Pre-final version

The role of letter features in visual-word recognition: Evidence from a delayed segment technique

Eva Rosa [1], Manuel Perea [2 3], and Peter Enneson [4]

[1] Universidad Católica de Valencia, Valencia, Spain

[2] Universitat de València, Valencia, Spain

[3] BCBL. Basque Center on Cognition, Brain, and Language, Donostia, Spain

[4] Peter Enneson Design Inc, Toronto, Ontario, Canada

Running head: visual features in word recognition

Word count: 7527 (main text)

Address for correspondence:

Eva Rosa

Departamento de Neuropsicobiología, Metodología y Psicología Social

Universidad Católica de Valencia

C/ Guillem de Castro, 175

46008, Valencia, Spain

Email: eva.rosa@ucv.es

Author's notes. The research reported in this article has been partially supported by Grant PSI2014-53444-P from the Spanish Ministry of Economy and Competitiveness. We thank William Berkson for manipulating the fonts to provide us with our stimuli, and for discussion of scientific methodology and terminology.

Abstract

Do all visual features in a word's constituent letters have the same importance during lexical access? Here we examined whether some components of a word's letters (midsegments, junctions, terminals) are more important than others. To that end, we conducted two lexical decision experiments using a delayed segment technique with lowercase stimuli. In this technique a partial preview appears for 50 ms and is immediately followed by the target item. In Experiment 1, the partial preview was composed of terminals+junctions, midsegments+junctions, or midsegments+terminals — a whole preview condition was used as a control. Results only revealed an advantage of the whole preview condition over the other three conditions. In Experiment 2, the partial preview was composed of the whole word except for the deletion of midsegments, junctions, or terminals — we again employed a whole preview condition as a control. Results showed the following pattern in the latency data: whole preview = delay of terminals < delay of junctions < delay of midsegments. Thus, some components of a word's constituent letters are more critical for word identification than others. We examine how the present findings help adjust current models of visual word identification or develop new ones.

Key words: visual-word recognition, priming, lexical decision, letter processing

In cognitive psychology today, a formidable consensus now exists that 1) a parallel letter recognition process involving explicit labelling at the letter level and 2) a mapping of these labelled entities onto abstract letter units mediates visual-word recognition. In line with this, a fundamental goal of computational models of visual-word recognition has been to specify in detail how, and in what sense — implicitly as feature conjunctions (footnote 1), or explicitly as labelled entities — the words' constituent letters are extracted from the visual features.

In the past years, there has been significant progress in our understanding of the response properties of the various layers in the visual and inferior temporal cortex. Sophisticated computational attempts to model orthographic processing and lexical access have been put forward and fitted to data. But *whether* or *to what extent* — in neuro-physical and cognitive-processing terms — explicit labelling occurs, or *where* and *how* the mapping unto abstract letter identities is attained, or *what* — in perceptual processing terms — the key components of the letter are during visual word recognition and reading, remains unclear. Theoretical framing options for staking-out the *what* and tracking-down the *how*, are many-fold and discrepant. Here we focus on the *what*, and try to establish the relative importance for *visual word recognition* of several components that have been presumed in previous experiments to be key components.

Many current computational models of visual-word recognition employ three processing levels: letter features, letters, and words (e.g., see Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Davis, 2010; Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981; Perry, Ziegler, & Zorzi, 2007; see also Carreiras, Armstrong, Perea, & Frost, 2014, for a recent review). But despite the intrinsic relevance of the widely acknowledged "feature detection" (feature-analytic) and "feature integration" processes at the perceptual processing front end, models of visual-word recognition have focused most assiduously on the intricacies of the hypothesized letter-level and word-level processing further downstream, and been satisfied to leave the structural particularities of the "letter feature level" (e.g., see Finkbeiner & Coltheart, 2009, for discussion), and the microprocesses occurring at the perceptual processing front end to remain comparatively under-specified. In computational modelling *per se* the prevailing strategy has been to “take a leap of faith and assume we have made it to the letter” where the computational prospect might be “a bit more tractable” (Balota, Yap, & Cortese, 2006, p. 289).

As a concomitant of this, all the above-cited models employ the — highly artificial from a typographical point of view — 14-feature uppercase-letter font created by Rumelhart and Siple (1974). In this font, the critical features of the letters correspond to straight-line segments or “quanta” that are location-specific and determinate with respect to orientation. Each of these oriented segments is numbered according to its position, so that the letter A would be represented by the binary feature pattern: 11111010100000 (see Figure 1).

Davis (2010) noted, “McClelland and Rumelhart’s (1981, p. 383) assumption that ‘the basic results do not depend on the font used’ seems like a reasonable starting point” (p. 725), but as Mewhort and Johns (1988, p. 139) point out, the Rumelhart /Siple scheme leaves the computational representation of the alphabetic system vastly overdetermined. So, while this assumption might be heuristically valuable as an exploratory principle, the artificiality of the scheme might not be inconsequential: it might not provide an operationally viable proxy for how the visual system actually breaks down the stimuli used in reading. Essentially, over-determination at the feature level might, for example, skew a calculation of the “capacity benefits” resulting from orthographic neighbourhood effects (see Houpt & Townsend, 2014, for a discussion of capacity benefits in visual word-recognition).

Thus, an unresolved issue for constructing realistic computational models of visual word-recognition is: what, if any, domain-specific perceptual processing primitives are critical in visual-word recognition.

Recently, in a connectionist computational model with a backpropagation routine by which the components a hidden layer between actual bitmapped stimuli and real words are constructed and revealed, presumably during letter level processing, Chang, Furber, and Welbourne (2012) used principal component analysis to define a set of eight crucial features spanning upper and lower case letters. The features identified in the Chang et al. (2012) model have some simple circular and angular shapes as well as combinations of line segments and direct line segments, arguably encompassing structural letter parts and relational. The list distinguishes a category of round curve-shaped features (as in G, O, and U); an n-shaped feature (A, K, R, X; a, h, k, n); a vertical center line feature; an inverted L-shaped feature; a v- shaped feature; a c-like feature; a hook-shaped feature; and a repeated vertical strokes feature. An unresolved issue in this account surrounds the fact that running the routines suggest that 50 units in the hidden layer seems to give the best fit to behavioural results, yet, only 8 features are identified and freely interpreted in the principal component analysis.

Over the last five or six decades, in psychophysical, behavioural and neurophysical, studies outside of the strictly computational modelling environment, various other — often incommensurate (though perhaps complimentary) — classes of "feature-level" operators have been proposed as candidates for what the perceptual processing primitives in letter identification and word recognition are. For example, edges or boundaries between light and dark, oriented bars and annular forms (see Hubel & Wiesel, 1959); aggregated segments of varying orientations and curvatures (Gibson, Gibson, Pick, & Osser, 1962; Gibson, 1965; Smith, 1969) — or discriminant parts and distinctive features of these segments, and their junctions (Fiset, Éthier-Majche, Arguin, Bub, & Gosselin, 2008; Petit & Grainger, 2002; Lanthier, Risko, Stolz, & Besner, 2009) —global features of letter wholes (Bouma, 1970; Chang, Furber, & Welbourne, 2012).

It appears then that there is too little consensus on the "identity" of the "key components" issue — the proper resolution of which the "relative importance," or "role" question appears to require.

The "role" issue

The identity and relative importance of different visual constituents in perceptual processing has been investigated previously in the area of object recognition (e.g., see Biederman, 1987). In the object recognition domain, Biederman proposed that, though the underlying visual processing is feature-based, object recognition is mediated by a segmentation into parts of an "intermediate complexity" between simple features and independent wholes. The segmentation and recognitional process occurs on the basis of structural and relational features of the input image. In the Biederman (1987) experiments, participants had to identify line-drawn three-dimensional pictorial representations of objects with midsegments or vertices deleted. Results revealed that the removal of vertices was more detrimental to object recognition than the removal of midsegments. However, one needs to be cautious at generalizing these findings to letter/word recognition. As Petit and Grainger (2002) indicated, "two-dimensional letter shapes are not segmented in a manner analogous to three-dimensional objects." (p. 352)

The literature on the role of vertices vs. midsegments vs. junctions in letter/word recognition is sparse. In a pioneering work, Petit and Grainger (2002) employed a masked prime paradigm using briefly presented, partial-letter primes that were followed by the target stimuli to determine which parts of letters play a critical role in the process of letter perception. Their experiments used letter naming and alphabetic decision tasks. The partial primes were created by deleting parts at different regions of the target letters and were composed of the same number of pixels in each condition: i) local segmental junction primes were composed of the pixels around the intersection between two lines plus pixels at the ends of the lines; ii) local segmental midsegment (or junctionless) primes were composed of pixels at regions between junctions; iii) global primes were constructed with pixels randomly distributed across the entire target; and iv) neutral primes were constructed with pixels randomly distributed across the rectangular space that a complete version of the prime would occupy (see Figure 1 in Petit & Grainger, 2002). Eighteen letters of the Roman alphabet in upper-case format (font: Courier, 24 pt) were used. Petit and Grainger found a significant advantage for the target letters when preceded by a global prime than when preceded by a junction or neutral prime (Experiment 1), but note that this experiment did not include the "midsegment" condition. In addition, the letter-naming task showed faster response times to targets when preceded by complete and junction primes compared with the targets preceded by a neutral prime. Importantly, in Experiments 2 and 3 Petit and Grainger found an advantage for the targets preceded by a midsegment prime over the targets preceded by a junction prime.

Other experimentalists have also used partially degraded stimuli in which some parts of the letters were removed (deletion) until the participant's response, but with different results.

Lanthier, Risko, Stolz, and Besner (2009) examined the relative importance of different letter features in recognition by comparing three experimental conditions in a naming task: intact, midsegment deletion and vertex deletion (see Figure 1 in Lanthier et al., 2009). Twenty-three upper-case only letters, set in 27-point Arial Narrow were used. When naming isolated letters, they found that removing the vertices was more detrimental than removing midsegments (Experiments 1 and 2). The same pattern occurred when participants had to name words with a very short presentation time (50 ms; Experiment 4), but not when these words were presented for unlimited time (Experiment 3).

At about the same time as the Lanthier et al. (2009) work, Fiset et al. (2008, 2009) used a classification image technique, *Bubbles*, in conjunction with an *ad hoc* and *a priori* list of featurally specific areas of upper and lower case letter forms to make inferences about which features of the letter were most important for letter identification. During each trial, several masks were successively placed over the stimuli, isolated letters, to modulate the availability of visual information across time (see Figure 1 in Fiset et al., 2009). Although the *Bubbles* method is pixel-based and does not require a priori definition of what the features for letter identification are, this allowed them to dissociate and compare salient and readily isolatable components of the letterform that have real and documentable cue value. All 26 letters of the Roman alphabet, displayed in lower-case Arial (152 points) and uppercase (117 points) were used. They found that terminations were the most important features for uppercase and lowercase letter identification, followed by horizontal lines. Note however, that in their tabulation, they didn’t treat midsegments as a group but treated each mid-segment type (example: horizontal lines) as on a par with the entire terminals group.

All of the above sets of experiments looked at the role that parts of letters play in letter perception, discrimination and identification outside of the context of words. Lanthier et al. (2009) also looked at the role that parts of letters play in word perception, and found identical results to their letter perception results.

It may be important to note that in a related Blais et al. (2009) study, the Bubbles technique was again employed, this time to study the information extraction strategy underlying the human efficiency at word recognition. Sample sizes consistent with the sample sizes used to draw conclusions about terminations, junctions and mid-segments in the Fiset et al. (2008, 2009) studies were employed, and results were calibrated in the time domain and according to spatial location, with the results in the spatial domain coarsely tabulated according to information in the letter body, ascender or descender zone, rather than according to feature type. Stimuli were 40 pt Courier lowercase words. Here, in contrast to the Lanthier et al. (2009) results for words, the graphic presentation of the Bubbles results suggest that the dynamics of information extraction in visual letter discrimination and identification might *differ* — when it comes to local segmental information — from the dynamics of information extraction and integration in visual word recognition.

Szwed et al. (2009) conducted three experiments in order to test if visual word recognition might be based on pre-existing mechanisms common to all visual object recognition. In a naming task, they presented partially deleted pictures of objects and printed words in which either the vertices or the line midsegments were preserved. The stimuli were of line drawings of objects (natural and artifacts); and 6–8 letter French words, made either exclusively or predominantly of 'non-curvy' letters, in an uppercase sans serif font with thin lines. The results showed that subjects made significantly less naming errors for objects presented in the vertex-preserved variant than in the midsegment-preserved variant. Reaction times showed a parallel tendency, although the effect was not significant. The pattern was identical with words (i.e., the subjects made fewer errors and were faster to respond when vertices were preserved).

In a study looking at specialization for written words over objects in the visual cortex, Szwed et al. (2011) employed degraded words in a one-back repetition detection task and an overt naming task. The words were degraded by partial deletion of some of their component lines (see Figure 1 in Szwed et al., 2011). The following conditions were tested: vertex-preserved versus midsegment-preserved, and two types of control stimuli were used: scrambled (randomly scrambled fragments of the word) and “gestalts” (fragments recomposed into pseudo-objects that have the same amount of collinearity and grouping as words). The results in the one-back repetition detection task revealed that the hit rate did not differ between vertex-preserved and midsegment-preserved words. However, in the naming task, the participants' accuracy was higher for the vertex-preserved stimuli than for the midsegment-preserved stimuli. Szwed et al. concluded that, as in object recognition, junctions play a particular role at the whole word level in reading. Szwed et al. also employed line drawings of objects as stimuli with the same conditions, and obtained a parallel finding. Finally, the fMRI data from Szwed et al. revealed that a restricted part of the object perception system in the left fusiform gyrus was sensitive to the presence of vertices and this area overlapped partially with the so-called “visual-word form area”.

In sum, the above-cited findings suggest that some localizable structural letter parts — or the information located there and at their junctions — might be more important than others during letter identification and/or word recognition, but the divergences across experimentally constructed stimuli, tasks and calibration procedures make it difficult to establish firm conclusions about the identity of the structural parts, and how relational features come into play. At this stage in the development of our understanding of processing at the front end, the issue of relative importance is largely empirical and testing has been exploratory. Theory-based models of feature pooling, averaging, summation or convolution haven't been developed with a view to making predictions or testing alternative theories that seek to dissociate and quantify the effects of delaying or deleting different sorts of localizable letter parts”.

In the present experiments, we examined the relative importance of mid-segments, junctions, and terminals of lower-case letters embedded in words. We decided to select mid-segments and junctions because, in previous literature on the “role” or “importance” or “potency” issue, they are the most frequently cited as important constituents of the letters in a word — note however that empirical evidence about the importance of each of them is not conclusive. Terminals were included — in contrast to the Petit and Grainger (2002), Lanthier et al. (2009), and Szwed et al. (2009, 2011) work — because of the importance inferred for them from the

Bubbles experiments. Unlike the features identified in the Chang, Furber, and Welbourne (2012) work — which are distinctive to discrete subsets of letters, so have an importance or crucial role only in those in which they occur — junctions (with some exceptions), midsegments and terminals are feature-bearing components of all letters. And in contrast to the Fiset, et. al. lists, where horizontal components, curved components of different kinds, vertical components and slanted components are treated analytically as on a par with "terminals" and "intersections" — while they are in fact "components" of a different order —, a focus on *midsegments* alongside of terminations and junctions appears to us more consistent from a "compositional anatomy of type" point of view. All the strokes in every letter have distinctive features on their paths, but not every letter has a curved or angled or vertical, etc., stroke.

Note, that all previous studies addressing the "role" or "importance" question employed (the less familiar) uppercase letters as stimuli (the only exceptions were the experiments of Fiset et al., 2008, and Blais et al., 2009, with the Bubbles technique), and most focused on letters in isolation, rather than embedded in words. A complete account would examine and compare both uppercase and lower case letters, and it would look at asymmetries in performance at a letter level versus at a word level.

Before describing the experiments, it may be appropriate to define some of the concepts used in this paper. Hinton (1981) thought of the quantized items in the "feature" layer of the computation models he helped pioneer as proxies for "stroke-units" or in the words of Franklin (1995) as "stroke-representing units". In the realm of writing, and by extension, type (Noordzij, 2005), stroke-units are readily identifiable as dimensional "swept-object" structures (Parida & Mudur, 1994) with unique and tractable gestural signatures. In cognitive science and perceptual psychology a functional ontology (see Price & Friston, 2005) of basic units below the letter level has, as indicated above, not so far been definitively formalized, and the tendency has been toward phanemic (i.e., pattern or appearance based), rather than kinemic (i.e., hand-writing or program based) description (see Watt, 1988). Furthermore, we recognize that the use of feature language in perceptual psychophysics, cognitive psychology and the theorizing around computational modelling is not optimally disciplined (see Feldman, 2015, for a review of various uses), and "visual word recognition" is a fuzzy generic construct, readily adapted from everyday usage, but not specific to the perceptual processing front end. For our presentation of the present experiments, and in the discussions that follow we employ the notions that functionally, in letter perception, orthographic processing and/or visual word-form resolution: i) terminations and midsegments are localizable *structural components* of *stroke-units*, and ii) that junctions or vertices are involved in a *local combination detection* such as is proposed by Dehaene, Cohen, Sigman, and Vinckier (2005). Local combination detection at a stroke-unit level functions on the basis of *relational features*, such as crossing or abutment, at terminations or mid-segments. As structural components of stroke-units, terminations and mid-segments carry information about *structural features* of stroke-units, such as extendedness, expressedness and closure. So in this scheme, the parts of letters — stroke-units — and their structural components — terminations and mid-segments — that the items in the so-called feature layers of computational accounts appear to stand as proxy for, *have* features or are featurally distinct, but are not themselves features. In conjunction with this, the "standard model" of object recognition proposes (see Põder, 2014): 1) a sampling by local feature detectors; 2) a pooling the output of these detectors over some second-level receptive fields and combining the results into second-order features, perhaps not unlike Biedermann's "structures of an intermediate complexity".

The delayed segment technique

In our study, we conducted two lexical decision experiments using a delayed segment technique with lowercase stimuli. In this technique, a partial preview appears for 50 ms and is immediately followed by the whole target item (Perea, Comesaña, & Soares, 2012; see also Carreiras, Gillon-Dowens, Vergara, & Perea, 2008; Lee, Rayner, & Pollatsek, 2001, for previous experiments with this technique). Compared with the paradigms employed in previous experiments, this constitutes a closer situation to the actual process of lexical access in normal reading because letters were presented in lowercase, embedded in words, and remained unaltered (i.e., non-degraded) except for the initial 50 ms of exposure. In contrast to lexical decision experiments that set out to orthographic processing and lexical access at a more abstract or cognitive level by using a cross-case priming technique, the delayed segment technique has the potential of addressing perceptual processing at the front end.

In Experiment 1, we compared four types of previews (see Figure 2 for illustration): i) the target itself (the whole preview as a control condition); ii) a partial preview composed of terminals and junctions; iii) a partial preview composed of some midsegments and junctions; and iv) a partial preview composed of some midsegments and terminals. Care was taken so that terminals, junctions and mid-segments were selected and sampled in a principled way and had approximately the same number of pixels. (This however may lead to some variations in pixel counts with more complex letter forms. and when letters are assembled into words.) The midsegment cuts were made at locations on the stroke that are thought to carrying distinctive information about the stroke's identity and may reflect the influence of prior studies of cue-value. The previews in Experiment 1 were composed of combinations of two feature types (midsegments plus terminals) because on review, we anticipated presenting just one feature type would probably fail to activate the stroke-unit-level or letter-level representations necessary for lexical retrieval.

Experiment 2 was prompted by the lack of different results for the three conditions in Experiment 1, which the standard sampling/pooling model and previous experiments had led us to expect. In Experiment 2, we employed a complementary strategy: The partial previews consisted of the whole word except for the deletion of pixels corresponding to the midsegments, junctions or terminals depending on the experimental condition (see Figure 3 for illustration). In this case, care was taken so that terminals, junctions and mid-segments *deletions* (rather than the sampling as in Experiment 1) were done in a principled way, and contained approximately the same number of pixels per deletion. The amount of pixels deleted in Experiment 2 visually relates to the number of pixels sampled in each location in Experiment 1. The effect is that the co-linearity of sub-sets of the sampled information is more evident, and the perceptual integrity of individual stroke units becomes more pronounced. It could thus be argued that the different conditions of Experiment 1 and 2 are equivalent with regard to the feature of the letter affected, and that the real difference between the two experiments is the percentage of pixels of the preview (~50% in Experiment 1 vs. ~75% in Experiment 2), however the effects are visually different, and as we shall see, the anticipated success of the sampling strategy of Experiment 1 might rely on a too simplistic "summation" or "pooling" assumption which doesn't take into account the higher order bottleneck we identify, a bottleneck which we propose the difference in results between the two experiments might reveal. If the sampling or removal of midsegments, junctions and terminals does not affect participant's performance in the same way across sampling levels and across conditions, visual-word recognition models should be able to explain the different importance of each component and the reason for the difference at different sampling levels in a principled way.

Experiment 1

Method

Participants. Thirty-two undergraduate students from the University of Valencia participated in the experiment in exchange of extra course credit. All of them were native speakers of Spanish and had normal or corrected-to-normal vision. None of them reported having any speech/reading problems.

Materials. We selected 240 words from the Spanish database B-Pal (Davis & Perea, 2005). The mean length was 6.2 letters (range: 5-8), and the mean word-frequency was 27.5 occurrences per million (range: 0.2-383.6). In addition, a set of 240 nonwords were created for the purposes of the lexical decision task; these nonwords had been created by changing one or two interior letters of words extracted from the same database. Nonwords were all pronounceable and orthographically correct, and they were equated to words in length. None of the words (or nonwords) had letters with accent/diacritical marks (e.g., "á"), punctuation marks (as in "ñ"), or letters without junctions ("c", "i", "o", "j" or "s"). The list of words and nonwords is presented in the Appendix. Four previews were created from each target (see Figure 2): i) the target itself (i.e., the whole preview condition); ii) a partial preview composed of terminals and junctions (Terminals+Junctions condition); iii) a partial preview composed of some midsegments and junctions (Midsegments+Junctions condition); and iv) a partial preview composed of some midsegments and terminals (Midsegments+Terminals condition). The guideline for preparing the letters in the Terminals+Junctions, Midsegments+Junctions, and Midsegments+Terminals conditions was to make it so that each terminal, junction and mid-segment in every letter contained approximately the same number of pixels. The previews were planned and created from a Minion font by font-design professionals using FontLab Studio software (available at http://www.fontlab.com/font-editor/fontlab-studio/). This strategy resulted in previews that retained about 50 percent of the stimulus information, as gauged by pixel counts, compared to the no-delay condition.

Procedure. Participants were tested in a quiet, well-lit room. Presentation of stimuli and collection of responses were controlled with DMDX software (Forster & Forster, 2003). Each trial started with the presentation of a fixation signal (+) for 500 ms. Then, the preview was presented for 50 ms and was immediately followed by the target. The 50-ms previews could be a terminal+junction preview (120 words and 120 nonwords), a midsegment+junction preview (120 words and 120 nonwords), a midsegment+terminal preview (120 words and 120 nonwords), or the full stimulus (120 words and 120 nonwords). RTs were measured from the onset of the target until the participant's response (see Figure 4). All stimuli were presented in the same spatial location. The target remained on the screen until the participant made a response or 2 sec had passed. Participants were instructed to press the "sí" [yes] button, with their right hand,when the letter string was a Spanish word and the "no" button, with their left hand, when the letter string was not a word. They were asked to make their responses as rapidly and as accurately as possible. Each participant was given a total of 480 experimental trials: 240 word trials and 240 nonword trials in a different random order. Prior to the experimental phase, the participants received 18 practice trials. The session lasted approximately 20 min.

Results and Discussion

Response times less than 250 or greater than 1500 ms were excluded from latency analyses (1.0 and 2.6% of the data for word and nonword trials, respectively). Mean correct RTs and errors rates were calculated across participants and across items. The ANOVAs were conducted on the basis of a 4 (preview type: Identity, Terminals+Junctions, Midsegments+Junctions, Midsegments+Terminals) x 4 (list: list 1, list 2, list 3, list 4) design. List was included in the analyses to partial out the error variance due to the counterbalancing lists (Pollatsek & Well, 1995). The mean correct RTs and percentage of error for each experimental condition are presented in Table 1.

Word data. There was a significant effect of preview type on RTs to word targets, $F1(3,84) = 39.69$, $p < .001$; $F2(3,708) = 37.71$, $p < .001$ Pairwise comparisons revealed that RTs were substantially shorter for the whole preview condition (627 ms) than in the other three preview conditions, all *ps* < .001, whereas the three partial preview conditions yielded comparable response times (668, 665, and 664 ms; for the midsegments+junctions, midsegments+terminals and terminals+junctions previews, respectively).

There was no reliable main effect of preview type on error rate to word targets, both *Fs* < 1.

Nonword data. The ANOVAs on the latency data showed a significant effect of preview type, $F1(3,84) = 17.0$, $p < .001$; $F2(3,708) = 16.84$, $p < .001$: RTs were shorter for the whole preview condition (740 ms) compared to the other three preview conditions, all *ps* < .001, whereas there were no trends of a difference across the three preview conditions (772, 770, and 771 ms; for the midsegments+junctions, midsegments+terminals and terminals+junctions preview conditions, respectively).

The ANOVAs on the error data showed no significant effects, both *F*s < 1.

In the present experiment, word (and nonword) identification times were substantially shorter for the (control) whole preview condition than in the three preview conditions. The large advantage of the identity condition over the other conditions was likely due to the fact that the partial previews were so fragmented that a simple pooling or summation of the visual information failed to activate higher, stroke-unit-level or letter-level representations.

Moreover, differences in response times between the whole preview condition and the delay conditions for words were somewhat larger than the differences between the whole preview condition and the delay conditions for nonwords (39 ms ± 2 vs. 31 ms ± 1 ms, respectively), but in the word conditions, lexical decision times were (unsurprisingly) lower. This probably indicates a larger impact of the failure to activate letter-level or stroke-unit-level representations on the integration of information or a relational filtering that typically attends lexicality (producing superiority effects).

Because we presumed the partial previews were so fragmented that the visual information failed to activate letter-level or stroke-unit-level representations, we used a complementary approach with larger segments in Experiment 2, in which we deleted parts of the word's constituent letters (midsegments, junctions, or terminals), thus making these stimuli letters more letter-like. The effect is that the co-linearity of sub-sets of the sampled information is more evident, and the perceptual integrity of individual stroke units becomes more pronounced. If a segmentation must occur for a differential effect to be observed in the different conditions, and for the effect to be significant, the possibility must be entertained that there is a "continuation of form constraint". We again employed a whole preview condition as a control (see Figure 3, for illustration).

Experiment 2

Method

Participants. Thirty-two additional psychology students from the same population as those in Experiment 1 took part in the experiment.

Materials. The 240 words and 240 nonwords were the same as in Experiment 1. Again the target itself, without any deletions was used as a preview. Three new partial preview versions were created from each target (see Figure 3): i) a partial preview with midsegments deleted (Midsegments-deleted), ii) a partial preview with junctions deleted (Junctions-deleted) and iii) a partial preview with terminals deleted (Terminals-deleted). Terminals, junctions and mid-segments deletions were selected in a principle way, and contained approximately the same number of pixels per deletion. As in Experiment 1, the previews were created using FontLab Studio software. This strategy resulted in previews that retained about 75 percent of the stimulus information, as gauged by pixel counts, compared to the no-delay condition. An illustration of the three degraded versions for each letter is provided in Figure 5—note that none of the three conditions altered letter confusability per se (e.g., each letter could be uniquely identified in all three conditions).

Procedure. It was the same as in Experiment 1.

Results and Discussion

RTs less than 250 ms or greater than 1500 ms were excluded from latency analyses (2.0 and 7.2% of the data for word and nonword trials, respectively). (footnote 2) Mean correct RTs and percentage of error for each experimental condition are presented in Table 2. The statistical analyses were parallel to those in Experiment 1.

Word data. There was a significant effect of preview type on RTs to word targets, $F1(3,84) = 19.95$, $p < .01$; $F2(3,708) = 19.19$, $p < .01$. This reflected that word identification times for the whole preview condition (701 ms) and the terminals-deleted condition (706 ms) were shorter than the other two conditions (junctions-deleted condition: 725 ms; midsegments-deleted condition, 739 ms), all *ps* < .01. In addition, the 14 ms advantage of the junctions-deleted condition over the midsegments-deleted condition was significant, both *ps* < .01.

There was no significant effect of preview type on error rate to word targets, both *Fs* < 1.

Nonword data. The ANOVAs on the latency data showed a significant effect of preview, $F1(3,84) = 3.73$, $p < .05$; $F2(3,708) = 4.78$, $p < .01$. This reflected that word identification times were shorter for the whole preview condition (832 ms) than in the other three preview conditions, all *p*s < .05, whereas there were no trends of a difference across the three preview conditions (847, 851 and 847 ms for the junctions-deleted, midsegments-deleted and terminal-deleted conditions, respectively).

The ANOVAs on the error data failed to reveal any significant effects, both *Fs* < 1.

This experiment showed that word identification times were shorter in the whole preview condition and the terminals-deleted condition than in the other two conditions (midsegments-deleted condition and junctions-deleted condition). Furthermore, word identification times in the junctions-deleted condition were shorter than for the midsegments-deleted condition. For the nonwords, we only found an advantage of the whole preview condition over the three partial preview conditions.

Unlike in our first experiment, a graded dissociation between performance on our three word conditions emerged. This confirms our suspicion that the lack of a positive delay-conditions effect in Experiment 1 was likely due to the fact that the partial previews were so fragmented that the visual information failed to activate letter-level or stroke-unit-level representations. Terminals and mid-segments are structural components of stroke-units, and junctions define the relationships of stroke-units. It seems reasonable to suppose that: i) our midsegment cuts removed centrally distinctive information at the stroke-unit level, affecting the resolution of the identity of the stroke-units; ii) our cuts at terminals preserved critical information about ascender, descender, relation-to-x-height status, and preserved centrally distinctive information about the identity of the stroke unit; and iii) our junction cuts obscured relational information and preserved centrally distinctive information about the identity of the stroke unit to a greater degree than our midsegment cuts, but to a lesser degree than our junction cuts.

So, from the dissociation between performance on our three word conditions in this Experiment, it appears that the co-operative and competitive pooling required for establishing the presence and identity of discrete structural components of the letters — specifically, stroke units —, and sorting out their relationships in embedded-letter and whole-word contexts was unable to operate effectively in Experiment 1.

We hypothesize that the visual system must be able to break down and process the visual information available in nonsense strings of letters or meaningful words in such a way that it can resolve the visual information into a set of mediating units — in this case, stroke-units — of an intermediate complexity between single unitary letters and primitive letter features and that the visual cortex has learned low-level representations for such items. We will call this mandatory breaking-down and resolution process (associated with the necessary mapping onto unitary mediating structures coded in memory this implies) *quantization*.

In Experiment 1 we suggested the disturbance to the perceptual integrity of the stroke units caused by the too fragmentary nature of the stimuli might be an issue. This is a continuation of form issue associated with the need to establish collinearity in the co-operative and competitive pooling and averaging presumed to occur in the primary visual cortex. In the present experiment, the biggest performance gap was between conditions in which the perceptual integrity of the form was disturbed by just such deletions. And in the deletion conditions that performed the worst the difference in performance is consonant with the number and (less critical) location of the deletions.

From the pattern of our results in this and the previous experiment then, there appears, to be a "continuation of form" constraint (see again our discussion in the "Results and Discussion" section of Experiment 1) in early vision at the level of stroke-units in visual word recognition, indicative of an early mandatory quantization *bottleneck*. We use the term bottleneck the way it is currently being used in psychophysics to indicate the mandatory nature of the process: it is a process that needs to be passed through for recognition to occur in a smooth and unimpeded way. We can then say that terminal, mid-segment and junction deletions affect this quantization differently in proportion to how disruptive they are of satisfying the continuation of form constraint, with the deletion of terminals being the least disruptive, at least in the first steps of the stimulus sampling or extraction process (i.e., the initial 50 ms). It appears that capacity benefits because of lexicality expressed in RT distributions can't accumulate unless a continuation of form constraint at the level of stroke–units is met.

Importantly, the differences across the terminals-deleted, midsegments-deleted and junctions-deleted conditions — highlighted above — did not appear for nonword stimuli, which suggests either that top-down processes from lexical levels may be at play here (e.g., see Perea, Jiménez, & Gomez, 2014; Perea, Marcet, & Vergara-Martínez, 2016; Vergara-Martínez, Gomez, Jiménez, & Perea, 2015), or that already well before 60 ms after stimulus onset (cf. van Leeuwen, 2015) the activity of early neurons in visual cortex may become dependent on that of their neighbours through within-layer horizontal connections for familiar words. Within-layer horizontal connections across receptive field boundaries might support lateral across-the-word accumulation of information processes that transgress graphemic boundaries and accumulate pre-orthographically. Previous findings suggest that the accumulation of information that supports "orthographic processing" could be most flexible in the context of words than in the context of non-words (e.g., Grainger, Bouttevin, Truc, Bastien, & Ziegler, 2003).

General Discussion

The goal of the present experiments was to examine the role of three potentially relevant components of the word's constituent letters (terminals, midsegments, and junctions) during visual-word recognition. To that end, we conducted two lexical decision experiments with a segment-delayed procedure in which partial information of terminals, midsegments or junctions was presented for 50 ms (see Figures 2 and 3).

In Experiment 1, the partial preview was composed of terminals+junctions, midsegments+junctions, or midsegments+terminals. Results revealed that the whole preview condition for the word stimuli had a substantial advantage over the preview conditions of approximately the same size (the differences were 41, 38 and 37 ms with the midsegments+junction conditions, midsegments+terminals and terminals+junction, respectively). Given that the whole preview was presented for 50 ms, this almost complete lack of performance advantage might imply that none of the partial previews was able to activate any kind of higher-level representations to a significant degree. In light of our initial discussion of Experiment 2, it appears that it is activation of higher level stroke-unit-level representations that failed.

In Experiment 2, the partial preview was composed of the whole word except for the deletion of midsegments, junctions, or terminals. Results for the word stimuli revealed that the difference between the whole preview condition and the terminals-deleted condition was very small (4 ms), thus suggesting that this type of partial preview itself is capable of activating higher-level representations to a degree similar to the identical preview. Furthermore, the difference between the whole preview condition and the midsegments-deleted condition was rather large (a 38-ms advantage of the whole preview condition), which suggests that this second type of partial preview is not able to activate higher-level representations as successfully. As indicated in our initial discussion of Experiment 2, and consistent with the idea presented above, that in Experiment 1 it is activation of stroke-unit-level representations that largely failed, in Experiment 2 it emerges that it is an activation of stroke-unit-level representations that is differentiated according to delay type. The results thus appear to stress the disproportionately large importance of the processing of midsegments: these provide distinctive information about stroke-units during letter and word recognition. Finally, the junctions-deleted condition led to faster word identification times than the midsegments-deleted condition, which suggests that the stroke-unit-identity-defining information at midsegments is more important — at least for visual word recognition — than the local combination information at junctions.

Globally considered, the present findings then appear to suggest that midsegments are more relevant than junctions in the early processing of words, as their deletion significantly decreased the preview facilitation effect on the lexical decision task, and our results help us make some headway in discerning why. Furthermore, terminals seem to be the least critical element during the early processing of words, as their deletion involves word identification times similar to those in the whole preview condition. Critical information for the higher-level quantization bottleneck we proposed is still present in the terminals-deleted condition (i.e., midsegments and junctions). This is consistent with eyetracking evidence showing a null difference during sentence reading between a serif font (Lucida) and a sans serif font (Lucida Sans) (see Perea, 2013).

Our results are, prima facia, in line with those obtained by Petit and Grainger (2002) in the context of isolated letter recognition for upper case letters, as we found evidence that the most relevant features for visual processing of words are contained midsegments. The current experiment has in common with that of Petit and Grainger that the two experiments used paradigms in which the target stimuli remained unaltered. In Petit and Grainger's experiments (2002) a masked priming task was used, in which the primes were created by deleting parts of the target letter.

However our results extend our understanding by revealing the existence of a "continuation of form" constrain, suggesting an explanation for the relevance of midsegment features. The suggestion that there is a mandatory quantization bottleneck at the stroke-unit level offers a new perspective on the encoding dynamics at the front end of visual word recognition. We suggested that there exists a continuation of form constraint indicative of a quantization bottleneck, and that this bottleneck exists around the components of letters we have tagged as stroke-units. Prima facia what this means is that the pooling or compulsory averaging or summation of feature-level information that leads to recognition at a letter- or word-level is indirect. We think that the pooling or averaging is not a simple summative process of detecting localizable low-level componential information such as junctions, terminations and mid-segments, constructed as independently represented entities, and then of combining them by virtue of their overlap in the larger receptive fields of a subsequent level to form letters. Such a process is proposed in, for example, Fukushima's (2013) neocognitron model and its antecedents. In Fukushima's account upper case letters illustrate the scheme. Instead, it might make sense to work from the notion that the visual system breaks stimulus words down into oriented lines and curves (simple — S1 — cells in level 1), to the point where responsiveness to aspect, closure, extendedness and expressedness can accumulate (in C1 cells with a complex sensitivity) and resolution or quantization into stroke-units occurs (S2). So we suggest, the quantization into stroke-unit-like operators constitutes a first-order encoding bottleneck. For computational modelling, this might mean abandoning the Rumelhart and Siple (1974) font and adapting a proposal closer to the representation of basic units formalized using the concept of "roles" in the computational work of Hofstadter and his collaborators around the Letter/Spirit project (e.g., Hofstadter, 1982; Hofstadter & McGraw, 1993; McGraw, Rehling & Goldstone, 1994; McGraw, 1995). Hofstadter and McGraw treat stroke-unit-defined parts of letters like bowls or cross-bars or extended/neutral stems as roles. Roles are fuzzy, or abstract or protoypical "concepts" that are "filled" in different fonts in different ways. In letters various roles are in role-relationships. Progress in font-development toward increasing legibility and readability, or toward creative variations in form for display come through strategic "norm-violations" in the way roles are filled. Experimental work on letter perception by Hofstadter and his collaborators indicate that letter-perception is based on a perceptual identification of roles and role-relationships.

In the present experiments, a partial preview was presented for 50 ms and was immediately followed by the target. This constitutes a closer situation to normal reading, compared with the paradigms employed in previous experiments where target stimuli were manipulated, either by removing some features until word response deteriorates (Lanthier et al., 2009; Szwed et al., 2011), or placing successively several masks over the stimuli (Fiset et al., 2008, 2009). Lexical decision times correlate moderately highly with eyetracking measures such as gaze duration (see Schilling, Rayner, & Chumbley, 1998, for early evidence; but see Kuperman, Drieghe, Keuleers, & Brysbaert, 2013, for a cautionary note). This correlation alone — aside from the specifics of the delay technique we employed — might make the lexical decision task ecologically valid for normal sentence reading. Delay techniques constrained to the first 50 ms in lexical decision task appear then to specifically target the stimulus sampling and perceptual encoding components occurring at the very front end of visual processing presumably within the visual cortex and along the ventral stream. Nonetheless, we acknowledge that further research should examine whether the present findings can be generalized to normal sentence reading (e.g., when processing parafoveal information using Rayner's, 1975, gaze-contingent boundary-change paradigm).

We started out by asking about the role of letter features in visual word recognition. We found that a continuation of form constraint had to be met before the effect of various types of deletions asserted themselves. We also found that in words – but not in nonwords– it mattered whether the deletions were within stroke-units or at their ends. Local combination detection seemed to matter somewhat less than the ability to quantize the information into discrete stroke units. To accommodate these results and explain the accumulation of superiority effects already at the very front end of visual word recognition it might be necessary to adjust current orthographic models of visual word recognition, or devise new ones. At the very least our results might lead to a better, more diversified understanding of the structural components of what has — in computational models — been labelled the feature level, and point the way to a more realistic set of computational units than the Rumelhart and Siple (1974) font provides.

References

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*, 115–147. doi: 10.1037/0033-295X.94.2.115

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). Amsterdam: Academic Press.

Blais, C., Fiset, D., Jolicoeur, P., Arguin, M., Bub, D. N., & Gosselin, F. (2009). Reading between eye saccades. *PLOS ONE*, *4*(7), e6448. doi:10.1371/journal.pone.0006448

Bouma, H. (1971). Visual recognition of isolated lower-case letters. *Vision Research, 11*, 459-474. doi: 10.1016/0042-6989(71)90087-3

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences, 18*, 90-98. doi: 10.1016/j.tics.2013.11.005

Carreiras, M., Gillon-Dowens, M., Vergara, M., & Perea, M. (2009). Are vowels and consonants processed differently? ERP evidence with a delayed letter paradigm. *Journal of Cognitive Neuroscience, 21*, 275-288. doi: 10.1162/jocn.2008.21023

Chang, Y. N., Furber, S. & Welbourne, S. (2012). Modelling normal and impaired letter recognition: Implications for understanding pure alexic reading. *Neuropsychologia, 50*, 2773-2788. doi: 10.1016/j.neuropsychologia.2012.07.031

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

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review, 177*, 713-758. doi: 10.1037/a0019738

Davis, C. J., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods, 37,* 665–671. doi: 10.3758/BF03192738

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

Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. *Quarterly Journal of Experimental Psychology, 66*, 563-580. doi: 10.1080/17470218.2012.658820

Feldman, J. (2015). Probabilistic models of perceptual features. In J. Wagemans (Ed.), *Oxford Handbook of Perceptual Organization* (pp. 1008-1026). Oxford, U.K.: Oxford University Press.

Finkbeiner, M., & Coltheart, M. (2009). Letter recognition: From perception to representation. *Cognitive Neuropsychology, 26*, 1-6. doi: 10.1080/02643290902905294

Fiset, D., Blais, C., Arguin, M., Tadros, K., Éthier-Majcher, C., Bub, D., & Gosselin, F. (2009). The spatio-temporal dynamics of visual letter recognition. *Cognitive Neuropsychology, 26,* 23–35. doi:10.1080/02643290802421160

Fiset, D., Blais, C., Éthier-Majcher, C., Arguin, M., Bub, D., & Gosselin, F. (2008). Features for identification of uppercase and lowercase letters. *Psychological Science, 19,* 1161–1168. doi: 10.1111/j.1467-9280.2008.02218.x

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35,* 116–124. doi: 10.3758/BF03195503

Franklin, S. (1995). *Artificial Minds: An Exploration of the Mechanisms of Mind*. Cambridge, MA: MIT Press.

Fukushima, K. (2013) Artificial vision by multi-layered neural networks: Neocognitron and its advances. *Neural Networks, 37*, 103–119. doi: 10.1016/j.neunet.2012.09.016.

Gibson, E. J., Gibson, J.J., Pick, A.D. & Osser, H. (1962). A developmental study of the discrimination of letter-like forms. *Journal of Comparative Physiology, 55*, 897-906. doi: 10.1037/h0043190

Gibson, E.J. (1965). Learning to read. *Science, 148*, 1066-1072. doi: 10.1126/science.148.3673.1066

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

Grainger, J., Bouttevin, S., Truc, C., Bastien, M., & Ziegler, J. (2003). Word superiority, pseudoword superiority, and learning to read: A comparison of dyslexic and normal readers. *Brain and Language*, *87*, 432-440. doi: 10.1016/s0093-934x(03)00145-7

Hinton, G. E. (1981). Shape representation in parallel systems. In *Proceedings of the Seventh International Joint Conference on Artificial Intelligence Vol 2*, Vancouver BC, Canada.

Hofstadter, D. R., (1982), Metafont, Metamathematics, and Metaphysics. Issue 136 of *Technical report* (Indiana University, Bloomington. Computer Science Department, available online); reprinted as Metafont, metamathematics and metaphysics: comments on Donald Knuth's article 'The concept of a meta-font.' *Visible language*, *16*, 3-27.

Hofstadter, D. R., & McGraw, G. (1993) Letter Spirit: An Emergent Model of the Perception and Creation of Alphabetic Style. *Technical Report 68, Center for Research on Concepts and Cognition*, reprinted as Douglas Hofstadter & Gary McGraw. (1995) Letter Spirit: Esthetic Perception and Creative Play in the Rich Microcosm of the Alphabet. *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*. New York: Basic Books.

Houpt, J. W., Townsend, J. T., & Donkin, C. (2014). A new perspective on visual word processing efficiency. *Acta Psychologica, 145*, 118-127. doi: 10.1016/j.actpsy.2013.10.013

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurons in the cats striate cortex. *Journal of Physiology*, *148*, 574–591. doi: 10.1113/jphysiol.1959.sp006308

Lanthier, S. N., Risko, E. F., Stolz, J. A., & Besner, D. (2009). Not all visual features are created equal: Early processing in letter and word recognition. *Psychonomic Bulletin & Review, 16,* 67-73. doi: 10.3758/PBR.16.1.67

Lee, H. W., Rayner, K., & Pollatsek, A. (2001). The relative contribution of consonants and vowels to word identification during reading. *Journal of Memory and Language, 44*, 189-205. doi: 10.1006/jmla.2000.2725

McClelland, J. L. (1996). Integration of Information: Reflections of the Theme of Attention and Performance XVI. In T. Inui and J. L. McClelland (Eds.), *Attention and Performance XVI: Information Integration in Perception and Communication* (pp. 633-656). Cambridge, MA: MIT Press.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88*, 375-407. doi: 10.1037%2F0033-295X.88.5.375

McGraw, G., Rehling, J., & Goldstone, R. (1994) Letter Perception: Toward a conceptual approach. In *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, 613-618, Atlanta, GA.

McGraw, G. (1995). Emergent Perception of Gridletters. In *Proceedings of the 1995 Midwest AI and Cognitive Science Conference*, 63-67.

Mewhort, D. J. K., & Johns, E. E. (1988). Some tests of the interactive-activation model for word identification. *Psychological Research*, *50*, 135-147. doi: 10.1007/BF00310174

Noordzij, G. (2005). *The stroke: Theory of writing*. London, UK: Hyphen Press.

Parida, L., & Mudur, S. P. (1994). Computational methods for evaluating swept object boundaries. *The Visual Computer, 10*, 266-276. doi: 10.1007/bf01901583

Perea, M. (2013). Why does the APA recommend the use of serif fonts? *Psicothema, 25*, 13-17. doi: 10.7334/psicothema2012.141

Perea, M., Comesaña, M., & Soares, A.P. (2012). Does the advantage of the upper part of words occur at the lexical level? *Memory and Cognition, 8*, 1257-1265. doi: 10.3758/s13421-012-0219-z

Perea, M., Jiménez, M., & Gómez, P. (2014). A challenging dissociation in masked identity priming with the lexical decision task. *Acta Psychologica, 148*, 130-135. doi: 10.1016/j.actpsy.2014.01.014

Perea, M., Marcet, A., & Vergara-Martínez, M. (2016). Phonological-lexical feedback during early abstract encoding: The case of deaf readers. *PLoS ONE, 11*(1), e0146265. doi:10.1371/journal.pone.0146265

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. doi: 10.1037/0033-295X.114.2.273

Petit, J. P., & Grainger, J. (2002). Masked partial priming of letter perception. *Visual Cognition, 9,* 337–354. doi: 10.1080/13506280042000207

Põder, E. (2014). Crowding by a single bar: Probing pattern recognition mechanisms in the visual periphery. *Journal of Vision, 14*, 1-12. doi: 10.1167/14.13.5

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis*. Journal of Experimental Psychology. Learning, Memory, and Cognition, 21,* 785–794. doi: 10.1037/0278-7393.21.3.785

Price, C. J., & Friston, K. J. (2005) Functional ontologies for cognition: The systematic definition of structure and function. *Cognitive Neuropsychology* 3, 262-275. doi: 10.1080/02643290442000095.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*, 65-81. doi: 10.1016/0010-0285(75)90005-5

Rumelhart, D. E., & Siple, P. (1974). Process of recognizing tachistoscopically presented words. *Psychological Review, 87*, 99-118. doi: 10.1037/h0036117

Schilling, H. E., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition, 26*, 1270-1281. doi: 10.3758/BF03201199

Smith, F. (1969). Familiarity of configuration vs. discriminability of features in the visual identification of words. *Psychonomic Science,* 14, 261-262. doi: 10.3758/bf03329112

Szwed, M., Cohen, L., Qiao, E., & Dehaene, S. (2009). The role of invariant line junctions in object and visual word recognition. *Vision Research*, *49*, 718-725. doi: 10.1016/j.visres.2009.01.003

Szwed, M., Dehaene, S., Eger E., Kleinschmidt A., Valabregue R., Amadon A., & Cohen L. (2011). Specialization for written words over objects in the visual cortex. *Neuroimage, 56,* 330-44. doi: 10.1016/j.neuroimage.2011.01.073

van Leeuwen, C. (2015). Hierarchical stages or emergence in perceptual integration? In J. Wagemans (Ed.), *Oxford Handbook of Perceptual Organization* (pp. 969-988). Oxford, U.K.: Oxford University Press.

Vergara-Martínez, M., Gomez, P., Jiménez, M., & Perea, M. (2015). Lexical enhancement during prime-target integration: ERP evidence from matched-case identity priming. *Cognitive, Affective, and Behavioral Neuroscience, 15*, 492-504. DOI: 10.3758/s13415-014-0330-7

Watt, W. C. (1988). Canons of Alphabetic Change. In D. de Kerckhove and C. J. Lumsden (Eds.), *The Alphabet and the Brain: The Lateralization of Writing*. Berlin: Springer-Verlag

Footnotes

Footnote 1. As discussed by McClelland (1996), in an intrinsic integration model, a letter is a pattern over feature units, and a word is a pattern over sets of feature units. Letters, words, and phrases appear as "simply the coherent and relatively independent sub matters of larger, more coherent whole patterns" (p. 647).

Footnote 2. The percentage of response times removed from the pseudowords stimuli was higher than in Experiment 1. Had we used a longer cutoff for pseudoword trials (e.g., 2000 ms which removes 1.1% of the data), the same pattern of data was exactly the same as that reported here (whole word = 882; junctions-deleted = 904 ms; midsegments-deleted = 901 ms, terminals-deleted = 900 ms).

Table 1. Mean lexical decision times (in ms) and accuracy (in parentheses) for words and nonwords in Experiment 1

	Identity	Midsegments-Junctions	Midsegments-Terminals	Terminals-Junctions
Words	627 (.94)	668 (.94)	665 (.94)	664 (.94)
Nonwords	740 (.94)	772 (.94)	770 (.95)	771 (.94)

Table 2. Mean lexical decision times (in ms) and accuracy (in parentheses) for words and nonwords in Experiment 2

	Identity	Junctions-deleted	Midsegments-deleted	Terminals-deleted
Words	701 (.95)	725 (.95)	739 (.94)	706 (.95)
Nonwords	832 (.93)	847 (.93)	851 (.93)	848 (.93)

Figure captions

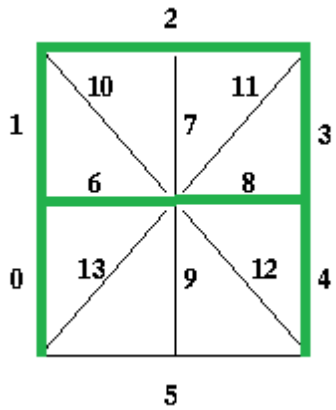
Figure 1. Representation of the letter "A" in the 14-feature uppercase-letter system created by Rumelhart and Siple (partially adapted from Figure 2 in Rumelhart & Siple, 1974).

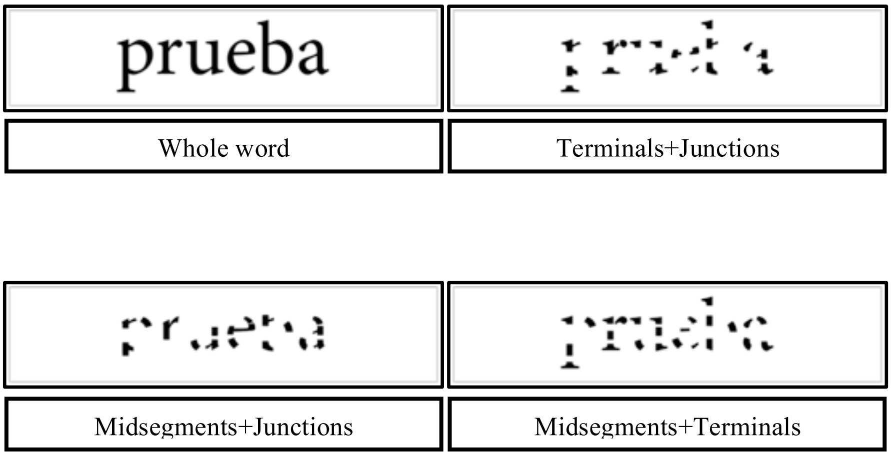
Figure 2. Types of previews used in Experiment 1(Identity, Terminals+Junctions, Midsegments+Junctions, and Midsegments+Terminals)

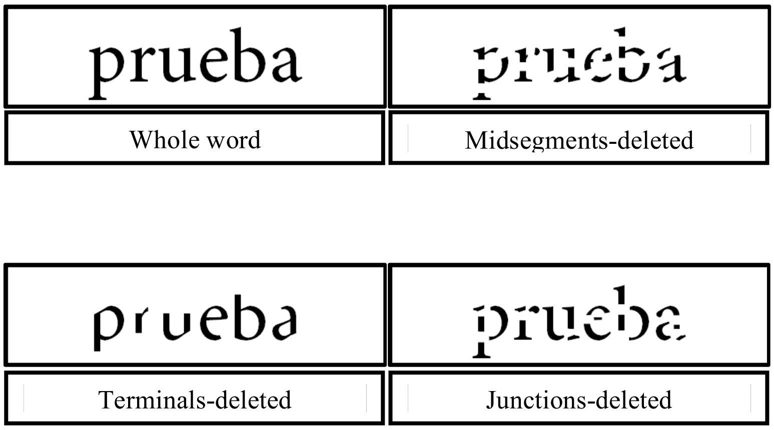
Figure 3. Types of previews used in Experiment 2 (Identity, Midsegments-deleted, Terminals-deleted, and Junctions-deleted)

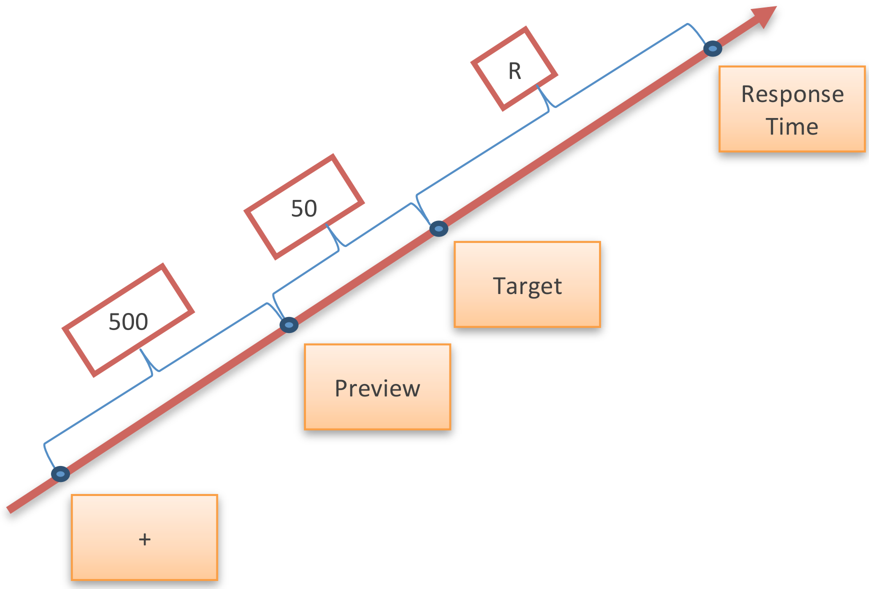
Figure 4. Scheme of a given trial with the Delayed Segment Technique

Figure 5. Alphabet in the various degraded versions of Experiment 2 (from top to bottom: terminals-deleted; junctions-deleted; midsegments-deleted; intact) (Note that the letters c, g, i, j, o, w, and s were not included in the set of stimuli.)


2
10
11
1
7
3
6
8
0
13
9
12
4
5


prueba
Whole word
Terminals+Junctions
Midsegments+Junctions
Midsegments+Terminals


prueba
Whole word
Midsegments-deleted
Terminals-deleted
Junctions-deleted


R
50
500
Response Time
Target
Preview
+

a b c d e f g h i j k l m n o p q r s t u v w x y z

a b c d e f g h i j k l m n o p q r s t u v w x y z

a b c d e f g h i j k l m n o p q r s t u v w x y z

a b c d e f g h i j k l m n o p q r s t u v w x y z

Appendix

Words and Nonwords in Experiments 1 and 2

Words: venta; murmurar; fabada; bandera; rueda; beata; prueba; avena; ayunar; rezar; querer; nadar; perfume; verde; arder; tarde; mente; matar; abertura; futura; rebanada; duradera; heredar; verdura; defender; prenda; patera; fumar; ventana; humana; pared; taburete; perpetua; dureza; mudanza; adaptar; trazar; buque; patentar; trufa; pradera; demandar; tardar; funda; abrazar; matanza; trenza; nueve; barrera; remate; armada; renta; puente; quemar; meter; brevedad; detener; depurar; huerta; panza; rampa; barata; empezar; ranura; veranear; humedad; puerta; eterna; ruptura; patada; menta; parra; madurez; bufanda; etapa; empate; enumerar; quebrar; nevar; enferma; tetera; radar; manera; demente; muerta; punzante; parar; pubertad; tenaz; pedrea; turbante; mudar; entrar; manta; pureza; madre; aparente; panda; paquete; atenta; amante; trama; embarque; fuera; fruta; apunte; breve; atrapar; redada; babear; temer; azafata; parada; zurda; duque; apartar; ataque; fraude; pedante; entrenar; duende; retratar; tardanza; hembra; ternera; nuera; verbena; fuerza; aduana; atraer; tarea; barba; armadura; madura; barra; tapar; hambre; entrada; barrer; heredera; frente; deber; bazar; reventar; tapadera; arena; nevera; tanque; mandar; tener; bandada; rareza; atender; antena; ayudar; pauta; neutra; armar; taberna; retrete; andar; apretar; fuerte; perder; examen; manzana; peruana; merendar; madera; aventura; tarta; banana; drama; faena; debate; maqueta; empanada; terraza; banquete; patente; arenque; deuda; durar; pereza; ayuda; trauma; banda; rumba; traza; tatuar; amenazar; marea; demanda; fuente; avanzar; derrame; errata; prudente; papaya; extraer; beber; textura; bruta; frenar; trampa; raqueta; tenaza; ternura; trazada; empatar; rebuznar; puntuar; retener; aprender; verdad; panadera; pantera; fauna; traer; mutante; hermana; reparar; parque; amenaza; vender; rumana; apuntar; azufre; padre; trepar; punta; ayudante; danza; patata; nevada; audaz; rematar; experta; dudar; remar

Nonwords: embretar; duvea; patenza; venvera; medar; bamurene; deferer; naubra; denurtar; anarbura; berrar; bamatera; ravar; quenar; rubenza; numedad; paveda; exdate; aberna; endrar; varena; tadrene; debarta; henar; punfurar; fruda; abenpar; veper; baunde; muenta; amerava; druda; vaxfera; ranque; apradar; nampre; aunaz; panar; evandura; merrube; munana; bamadear; damar; derrea; derday; nevunar; embarta; embruer; trumpe; amekana; temuna; parvana; defratar; averta; mupartad; tamea; atanpar; efrena; etanea; tarmera; merta; adear; tavuana; berruma; munte; naeta; apentar; madantar; munduar; merar; frunda; bevur; hunar; mevente; druepa; depuda; abrear; medeva; adeta; turda; purfante; dreba; mermer; envar; retetar; punar; davana; aepta; favena; mudatera; fexpana; adena; teder; fraede; ervatar; avarte; rumpa; perpe; rupeva; tuarena; fenvena; trupante; amedante; redentar; padar; amubenar; mueda; veada; vapeda; parut; rexde; mamatera; trerva; vunar; frate; fefuda; bumpa; embrada; advatura; teana; fuenva; reque; depantar; nuvuna; venfada; atremar; tavuna; mafrate; pruba; meneda; taprena; marante; tumandar; redante; ermateda; nevena;

fantuna; arpeda; ateravar; derate; prenar; redate; muedrar; mubra; peranda; ruvena; avunte; tenevar; braur; venzuete; tadeba; adapua; runta; prenta; bauda; etana; tunvanza; narear; envarza; reder; randar; truafa; deptar; patre; adatre; pafuata; nedeta; muena; huede; batuaz; marda; duemar; mapre; benar; arandue; enartar; tunvente; fendar; bedar; avandar; muebar; dafrata; darva; ferne; vanuta; numbra; menque; vuarte; brana; anvera; vuzarda; tundera; banza; pretedaz; tenta; bavar; benanera; pardar; funeda; rezpura; narda; zunta; tenfura; bavuna; hudae; faunte; afranbar; parvume; denta; tureba; tuarar; aprunar; papured; dupente; nuerta; pemarada; nabra; pavana; tantura; arter; edeva; dezantar; namea; drapar; nurta; derutar; neverar; exhadar; aerbunte; perpenet; renaz; fenta; prunte; paunte; murte; paprum; herta; devarar; emparpae; prahuda