LANGUAGE, COGNITION AND NEUROSCIENCE
https://doi.org/10.1080/23273798.2026.2701846

Reading through the lines: strikethroughs, shifted halves, and the tolerance to visual distortion in word recognition

Maria Fernández-López [a,b], Inka Romero-Ortells[c], Ana Marcet[d], Melanie Labusch[b,e,f] and Manuel Perea[b,c,f]

[a]Department of Basic Psychology, Universitat de València, València, Spain; [b]ERI-Lectura, Universitat de València, València, Spain; [c]Nebrija Research Center in Cognition, Universidad Nebrija, Madrid, Spain; [d]Department of Language and Literature Education, Universitat de València, València, Spain; [e]Health Research Institute La Fe, València, Spain; [f]Department of Methodology, Universitat de València, València, Spain

ABSTRACT

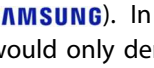
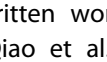
Skilled reading requires mapping variable visual input onto stable orthographic codes. A central question is how much visual distortion can be tolerated without increased lexical involvement. We investigated this issue using three lexical decision experiments with different forms of visual degradation. In Experiments 1-2, words were presented intact or with single or double strikethroughs, selectively occluding diagnostic letter features while preserving overall word shape. Both distortions slowed responses, with higher costs for the double strikethrough; critically, distortion effects were additive with word frequency. In Experiment 3, we misaligned the upper and lower halves of letters, disrupting the supraletter envelope. Once again, distortion effects were additive with word frequency. Together, these findings constrain models of visual word recognition that predict lexical involvement under moderate visual degradation. Instead, they indicate that the orthographic front-end is robust, extracting orthographic information despite visual noise.

ARTICLE HISTORY
Received 31 December 2025
Accepted 27 June 2026

KEYWORDS
Orthographic processing; visual distortion; letter detectors; lexical involvement

Introduction

Reading and writing are relatively recent cultural inventions that have repurposed visual cortical regions originally devoted to object recognition (Dehaene, 2009). These areas have been recycled to process arbitrary visual symbols (e.g. letters in alphabetic scripts), and to link them to linguistic representations. Within this framework, visual word recognition can be considered a special case of perceptual expertise. Just as individuals can rapidly categorise diverse instances of objects (e.g.,  ,) as instances of the same object category (i.e. table), skilled readers develop finely tuned recognition abilities for words, underpinned by the adaptation of the visual system to a restricted set of diagnostic features (see Grainger, 2008, 2017; Harel, 2016; McCandliss et al., 2003). In reading acquisition, letters transition from isolated objects to elements within multi-letter strings (e.g. words), enabling parallel orthographic processing (Grainger & Hannagan, 2014).

The ability to identify objects amid considerable variability is a defining characteristic of perceptual expertise, including skilled reading. Words can be recognised effortlessly regardless of whether they appear in different fonts, cases, or sizes (e.g. **dog**, *dog*, DOG). This feat is thought to be achieved through a flexible mapping from variable visual inputs to invariant abstract representations. Neurobiologically inspired models of visual word recognition (e.g. Local Combination Detector model, Dehaene et al., 2005; see also Grainger et al., 2008) characterise this process as a hierarchy of detectors with progressively larger receptive fields, ranging from orientation-selective units coding simple features, to letter detectors, to units coding letter clusters and entire words. A key property of these models is the inherent tolerance of the letter detectors to moderate levels of visual noise. Specifically, the arrays of letter detectors can activate their targets even when some visual details are partially missing or degraded (e.g. as in the "A" in ). In skilled readers, visual word recognition would only demand additional effort when distortions surpass these built-in tolerance thresholds (e.g. the handwritten word  ; see Cohen & Dehaene, 2010; Qiao et al., 2010), thereby requiring greater lexical involvement to support recognition.

One of the most robust diagnostic tools for examining lexical involvement during visual word recognition is the word-frequency effect. High-frequency words are recognised faster and more accurately than low-frequency words, a finding that is consistently found

CONTACT Maria Fernández-López maria.fernandez@uv.es

across tasks (Balota & Chumbley, 1984; Rayner & Duffy, 1986; see Brysbaert et al., 2018 for review). Critically, the interplay between the word-frequency effect and visual degradation may reveal the locus of the cost due to visual distortion. When visual degradation is small (e.g. easy-to-read handwritten text or moderate letter or word rotations) and remains within the tolerance range of letter detectors, the effects of distortion and word frequency tend to be additive; that is, the slowdown due to distortion is similar for high- and low-frequency words, consistent with an early (prelexical) locus of the distortion cost. However, when the degradation is severe enough to go beyond the tolerance of the letter detectors (e.g. CAPTCHA-like stimuli or difficult to read handwritten text), greater lexical support is needed to achieve recognition. As a result, the word-frequency effect is magnified, as high-frequency words can benefit more from lexical facilitation than low-frequency words (Barnhart & Goldinger, 2010; Fernández-López et al., 2023, 2025; Perea et al., 2015; see Vergara-Martínez et al., 2021, for ERP evidence).

A closely related benchmark comes from the literature on stimulus quality. In lexical decision, degrading the visual input (e.g. via contrast reduction or using masks of symbols alternating with the letter string) typically slows response times, with effects that are additive with the word-frequency effect (see Stanners et al., 1975, for the first demonstration, see also Balota et al., 2013; Besner & Young, 2022; Yap & Balota, 2007, for discussion).

A recent, seemingly divergent result appears to challenge this established pattern. Perea et al. (2023, Experiment 2) reported a reversal of the typical interaction between visual distortion and word frequency in the lexical decision task. Specifically, they introduced a distortion by horizontally shifting the upper halves of words relative to the lower halves (e.g. *sangre*; the Spanish word for *blood*), creating a misalignment that disrupted stroke continuity and introduced artificial intersections. Under this manipulation, high-frequency words were more impaired than low-frequency words, which is the opposite of what the above-cited accounts would predict. To explain this reversal, Perea and colleagues tentatively proposed that the manipulation could have selectively disrupted holistic processing of words as whole-word units, which is thought to be more prevalent for highly familiar words (Allen et al., 1995; Chen et al., 2016; Ventura et al., 2019, 2020). The idea is that when global word shape is disrupted, as in the half-shift manipulation, holistic processing becomes unreliable and high-frequency words may be particularly hindered. In contrast, low-frequency words, typically processed via a more analytic route (i.e. building word identity from individual letters), are relatively less affected. (A more direct evaluation of the holistic and hybrid accounts will be presented in the Introduction to Experiment 3.) While this interpretation is compelling, the half-shift manipulation simultaneously disrupted global word configuration (i.e. misaligning the letter halves) and letter-level continuity (i.e. breaking internal strokes). This dual impact makes it difficult to isolate the critical source of the effect.

The present experiments address this issue by introducing a novel distortion that primarily disrupts letter-level feature extraction while preserving global word shape. Specifically, we applied strikethrough lines across the stimuli, using the same words and nonwords as Perea et al. (2023). We employed a single line in Experiment 1 (e.g. ~~example~~) and a double line in Experiment 2 (e.g. ~~example~~). Strikethroughs introduce high-contrast visual noise that degrades the mapping from local diagnostic features (junctions, terminations) onto letter identities, while leaving the coarse word envelope relatively intact. This is an important contrast with the half-shift manipulation of Perea et al. (2023), which directly altered not only the local letter features, but also the global arrangement of letter parts. Critically, our design also implements a graded manipulation of feature occlusion. A single strikethrough provides a moderate level of disruption of the letter features, whereas a double strikethrough applies a more severe one. This sets up a direct test of where degradation costs arise in the processing stream.

We consider two broad classes of explanation for how the system handles visual degradation via feature occlusion and misalignment. In feedforward-dominant models of orthographic coding (e.g. Local Combinations Detector model; Dehaene et al., 2005), letter detectors have substantial tolerance to visual variability, such that low-to-moderate distortions can be handled by bottom–up accumulation and local normalisation processes (Cohen & Dehaene, 2010; Dehaene & Cohen, 2007). Importantly, these accounts do not deny the existence or usefulness of feedback; rather, they predict that – except when distortions exceed this tolerance – visual degradation should slow responses similarly for high- and low-frequency words.

In addition, other classes of models can accommodate increased lexical involvement under degraded visual formats. In interactive activation models (McClelland & Rumelhart, 1981; see also Davis, 2010; Grainger & Jacobs, 1996), this involvement may arise through bidirectional exchange between lexical and sublexical

levels; in cascaded feedforward models (e.g. Dual Route Cascaded model, Coltheart et al., 2001; CDP+, Perry et al., 2007), it may arise from differences in the temporal dynamics with which lexical representations accumulate evidence. For present purposes, the critical question is whether visual degradation increases the cost of low-frequency words relative to high-frequency. This interaction would indicate that the degraded input has become sufficiently ambiguous for lexical-level information to play a greater role during word recognition. This increased lexical contribution may arise either through interactive feedback mechanisms or through differences in the temporal dynamics of activation accumulation across lexical representations. To examine whether the effects of visual distortion were confined to the bulk of the response time distribution or also affected the slow responses, we modelled response times with an ex-Gaussian distribution. This allowed us to separate effects on the distributional mean, μ, from effects on the exponential scale parameter, β, which captures positive skew in the response time distribution (Balota & Yap, 2011; Heathcote et al., 1991).

We first report Experiments 1 and 2, which focus on single and double strikethroughs. These experiments test whether strikethroughs alone are sufficient to elicit increased lexical involvement. The pattern found in these experiments motivated Experiment 3, whose rationale and predictions are presented later.

Experiment 1

Method

Participants

A total of 50 native Spanish speakers (mean age = 28.3 years, SD = 4.8; 32 women) participated in the experiment. This sample size yielded 2,000 observations per condition, exceeding the recommendations for detecting small effects in mixed-effects models (Brysbaert & Stevens, 2018). All participants reported normal or corrected-to-normal vision and no history of reading or writing difficulties. They were recruited via the online platform Prolific Academic (www.prolific.co) and received monetary compensation for their participation. The study was approved by the Ethics Committee for Experimental Research at the University of València, and all participants provided informed consent prior to the start of the experiment.

Materials

The stimulus set consisted of 160 Spanish words (80 high-frequency and 80 low-frequency) and 160 orthographically legal nonwords, all six letters in length. These were the same stimuli used by Perea et al. (2023). For the high-frequency words, the mean Zipf frequency (as estimated from the subtitle-based EsPal database; Duchon et al., 2013) was 5.0 (range: 4.5–6.1). Their mean orthographic similarity, as measured by the Levenshtein distance to the 20 closest neighbours (OLD20), was 3.44 (range: 1.9–4.2), and their average imageability rating (on a 1–7 scale) was 5.5 (range: 2.7–6.6). The low-frequency words had a mean Zipf frequency of 3.4 (range: 3.0–3.7), an OLD20 of 3.36 (range: 1.7–4.25), and a mean imageability of 5.3 (range: 3.0–6.6). High- and low-frequency words were matched on OLD20 and imageability (both *ps* > .15). The set of nonwords was generated using the Wuggy software (Keuleers & Brysbaert, 2010), following the orthotactic constraints of Spanish. Their mean OLD20 was 3.46 (range 1.9-4.45), comparable to that of the word stimuli. Stimuli were presented in black 24-point Courier New font on a white background. Each item was presented either in its intact format (example) or with one strikethrough. The stimuli were distributed across two counterbalanced lists, and participants were randomly assigned to one of them.

Procedure

The experiment was programmed using PsychoPy 3 (Peirce et al., 2022) and conducted online via the Pavlovia platform (www.pavlovia.org). Demographic data were obtained from Prolific. Participants were instructed to complete the experiment in a quiet environment without any distractions. The task was a lexical decision task in which participants were asked to decide whether a given letter string was a real word or not by pressing the M key for "yes" and the Z key for "no". They were instructed to respond as quickly and accurately as possible. The programme recorded both response times and accuracy.

Each trial began with a fixation cross presented for 50 ms, immediately followed by the stimulus, which remained on screen until the participant responded or a response deadline of 2000ms was reached. The experiment comprised 336 trials, including 16 practice trials and 320 experimental trials. Feedback was provided during the practice block. There were breaks every 80 experimental trials. The completion time was approximately 20 min.

Results and discussion

For the analyses of the latency data, we excluded incorrect responses and response times faster than 250 ms (4.38% for both words and nonwords). Trials with no response within 2,000 ms were coded as errors. Table 1 shows the average response times and accuracy in each experimental condition.

Table 1. Average response times (ms) and % errors (in brackets) for single strikethrough and intact stimuli in Experiment 1.

Stimuli	Intact	Strikethrough	Distortion effect *(Strikethrough − Intact)*
Words			
High Frequency	537 (0.99)	560 (1.98)	23 (0.99)
Low Frequency	591 (5.01)	613 (7.49)	22 (2.48)
Nonwords	616 (4.88)	620 (4.95)	4 (0.07)

For the inferential statistics, we used the *brms* package (Bürkner, 2018) in R (R Core Team, 2022) to fit Bayesian linear mixed-effects models on the word and nonword data. The response time data were modelled with an ex-Gaussian distribution. In this parameterisation, μ indexes the overall mean of the fitted ex-Gaussian distribution, whereas β indexes the scale of the exponential component and captures positive skew in the RT distribution. Both μ and β were modelled as functions of Frequency, Format, and their interaction, whereas σ was estimated as a constant residual parameter across conditions. For accuracy, which was coded as binary (1 = correct; 0 = incorrect), we modelled the data with the Bernoulli distribution. In the word dataset, fixed effects included word frequency (coded as −0.5 for high-frequency words and 0.5 for low-frequency words) and format (coded as −0.5 for intact words and 0.5 for strikethrough words). We used the default priors from *brms* (Bürkner, 2018). The same procedure was applied to the nonword data, except that format was the only fixed factor. We specified the maximal random-effects structure allowed by the design:

$$[RT/\text{accuracy}] \sim \text{word\_frequency} * \text{format} + (1 + \text{word\_frequency} * \text{format} | \text{subject}) + (1 + \text{format} | \text{item})$$

Each model was run for 10,000 iterations (1,000 for warmup), across four chains. In both models, the chains converged successfully, and all $\hat{R}$ values (i.e. an index of convergence of the chains) were at their optimal value (i.e. 1.00). The model output included the coefficient (b) for each estimate (corresponding to the mean of the posterior distribution), the standard error, and the 95% credible interval (95% CrI). We considered evidence of an effect if the 95% CrI of its parameter estimate did not cross zero. The full outputs are presented in Table 2.

Word data

For the distributional mean of the RT model, we found that participants responded faster to high-frequency than to low-frequency words ($b = 57.99$, 95% CrI [51.42, 64.77]) and to intact words than to strikethrough words ($b = 24.71$, 95% CrI [19.93, 29.59]) but there were no signs of an interaction between the two factors ($b = 1.39$, 95% CrI [−7.42, 10.42]).

For the exponential component of the RT model, we only found a longer skew for low-frequency words than for high-frequency words ($b = 0.34$, 95% CrI [0.26, 0.41]); however, we did not find any clear signs for an effect of format ($b = 0.03$, 95% CrI [−0.04, 0.11]) or its interaction with frequency (b = 0.02, 95% CrI [−0.10, 0.13]).

Regarding the accuracy model, participants responded more accurately to high-frequency words than to low-frequency words ($b = -1.72$, 95% CrI [−2.22, −1.29]) and to intact words than to strikethrough words ($b = -0.63$, 95% CrI [−1.10, −0.18]). As with the latency data, there was no evidence for an interaction between frequency and format ($b = 0.20$, 95% CrI [−0.65, 1.02]).

Table 2. Posterior means (*b*), standard errors (SE), and 95% credible intervals (CrI) for fixed effects from Bayesian linear mixed-effects models (brms) for Experiment 1.

Stimulus/DV	Component	Parameter	Estimate (b)	SE	95% CrI [L, U]
Words					
RTs	Mean (μ)	Intercept	583.58	10.39	[562.81, 603.90]
		Frequency	57.99	3.41	[51.42, 64.77]
		Format	24.71	2.47	[19.93, 29.59]
		Frequency × Format	1.39	4.55	[−7.42, 10.42]
	Scale (β)	Intercept	4.50	0.06	[4.39, 4.61]
		Frequency	0.34	0.04	[0.26, 0.41]
		Format	0.03	0.04	[−0.04, 0.11]
		Frequency × Format	0.02	0.06	[−0.10, 0.13]
Accuracy	Log-Odds	Intercept	3.93	0.16	[3.64, 4.26]
		Frequency	−1.72	0.24	[−2.22, −1.29]
		Format	−0.63	0.23	[−1.10, −0.18]
		Frequency × Format	0.20	0.43	[−0.65, 1.02]
Nonwords					
RTs	Mean (μ)	Intercept	634.98	11.88	[611.74, 658.26]
		Format	8.31	3.06	[2.16, 14.21]
	Scale (β)	Intercept	4.53	0.06	[4.41, 4.65]
		Format	0.13	0.04	[0.05, 0.22]
Accuracy	Log-Odds	Intercept	4.01	0.19	[3.66, 4.38]
		Format	−0.03	0.20	[−0.43, 0.35]

Nonword data

In the RT analyses, participants responded faster to intact stimuli than to strikethrough nonwords in the distributional mean μ ($b = 8.31$, 95% CrI [2.16,14.21]). Furthermore, the exponential component indicated a longer tail for strikethrough nonwords compared to intact stimuli ($b = 0.13$, 95% CrI [0.05,0.22]). In accuracy, we found no evidence of a difference between strikethrough and intact nonwords ($b = -0.03$, 95% CrI [−0.43,0.35]).

Thus, the present experiment revealed a sizable effect of word frequency and a consistent advantage for intact stimuli over strikethrough stimuli. In the latency data, this distortion cost was primarily localised in the distributional mean (μ), while the exponential component (β) only showed the expected longer skew for low-frequency words (Ratcliff et al., 2004) but remained unaffected by the visual format. Critically, the frequency × format coefficient in RTs was centred near zero. The absence of an interaction in both distributional components suggests that letter detectors possess sufficient tolerance to resolve simple strikethrough distortions without requiring extra lexical support. As noted in the Introduction, an additive pattern between visual distortion and word frequency was found when comparing intact words to cursive words (Barnhart & Goldinger, 2010), easy-to-read handwritten words (Perea et al., 2016), words with letters rotated at low angles (Fernández-López et al., 2023), or words with the upper half of the letters moderately shifted (Perea et al., 2023; Experiment 1).

Experiment 2

A key question at this point is whether a stronger manipulation of letter distortion – through a more pronounced disruption of the visual input (a) – poses a challenge to the tolerance of letter detectors. In Experiment 2, we examined whether this greater level of distortion affects lexical processing, as reflected in a modulation of the word-frequency effect. As noted in the introduction, an interaction between the two factors would be consistent with greater lexical involvement in the normalisation of visual input. In this scenario, the reading cost due to letter distortion would be greater for low- than for high-frequency words – a pattern resembling that reported for difficult-to-read handwritten words (Barnhart & Goldinger, 2010; Vergara-Martínez et al., 2021) or words with rotated letters at angles of 67.5° (Fernández-López et al., 2023). Alternatively, another potential outcome is that letter detectors are still tolerant to this stronger distortion, yielding a pattern similar to that observed in Experiment 1 (i.e. additive effects of distortion and word-frequency).

Method

Participants

We recruited an additional sample of 50 participants (mean age = 29.32 years, SD = 4.81 years, 17 female) from the same population as Experiment 1. All participants provided informed consent before the experiment, which was approved by the Ethics Committee of the University of València.

Materials and procedure

The materials and procedure were the same as in Experiment 1, except that the stimuli were presented with double strikethrough (example) instead of simple strikethrough (Experiment 1).

Results and discussion

For the analysis of the response times, we excluded incorrect and extremely fast responses (8.2% for both words and nonwords). Table 3 shows the average response times and accuracy in each experimental condition.

For the inferential statistics, we followed a parallel procedure as in Experiment 1. Table 4 shows the full outputs of the models.

Word data

For the distributional mean of the RT model, we found that participants were faster to respond to high-frequency words than to low-frequency words ($b = 52.06$, 95% CrI [44.59, 59.54]), and to intact words over double-strikethrough words ($b = 36.27$, 95% CrI [29.94, 42.81]). In addition, both word-frequency ($b = 0.18$, 95% CrI [0.12, 0.24]) and format ($b = 0.09$,95% CrI [0.03, 0.15]) affected the exponential component, indicating that low-frequency and double-strikethrough words increased the positive skew of the RT distribution.

Table 3. Average Response Times (ms) and % Errors (in brackets) for double strikethrough and intact stimuli in Experiment 2.

Stimuli	Intact	Double Strikethrough	Distortion effect *(Double Strikethrough – Intact)*
Words			
High Frequency	554 (1.78)	587 (2.12)	33 (0.34)
Low Frequency	601 (4.61)	634 (6.94)	33 (2.33)
Nonwords	631 (5.12)	645 (4.38)	14 (−0.74)

Table 4. Posterior means (*b*), standard errors (SE), and 95% credible intervals (CrI) for fixed effects from Bayesian linear mixed-effects models (brms) for Experiment 2.

Stimulus/DV	Component	Parameter	Estimate (b)	SE	95% CrI [L, U]
Words					
RTs	Mean (μ)	Intercept	612.43	12.51	[588.04, 637.21]
		Frequency	52.06	3.83	[44.59, 59.54]
		Format	36.27	3.27	[29.94, 42.81]
		Frequency × Format	0.75	5.14	[−9.24, 10.94]
	Scale (β)	Intercept	4.66	0.06	[4.53, 4.78]
		Frequency	0.18	0.03	[0.12, 0.24]
		Format	0.09	0.03	[0.03, 0.15]
		Frequency × Format	−0.05	0.06	[−0.16, 0.07]
Accuracy	Log-Odds	Intercept	3.87	0.16	[3.56, 4.21]
		Frequency	−1.32	0.23	[−1.79, −0.89]
		Format	−0.63	0.23	[−1.10, −0.20]
		Frequency × Format	−0.04	0.36	[−0.71, 0.70]
Nonwords					
RTs	Mean (μ)	Intercept	666.84	13.55	[640.20, 693.45]
		Format	14.61	3.53	[7.66, 21.56]
	Scale (β)	Intercept	4.74	0.06	[4.63, 4.86]
		Format	0.11	0.03	[0.05, 0.17]
Accuracy	Log-Odds	Intercept	4.14	0.20	[3.76, 4.54]
		Format	0.24	0.20	[−0.14, 0.63]

There was no evidence of an interaction between the two factors in either the distributional mean component ($b = 0.75$, 95% CrI [−9.24, 10.94]) or the exponential component ($b = -0.05$, 95% CrI [−0.16, 0.07]).

In the accuracy model, participants were more accurate for high-frequency words than for low-frequency words ($b = -1.32$, 95% CrI [−1.79, −0.89]) and for intact words compared to double-strikethrough words ($b = -0.63$, 95% CrI [−1.10, −0.20]). We found no evidence of an interaction ($b = -0.04$, 95% CrI [−0.71, 0.70]).

Nonword data

The analysis of RTs indicated that participants responded faster to intact stimuli than to double-strikethrough nonwords in the distributional mean component ($b = 14.61$, 95% CrI [7.66, 21.56]). Additionally, the exponential component was longer in the double-strikethrough format ($b = 0.11$, 95% CrI [0.05, 0.17]). Regarding accuracy, no credible difference was found between intact and double-strikethrough nonwords ($b = 0.24$, 95% CrI [−0.14, 0.63]).

Experiment 2 replicated the qualitative pattern of results from Experiment 1, with a larger distortion cost for double strikethroughs (for words: 24 ms in Experiment 1 and 36 ms in Experiment 2). This was supported by a Format x Experiment interaction in a combined analysis including Experiment as a factor, with no evidence for an interaction with word frequency (see the OSF link for the full output). Notably, this stronger manipulation of visual distortion was reflected not only in the distributional mean (μ) but also in a longer exponential component (β) for double-strikethrough words than intact words. This indicates that while the double-strikethrough poses a greater challenge to the letter detector system – increasing the tail of the RT distribution – this additional slowing did not vary as a function of word frequency.

Once again, the absence of an interaction between word frequency and visual format in either distributional mean μ or the exponential component β indicates that the double-strikethrough format slows word recognition, relative to the intact format, regardless of word frequency.

In sum, Experiments 1 and 2 revealed the same qualitative pattern: additive effects of visual degradation (single and double strikethrough) and lexical access difficulty (high- vs. low-frequency words). This convergence strongly suggests that the visual distortion induced by the strikethrough manipulations is resolved at a prelexical, letter-based stage, within the tolerance range of the letter detector system.

Given this consistent pattern, it is natural to ask whether an interaction between distortion and word frequency would emerge for a manipulation that disrupts the global shape of the stimuli. A relevant case is Perea et al.'s (2023, Experiment 2) "letter-half" manipulation, in which the upper and lower halves of the words were horizontally shifted with respect to each other. In that experiment, high-frequency words were more impaired than low-frequency words when the halves were shifted; specifically, the distortion cost, relative to the intact words, was 13 ms for high-frequency words and 7 ms for low-frequency words.

As discussed in the Introduction, Perea et al. (2023) interpreted this small interaction as evidence for holistic word processing. The idea of holistic recognition for familiar forms has a long history in models of visual word recognition (see Davis, 2001, for review). In

particular, Allen et al.'s (1995) holistic-biased hybrid model proposes two partially independent routes: an analytic route that builds word identity from individual letters, and a holistic route that operates on stored visual patterns of entire words (see also Ventura et al., 2019, 2020). The balance between these routes is assumed to depend on familiarity. High-frequency words, encountered repeatedly, are more likely to be processed via the holistic route, whereas low-frequency words rely more heavily on analytic, letter-based decoding. Within this account, when global word shape is disrupted, as in the half-shift manipulation, one might expect the holistic route to become unreliable, such that high-frequency words (which rely on it most) are more hindered than low-frequency words, which are often processed via the analytic route.

To re-examine this issue, Experiment 3 revisited the half-shift manipulation introduced in Perea et al. (2023), except that we slightly increased the degree of misalignment (from 1.5 mm to 1.7 mm; e.g. abuelo the Spanish word *sangre* [blood] was presented as), thus eliminating the vertical continuity of vertical letter features (e.g. the letter "n" in *sangre*), breaking the potential local diagnostic features of letter detectors. This provides a more stringent test of the previously reported frequency × format interaction: if the frequency-dependent sensitivity to letter misalignment is real, we should replicate it with more extreme visual distortion; alternatively, if the small-sized interaction reported by Perea et al. (2023) had been an empirical anomaly, the distortion costs should be comparable for high- and low-frequency words – note that Perea et al. (2023) cautioned "about drawing firm conclusions from this interaction" (p. 10).

Experiment 3: Shifted letter halves

Method

Participants

A total of 50 native Spanish speakers (mean age = 29.4 years, SD = 4.9; 37 women) were recruited. This sample size yielded 2,000 observations per condition, which exceeds the recommendations for detecting small effects in linear mixed-effects models (Brysbaert & Stevens, 2018). Recruitment procedures, compensation, and ethical approval were identical to those used in Experiments 1 and 2, and all participants provided informed consent prior to beginning the task.

Materials

The materials were the same as in Perea et al. (2023). The stimuli were presented in Calibri font, size 24-point, with black text on a white background. Each item was converted into an image. It could be presented either intact (e.g. abuelo, the Spanish word for *grandfather*) or with the lower half of the item's letters 1.7 mm to the left or right (e.g. SAMSUNG and allianz) –participants could not anticipate the direction of the shift. The shifted items were created using a Python program written for this purpose. We created three lists to counterbalance the items across the three versions. We randomly assigned participants to the three lists.

Procedure

The procedure was the same as in Experiments 1 and 2, as well as in Perea et al. (2023).

Results and discussion

For the analysis of the response times, we excluded incorrect and extremely fast responses (4.1% for both words and nonwords). Table 5 shows the average response times and accuracy in each experimental condition.

For the inferential statistics, we followed a procedure similar to Experiment 1, with an adjustment to the contrast coding to account for the unbalanced design (1/3 intact vs. 2/3 shifted stimuli). Word frequency was coded as in the previous experiments (high-frequency: −0.5; low-frequency: 0.5), while the visual format was coded using weighted effect coding (intact: −0.66; shifted: 0.33) to ensure the intercept represented the grand mean of the design. Table 6 presents the full outputs of the models.

Word data

For the distributional mean μ, participants were faster to respond to high-frequency words than to low-frequency words ($b = 55.63$, 95% CrI [49.77, 61.53]), and to intact words compared to shifted words ($b = 14.01$, 95% CrI [10.00, 17.98]). Regarding the exponential component, β, word-frequency increased the positive skew of the RT distribution ($b = 0.30$, 95% CrI [0.24, 0.37]), whereas the visual format did not credibly affect the tail ($b = 0.03$, 95% CrI [−0.02, 0.09]). No evidence of an

Table 5. Average response times (ms) and % errors (in brackets) for shifted and intact stimuli in Experiment 3.

Stimuli	Intact	Shifted	Distortion effect *(Shifted – Intact)*
Words			
High Frequency	539 (1.60)	556 (1.51)	16 (−0.09)
Low Frequency	597 (7.32)	612 (6.45)	15 (−0.87)
Nonwords	620 (4.00)	643 (3.99)	23 (−0.01)

Table 6. Posterior means (*b*), standard errors (SE), and 95% credible intervals (CrI) for fixed effects from Bayesian linear mixed-effects models (brms) for Experiment 3.

Stimulus/DV	Component	Parameter	Estimate (b)	SE	95% CrI [L, U]
Words					
RTs	Mean (μ)	Intercept	578.05	9.48	[559.31, 596.46]
		Frequency	55.63	3.00	[49.77, 61.53]
		Format	14.01	2.03	[10.00, 17.98]
		Frequency × Format	−3.13	3.71	[−10.43, 4.18]
	Scale (β)	Intercept	4.50	0.05	[4.41, 4.60]
		Frequency	0.30	0.03	[0.24, 0.37]
		Format	0.03	0.03	[−0.02, 0.09]
		Frequency × Format	−0.01	0.06	[−0.12, 0.10]
Accuracy	Log-Odds	Intercept	3.79	0.12	[3.56, 4.04]
		Frequency	−1.61	0.18	[−1.98, −1.27]
		Format	−0.04	0.17	[−0.38, 0.27]
		Frequency × Format	0.15	0.28	[−0.38, 0.74]
Nonwords					
RTs	Mean (μ)	Intercept	633.98	10.45	[613.58, 654.28]
		Format	22.54	2.51	[17.68, 27.55]
	Scale (β)	Intercept	4.58	0.05	[4.48, 4.68]
		Format	0.14	0.03	[0.09, 0.19]
Accuracy	Log-Odds	Intercept	3.87	0.17	[3.55, 4.22]
		Format	−0.07	0.14	[−0.37, 0.20]

interaction between the two factors was found in either the distributional mean μ ($b = -3.13$, 95% CrI [−10.43, 4.18]) or the exponential component β ($b = -0.01$, 95% CrI [−0.12, 0.10]).

In the accuracy model, participants responded more accurately to high-frequency words than to low-frequency words ($b = -1.61$, 95% CrI [−1.98, −1.27]). However, we found no differences in accuracy between intact and shifted words ($b = -0.04$, 95% CrI [−0.38, 0.27]), nor an interaction between the two factors ($b = 0.15$, SE = 0.28, 95% CrI [−0.38, 0.74]).

Nonword data

Participants showed faster response times for intact nonwords compared to shifted nonwords in the distributional mean component ($b = 22.54$, 95% CrI [17.68, 27.55]) and the exponential component ($b = 0.14$, 95% CrI [0.09, 0.19]). In the accuracy data, no differences were found between intact and shifted nonwords ($b = -0.07$, 95% CrI [−0.37, 0.20]).

In sum, as in Experiments 1 and 2, we again found additive effects of visual degradation – this time horizontal shifting – and lexical access difficulty (word frequency). This pattern is consistent with the idea that, when within the tolerance range of the letter detector system, visual distortion can be resolved at a prelexical, letter-based stage, without substantial reliance on lexical information.

One remaining issue is that Perea et al. (2023, Experiment 2), using the same materials but with a smaller misalignment between the upper and lower parts of the items, reported that high-frequency words were slightly more impaired by misalignment than low-frequency words (13 vs. 7 ms, respectively). In the present experiment, where the manipulation was more extreme, any potential interaction between lexical and perceptual factors should, if anything, have been amplified. However, we found no signs of an interaction (the effect of distortion was 16 vs. 15 ms, for high- and low-frequency words, respectively). The consistency of the additive patterns across all three experiments suggests that letter detectors are remarkably resilient to visual degradation.

General discussion

Skilled adult readers can read words with ease even when they appear in atypical visual formats such as abuelo, with rotated letters (handwriting), in low contrast, or when using distorted (e.g. ᑎᑕᖇᕼᑐᑕᔕ) fonts (e.g. see Grainger, 2022; Hannagan et al., 2012; Qiao et al., 2010). This outcome is often attributed to the idea that letter detectors within the word recognition system have developed a high degree of tolerance to visual degradation (e.g. Cohen et al., 2008; Dehaene et al., 2005). However, in situations where the visual input is highly degraded, automatic recognition becomes less efficient, and additional resources may be required to support lexical access, producing a smaller cost of distortion for familiar high-frequency words than for low-frequency words (see Barnhart & Goldinger, 2010; Dehaene & Cohen, 2007; Fernández-López et al., 2023; Perea et al., 2023; Vergara-Martínez et al., 2021).

In Experiments 1 and 2, we introduced a novel manipulation that degrades key visual features of letters via strikethroughs (Experiment 1 and Experiment 2). To assess whether the effects of the visual distortion introduced by the strikethroughs primarily affect

processing strictly at an encoding, prelexical level, or also involve higher-level lexical processes, we orthogonally manipulated word frequency (comparing the effects of strikethroughs on low- and high-frequency words). Both experiments revealed a cost of letter distortion in the latency data. Critically, however, the magnitude of the visual degradation effect was virtually the same for low- and high-frequency words in both experiments. This additive pattern suggests that the cost associated with strikethroughs is resolved at a prelexical stage of processing, likely at the level of letter feature extraction or identification, without requiring increased lexical involvement during word recognition. This is consistent with the built-in noise tolerance assumed for letter detectors (Cohen et al., 2008). In other words, we found no evidence that lexical-level representations compensated for the visual degradation, even when the degradation was relatively substantial (e.g. the double strikethrough). An important cue is revealed by the ex-Gaussian components. The double strikethrough manipulation did not solely shift the bulk of the distribution (μ), it also increased the right tail (β) for double strikethrough words. In functional terms, this pattern suggests that, for a subset of trials, early orthographic normalisation took longer to settle, presumably because the available letter-level evidence was temporarily too noisy or too inconsistent. Critically, this slowdown remained additive with word frequency, suggesting that the distortion was prolonging prelexical evidence formation without inducing additional lexical involvement.

Given the clear pattern of results from Experiments 1 and 2, showing consistent additive effects of visual degradation and word frequency, we sought to re-examine the letter-half manipulation introduced by Perea et al. (2023, Experiment 2), where high-frequency words showed greater impairment than low-frequency words when the upper and lower halves of the words were horizontally shifted (13 ms for high-frequency words and 7 ms for low-frequency words). Specifically, if the interaction they observed truly reflects reliance on holistic representations, as they proposed, then increasing the degree of misalignment should amplify the interaction effect. However, in Experiment 3, using a more extreme manipulation of misalignment than Perea et al. (2023), we again found additive effects of word-frequency and distortion (the cost was 16 ms for high-frequency words, and 15 ms for low-frequency words), in line with those obtained in Experiments 1 and 2, and consistent with an early locus of the distortion effect. This finding is important because the interaction reported by Perea et al. (2023) was numerically small (a 6-ms difference in cost between high- and low-frequency words), and those authors cautioned against drawing firm conclusions from it. Thus, Experiment 3 examined whether this pattern would generalise under stronger perceptual distortion. It did not; increasing the degree of misalignment yielded virtually identical distortion costs for high- and low-frequency words.

Consequently, our findings place tight constraints on hybrid models that assume a shift to holistic shape processing for high-frequency words (Allen et al., 1995). Any such model must accommodate the additivity between word frequency and visual distortion observed here, even under manipulations that strongly perturb letter features and the supraletter envelope. Instead, the additive effect across experiments reinforces the generality of a highly noise-tolerant, letter-based route of word recognition.

Notably, the additivity we found between distortion and frequency in the three experiments echoes the additivity between stimulus-quality (via low contrast or alternating masks) and word frequency in lexical decision (see Stanners et al., 1975; Yap & Balota, 2007). The logic of these experiments is, to some degree, parallel to that employed here, namely, whether reductions in the quality of letter-level evidence require increased lexical involvement. In this sense, the current manipulation of visual degradation using strikethroughs and half-shifts pushes the reading system into a regime in which early orthographic normalisation incurs a cost, but before lexical frequency exerts its influence.

Taken together, the evidence from the three lexical decision experiments reported here converges on the conclusion that physical distortions of letter features and letter alignment primarily affect prelexical processing, with no behavioural indication that lexical-level information (via the word-frequency effect) selectively contributes to normalising these distortions under the present conditions. These results can be parsimoniously explained by hierarchical letter-detector models in which successive layers of visual detectors exhibit substantial tolerance to visual noise (Dehaene et al., 2005; Grainger et al., 2008). Both the strikethrough and shifting manipulations introduce spurious junctions and partial occlusions of letter strokes, but the reading cost remained relatively small and remarkably uniform regardless of word frequency (approximately 22-23 ms for single strikethroughs, 33 ms for double strikethroughs, and 15-16 ms for shifted halves). This pattern suggests that the system of letter detectors possesses sufficient built-in tolerance to resolve feature degradation via predominantly feedforward processing (see Vergara-Martínez et al., 2021, for ERP evidence with easy-to-read handwritten words).

We acknowledge, however, that the absence of a visual distortion × word-frequency interaction in behavioural measures does not imply that lexical factors never modulate the processing of degraded input. Time-resolved measures (e.g. event-related potentials) can reveal a transient distortion × word-frequency interaction even when response times are additive (e.g. see Civera et al., 2026, for evidence comparing intact words vs. upper-only words). This point is also illustrated by Grainger et al. (2012), who showed that additive or weakly interactive patterns in behavioural measures can coexist with temporally specific modulations in ERP components during visual word recognition. Thus, the present additivity in response times should be interpreted as a behavioural constraint on the locus of the distortion cost, rather than as evidence that perceptual and lexical variables never interact during the unfolding of word recognition. A similar pattern could occur here, with short-lived frequency-dependent sensitivity to distortion that does not survive early word processing stages. However, this is an issue for future research, beyond the scope of the present experiments. The current findings provide a clear behavioural constraint for models of visual word recognition, showing that moderate feature occlusion and misalignment in a standard lexical decision task (i.e. the most commonly used paradigm in visual word recognition research) produce a reading cost that is approximately the same for high- and low-frequency words.

To sum up, our findings indicate that moderate distortions, whether through feature occlusion (e.g. strike-throughs) or spatial perturbation (e.g. horizontal displacement of the letter halves), are largely accommodated during early encoding. This pattern is consistent with the view that the feedforward sweep of orthographic processing is remarkably tolerant to noise, as proposed by neurobiologically inspired models of letter and word recognition (Dehaene et al., 2005; Grainger et al., 2008). At the same time, these results help delineate when lexical-level activation contributes to perceptual normalisation, and when automatic reading can be achieved based on bottom-up evidence supporting abstract letter identities. More generally, our findings favour the view that orthographic processing constitutes a critical mid-level visual interface between early visual analysis and linguistic processing during reading (Grainger, 2017).

Acknowledgements

The present research was supported by a Grant from the Spanish Ministry of Science, Innovation, and Universities (PID2023-152078NB-I00) to Manuel Perea, and Grants from Department of Education, Culture, Universities, and Employment of the Valencian Government (CIAICO/2024/180 to Manuel Perea, CIGE/2023/20 to Maria Fernández-López, and CIAPOS/2024/171 to Melanie Labusch).

Disclosure statement

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

Funding

This work was supported by Generalitat Valenciana: [Grant Number CIGE/2023/20 and CIAICO/2024/180]; Ministerio de Ciencia, Innovación y Universidades: [Grant Number PID2023-152078NB-I00].

Data availability statement

Experiments 1 and 2 were pre-registered. The data, scripts and full output that support the findings of this study are openly available on the Open Science Framework (OSF) at https://osf.io/bxmc6/overview?view_only=1cbf9ea3ce3a46b499cc98d49f9c0538.

ORCID

Maria Fernández-López http://orcid.org/0000-0001-7419-7369

References

Allen, P. A., Wallace, B., & Weber, T. A. (1995). Influence of case type, word frequency, and exposure duration on visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 914–934. https://doi.org/10.1037/0096-1523.21.4.914

Balota, D. A., Aschenbrenner, A. J., & Yap, M. J. (2013). Additive effects of word frequency and stimulus quality: The influence of trial history and data transformations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1563–1571. https://doi.org/10.1037/a0032186

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 340–357. https://doi.org/10.1037/0096-1523.10.3.340

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*, 160–166. https://doi.org/10.1177/0963721411408885

Barnhart, A. S., & Goldinger, S. D. (2010). Interpreting chicken-scratch: Lexical access for handwritten words. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 906–923. https://doi.org/10.1037/a0019258

Besner, D., & Young, T. (2022). On the joint effects of stimulus quality and word frequency in lexical decision: Conditions that promote staged versus cascaded processing. *Canadian Journal of Experimental Psychology*, *76*, 122–131. https://doi.org/10.1037/cep0000266

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, *27*, 45–50. https://doi.org/10.1177/0963721417727521

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*, 1–20. https://doi.org/10.5334/joc.10

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*, 395. https://doi.org/10.32614/RJ-2018-017

Chen, C., Abbasi, N. u. H., Song, S., Chen, J., & Li, H. (2016). The limited impact of exposure duration on holistic word processing. *Frontiers in Psychology*, *7*, 646. https://doi.org/10.3389/fpsyg.2016.00646

Civera, T., Perea, M., Comesaña, M., Gutierrez-Sigut, E., & Vergara-Martínez, M. (2026). Neural signatures of perceptual closure and top-down feedback in lexical access: An ERP study. *Language, Cognition, and Neuroscience*, *41*, 175–193. https://doi.org/10.1080/23273798.2025.2601156

Cohen, L., & Dehaene, S. (2010). Specialization within the ventral stream: The case for the visual word form area. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences (4th ed.)* (pp. 551–562). MIT Press.

Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *NeuroImage*, *40*, 353–366. https://doi.org/10.1016/j.neuroimage.2007.11.036

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

Davis, C. J. (2001). The self-organising lexical acquisition and recognition (SOLAR) model of visual word recognition (Doctoral Dissertation). University of New South Wales, Sydney, Australia.

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

Dehaene, S. (2009). *Reading in the brain: The science and evolution of a human invention (Vol. 7)*. Viking.

Dehaene, S., & Cohen, L. (2007). Response to Carreiras et al.: The role of visual similarity, feedforward, feedback and lateral pathways in reading. *Trends in Cognitive Sciences*, *11*, 456–457. https://doi.org/10.1016/j.tics.2007.08.009

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

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

Fernández-López, M., Gomez, P., & Perea, M. (2023). Letter rotations: Through the magnifying glass and what evidence found there. *Language, Cognition, and Neuroscience*, *38*, 127–138. https://doi.org/10.1080/23273798.2022.2093390

Fernández-López, M., Solaja, O., Crepaldi, D., & Perea, M. (2025). Top-down feedback normalizes distortion in early visual word recognition: Insights from masked priming. *Psychonomic Bulletin & Review*, *32*, 920–929. https://doi.org/10.3758/s13423-024-02585-2

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

Grainger, J. (2017). Orthographic processing: A "mid-level" vision of reading. *Quarterly Journal of Experimental Psychology*, *70*, 1003–1016. https://doi.org/10.1080/17470218.2017.1314515

Grainger, J. (2022). Word recognition I: Visual and orthographic processing. In M. J. Snowling, C. Hulme, & K. Nation (Eds.), *The science of reading (2nd ed.)* (pp. 60–78). Wiley. https://doi.org/10.1002/9781119705116.ch3

Grainger, J., & Hannagan, T. (2014). What is special about orthographic processing? *Written Language & Literacy*, *17*, 225–252. https://doi.org/10.1075/wll.17.2.03gra

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

Grainger, J., Lopez, D., Eddy, M., Dufau, S., & Holcomb, P. J. (2012). How word frequency modulates masked repetition priming: An ERP investigation. *Psychophysiology*, *49*(5), 604–616. https://doi.org/10.1111/j.1469-8986.2011.01337.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

Hannagan, T., Ktori, M., Chanceaux, M., & Grainger, J. (2012). Deciphering CAPTCHAs: What a Turing test reveals about human cognition. *PLoS One*, *7*, e32121. https://doi.org/10.1371/journal.pone.0032121

Harel, A. (2016). What is special about expertise? Visual expertise reveals the interactive nature of real-world object recognition. *Neuropsychologia*, *83*, 88–99. https://doi.org/10.1016/j.neuropsychologia.2015.06.004

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin*, *109*, 340–347. https://doi.org/10.1037/0033-2909.109.2.340

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*, 627–633. https://doi.org/10.3758/BRM.42.3.627

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*, 293–299. https://doi.org/10.1016/S1364-6613(03)00134-7

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

Peirce, J. W., Hirst, R. J., & MacAskill, M. R. (2022). *Building experiments in PsychoPy*. Sage.

Perea, M., Gil-López, C., Beléndez, V., & Carreiras, M. (2016). Do handwritten words magnify lexical effects in visual word recognition? *Quarterly Journal of Experimental Psychology*, *69*, 1631–1647. https://doi.org/10.1080/17470218.2015.1091016

Perea, M., Romero-Ortells, I., Labusch, M., Fernández-López, M., & Marcet, A. (2023). Examining letter detector tolerance through offset letter halves: Evidence from lexical decision. *Journal of Cognition*, *6*, 56. https://doi.org/10.5334/joc.322

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

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, *114*, 273–315. https://doi.org/10.1037/0033-295X.114.2.273.

Qiao, E., Vinckier, F., Szwed, M., Naccache, L., Cohen, L., & Dehaene, S. (2010). Unconsciously deciphering handwriting: Subliminal invariance for handwritten words in the visual word form area. *NeuroImage*, *49*, 1786–1799. https://doi.org/10.1016/j.neuroimage.2009.09.034

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, *111*, 159–182. https://doi.org/10.1037/0033-295X.111.1.159

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*, 191–201. https://doi.org/10.3758/BF03197692

R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.r-project.org/.

Stanners, R. F., Jastrzembski, J. E., & Westbrook, A. (1975). Frequency and visual quality in a word-nonword classification task. *Journal of Verbal Learning and Verbal Behavior*, *14*, 259–264. https://doi.org/10.1016/S0022-5371(75)80069-7

Ventura, P., Fernandes, T., Pereira, A., Guerreiro, J. C., Farinha-Fernandes, A., Delgado, J., Ferreira, M. F., Faustino, B., Raposo, I., & Wong, A. C.-N. (2020). Holistic word processing is correlated with efficiency in visual word recognition. *Attention, Perception, & Psychophysics*, *82*, 2739–2750. https://doi.org/10.3758/s13414-020-01988-2

Ventura, P., Pereira, A., Xufre, E., Pereira, M., Ribeiro, S., Ferreira, I., Madeira, M., Martins, A., & Domingues, M. (2019). Holistic word context does not influence holistic processing of artificial objects in an interleaved composite task. *Attention, Perception, & Psychophysics*, *81*, 1767–1780. https://doi.org/10.3758/s13414-019-01812-6

Vergara-Martínez, M., Gutierrez-Sigut, E., Perea, M., Gil-López, C., & Carreiras, M. (2021). The time course of processing handwritten words: An ERP investigation. *Neuropsychologia*, *159*, 107924. https://doi.org/10.1016/j.neuropsychologia.2021.107924

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 274–296. https://doi.org/10.1037/0278-7393.33.2.274