacritical and non-diacritical vowels are represented as distinct letter units, and that feature-level differences at the orthographic front end continue to shape lexical access.

<u>Cross-linguistic implications</u>

At the language-specific level, our findings reinforce that diacritical vowels are represented as distinct letter units in German and Finnish. This assumption has long been present in models of visual word recognition in German (e.g., Hutzler et al., 2004; Ziegler et al., 2000), and our results extend it to Finnish: vowels such as ä and ö must be treated as separate graphemes from a and o, despite their visual similarity. In other words, they are encoded as different letters, much like conventional pairs with high feature overlap (e.g., i–j, n–m). By extension, the same principle likely applies to other orthographies where diacritics mark phonemic contrasts, such as Hungarian (Benyhe et al., 2023).

In contrast, languages in which diacritics primarily mark prosodic features such as stress (e.g., Spanish, Italian) show a different pattern: in Spanish, “á” and “a” correspond to the same phoneme /a/, and evidence from masked priming and single-presentation studies suggests they share a single abstract representation (Chetail & Boursain, 2019; Marcet & Perea, 2022). Omission of the accent has only a minimal cost (approximately 4 ms), while addition produces a slightly larger but still small cost (approximately 6–7 ms; Labusch et al., 2022; Perea et al., 2022b). Thus, in stress-marking languages, diacritics carry relatively little weight in visual word recognition.

French occupies an intermediate position. Some diacritical vowels correspond to distinct phonemes (é vs. è), while others do not. Behavioral evidence shows small omission costs (5–6 ms; *chèvre* [goat] vs. *chevre*) but larger addition costs (up to 19 ms; *chèval* vs. *cheval* [horse], Labusch et al., 2023). The larger addition costs may partly reflect orthographic conventions, since readers frequently encounter omission of diacritics in informal French writing (*chevre*), but rarely encounter additions of diacritics (*chèval*). Thus, French illustrates how orthographic conventions (optional vs. obligatory diacritics) shape perceptual processing.

Taken together, additions of diacritics are more disruptive than omissions across orthographies. The magnitude varies with function: strongest when they mark phonemic contrasts (German, Finnish), weaker when they mark prosody (Spanish), and intermediate when omission of diacritics is frequent in informal writing (French). We next consider how these findings inform models of visual word recognition

Theoretical implications

Our findings speak directly to competing theoretical accounts of visual word recognition. On the one hand, abstractionist models such as the Local Combination Detectors model (Dehaene et al., 2005; see also Grainger et al., 2008) propose that perceptual variations in orthographic input (e.g., case, font, color, size) are rapidly normalized during early processing stages. In German and Finnish, however, diacritics mark distinct phonemes and therefore should correspond to different abstract letter units (see Ziegler et al., 2000). On this view, addition and omission of diacritics should yield comparable costs, since “a” and “ä” activate different letter units. Our data contradict this: we found asymmetric costs in both languages, with the addition of diacritics consistently producing greater disruption than their omission. Thus, this asymmetry challenges the assumption of rapid normalization for letters with diacritics in phoneme-marking orthographies.

On the other hand, noisy-channel models provide a more natural fit. In the Bayesian Reader model (Norris & Kinoshita, 2012), visual word recognition is treated as inference from noisy input: the system combines the likelihood of the observed features with priors about common error types (e.g., substitutions, transpositions, insertions, deletions). In the original account, insertions and deletions were assumed to be equally probable, predicting symmetric costs for diacritical versus non-diacritical letters. More recently, however, Kinoshita et al. (2021) argued that features are generally more likely to be lost than spuriously added, due to factors such as attentional filtering or peripheral degradation. On this view, additions are more costly than omissions (e.g., *Schwan → Schwän* vs. *Kröte → Krote*), consistent with the 12–18 ms asymmetry observed in German and Finnish. Importantly, this asymmetric pattern may depend on language-specific experience: in orthographies where diacritics mark prosodic features and are often omitted in informal writing (e.g., Spanish, French), the system may treat missing diacritics as more probable, yielding very small differences between addition and omission. In contrast, in phoneme-marking systems such as German and Finnish, added diacritics more strongly violate expectations, increasing the reading cost. Current noisy-channel models underspecify how visual features map onto letter representations (see Snell, 2025), an issue that remains to be addressed in future work.

Finally, a third perspective is offered by interactive activation models (McClelland & Rumelhart, 1981; see also Rey et al., 2009), which integrate bottom-up activation from letter features, lateral inhibition at the letter level, and top-down feedback from lexical representations. In these frameworks, diacritical and non-diacritical vowels in German and Finnish are encoded as distinct letter representations (e.g., a vs ä; see Ziegler et al., 2000). The addition of a diacritic (e.g., *Schwan → Schwän*) activates the wrong node (“ä”), triggering inhibition of the correct letter unit (“a”). In contrast, omitting a diacritic (e.g., Kröte → Krote) activates a visually similar node (“o”), while the intended target letter (“ö”) still receives partial bottom-up support from overlapping features, attenuating the cost. This asymmetry follows naturally from feature-to-letter dynamics and aligns with the observed reading cost. Cross-linguistic patterns fit as well: in stress-marking systems such as Spanish, where accented and unaccented vowels correspond to the same phoneme and are presumed to share the same abstract letter unit (Conrad et al., 2010; Labusch et al., in press), the asymmetry is minimized, whereas in phoneme-marking systems, vowels with and without diacritics map onto separate letter representations, thereby amplifying it. Although compatible with our findings, interactive activation models may need parameter adjustments (e.g., stronger inhibition for extra features) and, more fundamentally, a more realistic letter layer than the Rumelhart and Siple (1974) font to capture diacritic-specific processing (see Davis, 2010; Snell, 2025, for limitations of the orthographic front-end of these models).

At a general level, most implemented models of visual word recognition focus on the interplay between activated lexical units rather than on the orthographic front-end of the models (McClelland & Rumelhart, 1981). In recent front-end models of orthographic processing, the focus has been on letter position coding rather than letter identity (e.g., LTRS: Adelman, 2011; Spatial Coding: Davis, 2010; PONG: Snell, 2025; Über-Reader: Reichle, 2020). Indeed, they underspecify how visual features become letter identities (see Davis, 2010; Snell, 2025, for discussion). Our findings show that this mapping is nontrivial: a diacritic is not an optional stroke but a feature that changes letter identity and lexical access. Furthermore, diacritic manipulations are particularly useful because they create clean, categorical contrasts (a vs. ä; n vs. ñ; s vs. š) that are easier to formalize than confusable pairs like i–j or u–v (Marcet & Perea, 2017). They thus provide an ideal testing ground for future computational models of visual word recognition to incorporate feature-level coding.

In addition, one important next step is to determine when in processing the asymmetry emerges. Masked priming studies show early insensitivity to omission (*fácil–FÁCIL* [easy] produces similar response times as *facil–FÁCIL*) but reduced priming for addition across languages (e.g., *feliz–FELIZ* [happy] yields faster responses than *féliz–FELIZ*; Benyhe et al., 2023; Perea et al., 2020; see also Marcet et al., 2021, for diacritical consonants), suggesting that the asymmetry arises during perceptual encoding. ERP studies targeting components linked to orthographic processing (Holcomb & Grainger, 2002) could test whether the extra cost of adding a diacritic emerges at the feature-to-letter mapping stage. Eye-tracking during sentence reading could further assess whether the asymmetry generalizes to natural reading, particularly via parafoveal preview effects (Rayner, 2009; Marcet & Perea, 2022).

Conclusions

In sum, orthographies that use diacritics to encode phonemic distinctions (e.g., German, Finnish) are best described by models in which diacritical and non-diacritical vowels are represented as distinct letter units. In both German and Finnish, adding a diacritical mark led to greater reading costs than omitting one, demonstrating that low-level perceptual features exert a persistent influence on lexical access. This challenges purely abstractionist accounts and favors models incorporating perceptual noise or feature-level inhibition. Our results also contribute to Hofstadter's (1985) enduring question: what counts as a letter? In German and Finnish, the answer is clear: the "a" in *Handel* is not the same as the "ä" in *Händel*, and an added diacritic disrupts reading more than its omission. Future computational models must therefore incorporate feature-level coding to capture these asymmetries.

References

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

AlJassmi, M. A. & Perea, M. (2024). Visual similarity effects in the identification of Arabic letters: Evidence with masked priming. *Language and Cognition*, *16*, 1618–1638. https://doi.org/10.1017/langcog.2024.20

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4*, 439–461. https://doi.org/10.3758/BF03214334

Aust, F., & Barth, M. (2022). papaja: Prepare reproducible A*PA journal articles with R-Markdown*. https://github.com/crsh/papaja

Bae, S., Lee, C. H., & Pae, H. K. (2024). Visual letter similarity effects in Korean word recognition: The role of distinctive strokes. *Quarterly Journal of Experimental Psychology*, *78*, 1369–1378. https://doi.org/10.1177/17470218241278600

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science, 20*(3), 160–166. https://doi.org/10.1177/0963721411408885

Balota, D., Yap, M. J., & Cortese, M. J. (2006). Visual word recognition: The journey from features to meaning (A travel update). In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics* (2nd Edition) (pp. 285–375). Academic Press.

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*, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Benyhe, A., Labusch, M., & Perea, M. (2023). Just a mark: Diacritic function does not play a role in the early stages of visual word recognition. *Psychonomic Bulletin & Review*, *30*, 1530–1538. https://doi.org/10.3758/s13423-022-02244-4

Bertram, R., Pollatsek, A., & Hyönä, J. (2004). Morphological parsing and the use of segmentation cues in reading Finnish compounds. *Journal of Memory and Language, 51,* 325–345. https://doi.org/10.1016/j.jml.2004.06.005

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

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*(1), 9. https://doi.org/10.5334/joc.10

Chauncey, K., Holcomb, P. J., & Grainger, J. (2008). Effects of stimulus font and size on masked repetition priming: An event-related potentials (ERP) investigation. *Language and Cognitive Processes, 23,* 183–200. https://doi.org/10.1080/01690960701579839

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

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256. https://doi.org/10.1037/0033-295x.108.1.204

Colombo, L., & Sulpizio, S. (2021). The role of orthographic cues to stress in Italian visual word recognition. *Quarterly Journal of Experimental Psychology, 74*, 1631–1641. 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*, 861–893. https://doi.org/10.1080/09541440903356777

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

Davis, C. J., & Andrews, S. (2001). Inhibitory effects of transposed-letter similarity for words and nonwords of different lengths. Australian Journal of Psychology, *53*(Suppl. 1), 50, https://doi.org/10.1080/00049530.2001.10600104

Davis, C. J., Perea, M., & Acha, J. (2009). Re(de)fining the orthographic neighbourhood: The role of addition and deletion neighbours in lexical decision and reading. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1550–1570. https://doi.org/10.1037/a0014253

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

Duñabeitia, J. A., Perea. M., & Labusch, M. (2023). Rëâdīńg wõrdš wîth ōrńåmêńtš: is there a cost? *Frontiers in Psychology, 14*, 1168471. https://doi.org/10.3389/fpsyg.2023.1168471

Fernández-López, M., & Perea, M. (2023). A letter is a letter and its co-occurrences: Cracking the emergence of position-invariance processing. *Psychonomic Bulletin & Review, 30*, 2328–2337. https://doi.org/10.3758/s13423-023-02265-7

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 680–698. https://doi.org/10.1037/0278-7393.10.4.680

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences, 110*(20), 8051–8056. https://doi.org/10.1073/pnas.1216438110

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, 74*, 1542–1552. https://doi.org/10.1177/17470218211012960

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes, 23*, 1–35. https://doi.org/10.1080/01690960701578013

Grainger, J. (2018). Orthographic processing: A "mid-level" vision of reading. *Quarterly Journal of Experimental Psychology, 71*, 335–359. https://doi.org/10.1080/17470218.2017.1314515

Grainger, J. (2024). Letters, words, sentences, and reading. *Journal of Cognition, 7*(1), 66. https://doi.org/10.5334/joc.396.

Grainger, J., Kiyonaga, K., & Holcomb, P. J. (2006). The time course of orthographic and phonological code activation. *Psychological Science, 1*7, 1021–1026. https://doi.org/10.1111/j.1467-9280.2006.01821.x

Grainger, J., Rey, A., & Dufau, S. (2008). Letter perception: From pixels to pandemonium. *Trends in Cognitive Sciences, 12*, 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*, 966–984. https://doi.org/10.3758/s13415-019-00696-1

Grainger, J., & Holcomb, P. J. (2009). Watching the word go by: On the time-course of component processes in visual word recognition. *Language and Linguistics Compass, 3*, 128–156. https://doi.org/10.1111/j.1749-818X.2008.00121.x

Hofstadter, D. (1985). *Metamagical Themas: Questing for the Essence of Mind and Pattern*. Basic Books.

Holcomb, P. J., & Grainger, J. (2002). On the time course of visual word recognition: An event-related potential investigation using masked repetition priming. *Journal of Cognitive Neuroscience, 14*(3), 433–444. https://doi.org/10.1162/089892902317361967

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*, 273–296. https://doi.org/10.1016/j.cognition.2003.09.006

Kinoshita, S., Yu, L., Verdonschot, R. G., & Norris, D. (2021). Letter identity and visual similarity in the processing of diacritic letters. *Memory & Cognition, 49*, 815–825. https://doi.org/10.3758/s13421-020-01125-2

Jacobs, A. M., Ferrand, L., & Grainger, J. (1995). The incremental priming technique: A method for determining within-condition priming effects. *Perception & Psychophysics, 57*(8), 1101–1110. https://doi.org/10.3758/BF03208369

Jouravlev, O., McPhedran, M., Hodgins, V., & Jared, D. (2023). Cross-language semantic parafoveal preview benefits in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 49,* 1683–1697. https://dx.doi.org/10.1037/xlm0001238

Labusch, M., Gómez, P., & Perea, M. (2022). Does adding an accent mark hinder lexical access? Evidence from Spanish. *Journal of Cultural Cognitive Science, 6,* 219–228. https://doi.org/10.1007/s41809-022-00104-0

Labusch, M., Massol, S., Marcet, A., & Perea, M. (2023). Are goats chèvres, chévres, chēvres, and chevres? Unveiling the orthographic code of diacritical vowels. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 49,* 301–319. https://doi.org/10.1037/xlm0001212

Labusch, M., Perea, M., & Hyönä, J. (2025). *The Costs of Adding versus Omitting Diacritics in Visual Word Recognition: Evidence from German and Finnish*. https://osf.io/jtda8

Labusch, M., Perea, M., Marcet, A., Baciero, A., & Fernández-López, M. (in press). The Role of Diacritics in the Recognition of Words across Different Writing Systems. In H. Winskel and H. Pae (Eds.), *Handbook of Nonlinear Writing Systems – Complex Processes and Learning Challenges*. Springer.

Laine, M., & Virtanen, P. (1999). *WordMill lexical search program*. Center for Cognitive Neuroscience, University of Turku.

Levy, R. (2008). A noisy-channel model of rational human sentence comprehension under uncertain input. In M. Lapata and H. T. Ng (Eds.), *Proceedings of the 2008 conference on empirical methods in natural language processing* (pp. 234–243). Association for Computational Linguistics.

Lüdecke, D., Bartel, A., Schwemmer, C., Powell, C., Djalovski, A., & Titz J. (2021a). *Package 'sjPlot'*. https://cran.r-project.org/web/packages/sjPlot/

Lüdecke, D., Patil. I., Ben-Shachar, M. S., Wiernik. B. M., Waggoner, P., & Makowski, D. (2021b). see: An R package for visualizing statistical models. *Journal of Open Source Software, 6*, 3393. https://doi.org/10.21105/joss.03393

Marcet, A., Fernández-López, M., Baciero, A., Sesé, A., & Perea, M. (2022). What are the letters e and é in a language with vowel reduction? The case of Catalan. *Applied Psycholinguistics, 43*, 193–210. https://doi.org/10.1017/S0142716421000497

Marcet, A., Fernández-López, M., Labusch, M., & Perea, M. (2021). The omission of accent marks does not hinder word recognition: Evidence from Spanish. *Frontiers in Psychology*, *12*, 794923. https://doi.org/10.3389/fpsyg.2021.794923

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*, 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*, 1180–1185. https://doi.org/10.3758/s13423-016-1180-9

Marcet, A., & Perea, M. (2022). Does omitting the accent mark in a word affect sentence reading? Evidence from Spanish. *Quarterly Journal of Experimental Psychology*, *75*, 148–155. https://doi.org/10.1177/17470218211044694

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. Psychological Review, 88, 375–407. https://doi.org/10.1037/0033-295x.88.5.375

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

O’Connor, R. E., & Forster, K. I. (1981). Criterion bias and search sequence bias in word recognition. *Memory & Cognition, 9*, 78–92. https://doi.org/10.3758/BF03196953

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, 578–596. https://doi.org/10.1017/S1366728915000292

Özkan, Z. G., Özdemir, B., Gómez, P., & Perea, M. (2025). The distinctive role of vowel harmony in visual word recognition: The case of Turkish. *Quarterly Journal of Experimental Psychology*, *78*, 1539–1546. https://doi.org/10.1177/17470218241277683

Peirce, J. W., & MacAskill, M. R. (2018). *Building Experiments in PsychoPy.* Sage.

Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language, 51*, 231–246. https://doi.org/10.1016/j.jml.2004.05.005

Perea, M., Baciero, A., & Marcet, A. (2021). Does a mark make a difference? Visual similarity effects with accented vowels. *Psychological Research, 85*, 2279–2290. https://doi.org/10.1007/s00426-020-01405-1

Perea, M., Fernández-López, M., & Marcet, A. (2020). What is the letter é? *Scientific Studies of Reading, 24*, 434–443. https://doi.org/10.1080/10888438.2019.1689570

Perea, M., Hyönä, J., & Marcet, A. (2022a). Does vowel harmony affect visual word recognition? Evidence from Finnish. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*, 2004–2014. https://doi.org/10.1037/xlm0000907

Perea, M., Labusch, M., & Marcet, A. (2022b). How are words with diacritical vowels represented in the mental lexicon? Evidence from Spanish and German. *Language, Cognition, and Neuroscience, 37*, 457–468. https://doi.org/10.1080/23273798.2021.1985536

Perea, M., Gomez, P., & Baciero, A. (2023). Do diacritics entail an early processing cost in the absence of abstract representations? Evidence from masked priming in English. *Language and Speech, 66,* 105–117. https://doi.org/10.1177/00238309221078321

Perea, M., Vergara-Martínez, M., & Gomez, P. (2015). Resolving the locus of cAsE aLtErNaTiOn effects in visual word recognition: Evidence from masked priming. *Cognition, 142*, 39–43. https://doi.org/10.1016/j.cognition.2015.05.007

Protopapas, A., & Gerakaki, S. (2009). Development of processing stress diacritics in reading Greek. *Scientific Studies of Reading, 13*, 453–483. https://doi.org/10.1080/10888430903034788

R Core Team (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*, 1457–1506. https://doi.org/10.1080/17470210902816461

Rayner, K., & Kaiser, J. S. (1975). Reading mutilated text. Journal of Educational Psychology, 67, 301–306. https://doi.org/10.1037/h0077015

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*(3), 446–461. https://doi.org/10.1037/0033-2909.86.3.446

Reichle, E. D. (2020). *Computational models of reading: A handbook*. Oxford University Press.

Rey, A., Dufau, S., Massol, S., & Grainger, J. (2009). Testing computational models of letter perception with item-level event-related potentials. *Cognitive Neuropsychology, 26*, 7–22. https://doi.org/10.1080/09541440802176300

Rumelhart, D. E. (1985). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 573–603). Erlbaum.

Schad, D. J., Nicenboim, B., Bürkner, P. C., Betancourt, M., & Vasishth, S. (2021). Workflow techniques for the robust use of Bayes factors. *Psychological Methods, 26*(4), 429–455. https://doi.org/10.1037/met0000275

Scholz, M., & Bürkner, P. C. (2023). *Posterior accuracy and calibration under misspecification in Bayesian generalized linear models*. arXiv preprint arXiv:2311.09081.

Schwab, S. (2015). Accent mark and visual word recognition in Spanish. *Loquens, 2,* e018. https://doi.org/10.3989/loquens.2015.018

Snell, J. (2025). PONG: A computational model of visual word recognition through bi-hemispheric activation. *Psychological Review*, *132*, 505–527. https://doi.org/10.1037/rev0000461

Stan Development Team (2023). *RStan: The R interface to Stan.* R package version 2.26.22, https://mc-stan.org/

Suomi, K., McQueen, J.M., & Cutler, A. (1997). Vowel harmony and speech segmentation in Finnish. *Journal of Memory and Language, 36,* 422–444. https://doi.org/10.1006/jmla.1996.2495

Tuomainen, J. (2001). *Language specific cues to segmentation of spoken words in Finnish: Behavioral and event-related brain potential studie*s [Ph.D. thesis]. de Katholieke Universiteit Brabant.

Treisman, A. (1986). Features and objects in visual processing. *Scientific American, 255*, 114–125. https://www.jstor.org/stable/24976089

Vasishth, S., Nicenboim, B., Beckman, M. E., Li, F., & Kong, E. J. (2018). Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics, 71*, 147–161. https://doi.org/10.1016/j.wocn.2018.07.008

Wells, J. C. (2000). Orthographic diacritics and multilingual computing. *Language Problems and Language Planning, 24*, 249–272. https://doi.org/10.1075/lplp.24.3.04wel

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15*, 971–979. https://doi.org/10.3758/PBR.15.5.971

Ziegler, J. C., Perry, C., & Coltheart, M. (2000). The DRC model of visual word recognition and reading aloud: An extension to German. *European Journal of Cognitive Psychology, 12,* 413–430. https://doi.org/10.1080/09541440050114570

Footnotes

Footnote 1. Some noisy-channel models of reading predict greater reading costs for addition than for deletion (e.g., Gibson et al., 2013), whereas others assume a symmetric cost structure (e.g., Levy, 2008).

Footnote 2. A similar asymmetry has been reported in letter recognition tasks (e.g., a–Á yields similar response times to á–Á, but a–A produces faster responses than á–A; Perea et al., 2020; see also Kinoshita et al., 2021, for Japanese kana). However, this asymmetry appears task-dependent (e.g., it is absent in alphabetic decision tasks; see AlJassmi & Perea, 2024; Bae et al., 2024; Marcet et al., 2022).

Footnote 3. German and Finnish readers may occasionally encounter words with omitted diacritics, especially in English media (e.g., Räikkönen → Raikkonen; Schröder → Schroeder). By contrast, they rarely see words where diacritics are added to letters that do not normally carry them, except for stylistic effect, as in certain heavy metal band names (e.g., Motörhead).

Footnote 4. We employed version 4.3.1 of R (R Core Team, 2023) with the packages *brms* (version 2.19.0; Bürkner, 2017, 2018, 2021), *papaja* (version 0.1.1; Aust & Barth, 2022), *sjPlot* (version 2.8.14; Lüdecke et al., 2021a), and *see* (version 0.8.0; Lüdecke et al., 2021).

Footnote 5. A complementary analysis using a reciprocal transformation of RTs (–1000/RT) produced a 95% credible interval that excluded zero, providing converging evidence for the asymmetry (see OSF for the full model output).

Footnote 6. While word length was controlled across conditions, one might argue that the relative impact of adding or omitting a diacritic could be smaller in shorter than in longer words (see Davis & Andrews, 2001, for mixed evidence on whether transposed-letter similarity effects are modulated by length). To test this, we computed Bayesian Pearson correlations between mean item RTs and number of letters in all three experiments. The analyses showed weak negative associations (Experiment 1: $r = -.182$, 95% CrI [−.346, −.002]; Experiment 2: $r = -.254$, 95% CrI [−.397, −.093]; Experiment 3: $r = -.278$, 95% CrI [−.419, −.118]), indicating that longer words tended to show slightly smaller misspelling costs. Importantly, the pattern was similar when correlations were computed separately for omission and addition misspellings (Experiment 2: $r = -.317$ for omission, $r = -.212$ for addition; Experiment 3: $r = -.284$ for omission, $r = -.292$ for addition). While beyond the scope of this paper, the interplay between word length and orthographic processing merits further investigation.

Table 1. Example stimuli of Experiment 1 (German). The critical comparisons were between the correct and incorrect spellings of each condition. Two blocks were used to present the items with diacritical and non-diacritical base words separately to each participant.

	Diacritical base words			**Non-diacritical base words**		
spelling	***Intact***		***Misspelled***	***Intact***		***Misspelled***
Animal words	Kröte [toad]	⟷	Krote	Schwan [swan]	⟷	Schwän
Non-Animal words	Hügel [hill]	⟷	Hugel	Korken [cork]	⟷	Körken

Table 2. Mean response times (in ms) and error rates (in percentages) for non-animal and animal names with (1) the intact format or the added diacritic and (2) the intact format or the removed diacritic in German (Experiment 1).

	Correct spelling		Incorrect spelling	
	Response time	**% Errors**	**Response time**	**% Errors**
		Non-diacritical base words		
Non-Animal	631	1.8	670	2.6
Animal	634	4.6	650	4.3
		Diacritical base words		
Non-Animal	639	2.4	664	2.7
Animal	641	4.7	668	7.4

Table 3. Mean response times (in ms) and error rates (in percentages) for non-animal and animal names with (1) the intact format or the added diacritic and (2) the intact format or the removed diacritic in Finnish (blocked presentations, Experiment 2).

	Correct spelling		**Incorrect spelling**	
	Response time	**% Errors**	**Response time**	**% Errors**
		Non-diacritical base words		
Non-Animal	643	3.1	672	3.9
Animal	645	6.7	673	8.1
		Diacritical base words		
Non-Animal	630	2.2	643	3.9
Animal	636	8.8	651	10.7

Table 4. Mean response times (in ms) and error rates (in percentages) for non-animal and animal names with (1) the intact format or the added diacritic and (2) the intact format or the removed diacritic in Finnish (unblocked presentations, Experiment 3).

	Correct spelling		**Incorrect spelling**	
	Response time	**% Errors**	**Response time**	**% Errors**
	Non-diacritical base words			
Non-Animal	625	3.6	670	3.8
Animal	651	8.3	689	6.8
	Diacritical base words			
Non-Animal	628	1.9	655	2.3
Animal	649	7.1	670	8.9

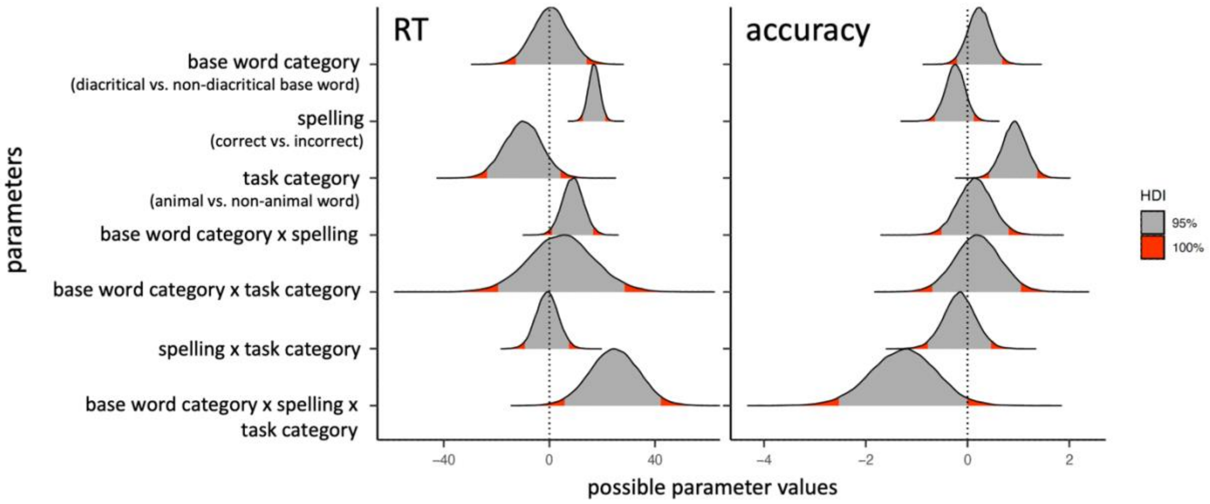

Figure 1. Posterior distributions for each of the estimates of the Bayesian Linear Mixed-Effects models on response time (left panel) and accuracy (right panel) for German words in Experiment 1. The grey areas correspond to the 95% Credible Intervals.

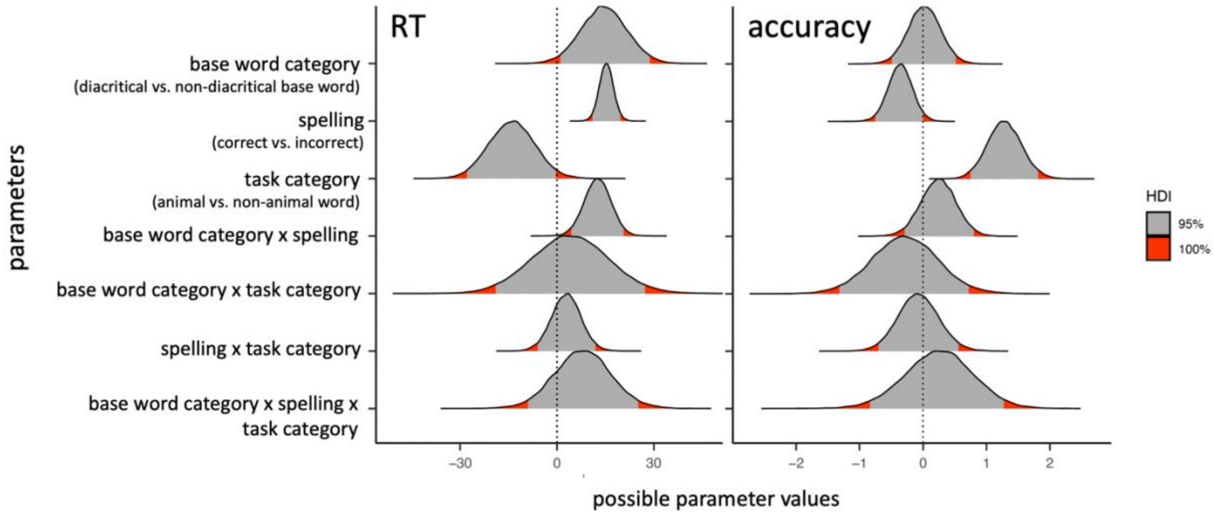

Figure 2. Posterior distributions for each of the estimates of the Bayesian Linear Mixed-Effects models on response time (left panel) and accuracy (right panel) for Finnish words in Experiment 2. The grey areas correspond to the 95% Credible Intervals.

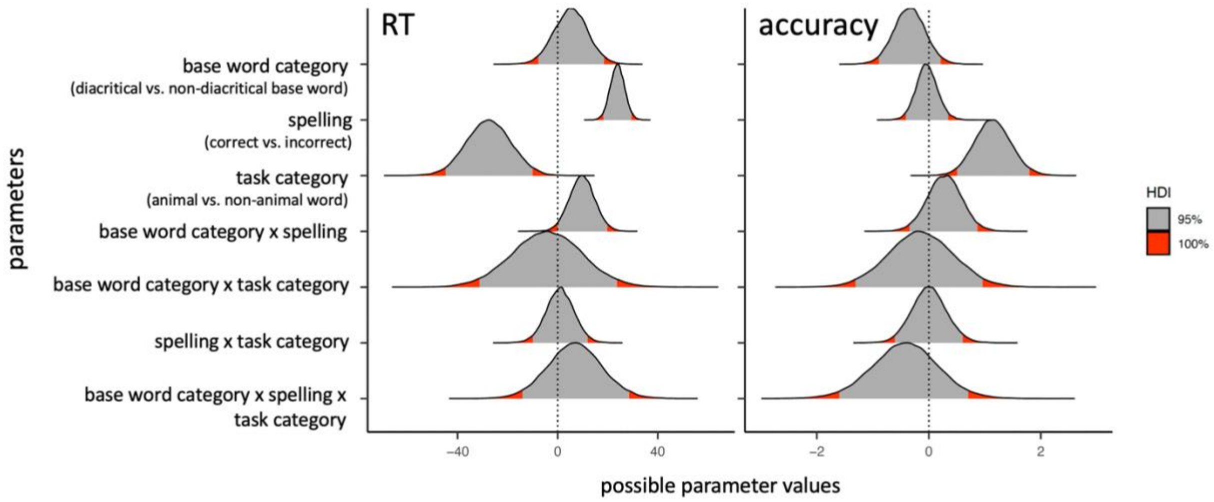

Figure 3. Posterior distributions for each of the estimates of the Bayesian Linear Mixed-Effects models on response time (left panel) and accuracy (right panel) for Finnish words in Experiment 3 (fully randomized presentations). The grey areas correspond to the 95% Credible Intervals.

Appendix 1. Materials of the experiments

Experiment 1 (German)

Diacritical non-animals: Töpferei, Ähre, Geländer, Jubiläum, Tonhöhe, Glühwein, Küchenherd, Unglück, Getränk, Gefäß, Südpol, Baulärm, Porträt, Speiseöl, Hörbuch, Köstlichkeiten, Behörde, Hügel, Ostküste, Gewürz, Sägemehl, Hochzeitspläne, Fremdkörper, Reiseführer, Löffel, Gemüse, Hutgröße, Münze, Diät, Molekül, Überstunde, Bustür, Schiebetür, Gemälde, Frisör, Knödel, Bücherei, Möbelhaus, Süßigkeit, Föhn

Non-diacritical non-animals: Espresso, Altstadt, Mappe, Tuben, Herstellung, Kannen, Lexikon, Putzfrau, Festival, Pirat, Gasthof, Waggon, Entspannung, Konsulat, Sprichwort, Topf, Monument, Notsitz, Ratgeber, Sirup, Magnet, Andenken, Umwelt, Korken, Korruption, Unterhose, Logo, Sanduhr, Kronleuchter, Badematte, Taschenlampe, Halstuch, Feuerwehr, Rechtsanwalt, Nadelkopf, Haarkur, Marmor, Ohrring, Quiz, Waage

Diacritical animals: Krähe, Mäuse, Siebenschläfer, Schildkröte, Gürteltier, Kröte, Marienkäfer, Eichhörnchen, Möwe, Maikäfer, Büffel, Wühlmaus, Hyäne, Küken, Känguru, Hündin, Chamäleon, Löwe, Erdmännchen, Seelöwe

Non-diacritical animals: Faultier, Zebra, Skorpione, Fledermaus, Krokodil, Ameise, Schwan, Otter, Bussard, Kaninchen, Maulesel, Flamingo, Giraffe, Kanarienvogel, Gans, Salamander, Regenwurm, Truthahn, Ratte, Gazelle

Experiments 2 and 3 (Finnish)

Diacritical non-animals: jäätelö, kärry, tähkä, päre, pöytäliina, jäkälä, kääretorttu, höylä, käsipyyhe, lähiö, hävittäjä, säiliö, sähkökiuas, mänty, käsijarru, likööri, käpy, käsine, köysi, häkki, levä, käyttötili, jälkiruoka, köynnös, läppä, näkötorni, tölkki, pönttö, päärynä, pääkallo, työpöytä, lätäkkö, vyölaukku, lääni, pöytälamppu, ränni, kävelykeppi, kävelysilta, sähköliesi, vehnä, höyry, mämmi, pölkky, käsiohjelma, kypärä, kylpylä, päähine, pyöräteline

Non-diacritical non-animals: Karpalo, hamppu, rannekello, lossi, appelsiini, hauli, tulostin, aprikoosi, urheiluhalli, sitruuna, uimapuku, kaulahuivi, harppu, laboratorio, lippahattu, mustikka, sora, taateli, kirjahylly, hanuri, lusikka, puolukka, haarukka, kantarelli, haarniska, lato, saavi, kauppakeskus, hiutale, vitamiini, laine, lipasto, nojatuoli, kaali, saniainen, katos, raitiovaunu, kasarmi, motelli, solmio, holvi, kulho, komero, ohra, navetta, harppi, kaulus, vesipullo

Diacritical animals: jääkarhu, kärpänen, leppäkerttu, sääksi, mehiläinen, västäräkki, särki, lehmä, päästäinen, kärppä, käki, jänis, mäyrä, pääskynen, haisunäätä, myyrä, käärme, näätä, pöllö, räkättirastas, pyöriäinen, härkä, telkkä, heinäsirkka

Non-diacritical animals: muurahainen, kilpikonna, nahkiainen, itikka, antilooppi, hauki, varis, sarvikuono, kuikka, ahma, korppikotka, kovakuoriainen, lisko, mustekala, lokki, lammas, ahven, tikka, varpunen, krokotiili, ankerias, perhonen, karhu, siika

Appendix 2. Population-level estimates of the response time (RT) models

Table A1. Population-level estimates for the RT model in Experiment 1

Predictor	Estimate (b)	SE	95% CrI LL	95% CrI UL
Intercept	650.96	8.04	635.35	666.83
Word type (target1)	0.47	6.81	–13.01	13.81
Spelling (spelling1)	16.93	2.16	12.67	21.17
Category (category1)	–9.97	7.10	–23.83	3.97
Word type × Spelling	8.89	3.96	1.15	16.71
Word type × Category	4.60	12.21	–19.28	28.78
Spelling × Category	–0.83	4.23	–9.17	7.44
Word type × Spelling × Category	24.57	9.31	6.12	42.63

Note. Predictor coding: “Word type (target1)” contrasts diacritical (–0.5) vs. non-diacritical (+0.5) base words; “Spelling (spelling1)” contrasts correct (–0.5) vs. misspelled (+0.5) forms; “Category (category1)” codes animal (–2/3) vs. non-animal (+1/3) items, reflecting the 1:3 ratio of animals to non-animals in the design. Estimates are posterior means from Bayesian linear mixed-effects models. SE = posterior standard error; CrI = credible interval; LL = lower limit; UL = upper limit.

Table A2. Population-level estimates for the RT model in Experiment 2

Predictor	Estimate (b)	SE	95% CrI LL	95% CrI UL
Intercept	649.63	8.13	633.69	665.75
Word type (target1)	14.59	7.08	0.75	28.48
Spelling (spelling1)	15.32	2.20	11.04	19.68
Category (category1)	–13.97	7.07	–27.62	–0.03
Word type × Spelling	12.59	4.17	4.38	20.64
Word type × Category	4.14	11.85	–18.91	27.50
Spelling × Category	2.88	4.54	–6.12	11.76
Word type × Spelling × Category	7.98	8.73	–9.14	25.15

Note. We used the same predictor coding and abbreviations as in Table A1.

Table A3. Population-level estimates for the RT model in Experiment 3

Predictor	Estimate (b)	SE	95% CrI LL	95% CrI UL
Intercept	652.84	8.36	636.23	669.02
Word type (target1)	5.28	6.70	–7.90	18.40
Spelling (spelling1)	23.90	2.85	18.38	29.56
Category (category1)	–27.56	8.87	–44.93	–10.20
Word type × Spelling	9.72	5.06	–0.16	19.74
Word type × Category	–4.18	13.96	–31.67	23.36
Spelling × Category	0.98	5.57	–10.05	11.91
Word type × Spelling × Category	6.87	10.76	–14.21	28.30

Note. We used the same predictor coding and abbreviations as in Table A1.