

# Is there a genuine advantage to the upper part of words during lexical access? Evidence from the Stroop task

Pilar Tejero · Manuel Perea · María Jiménez

Published online: 10 January 2014  
© Psychonomic Society, Inc. 2014

**Abstract** A number of recent visual-word recognition and reading experiments have concluded that the upper part of words is more important for lexical access than is the lower part, which conforms with Huey's (1908) observation. Here, we examined whether this phenomenon may simply be due to the fact that words in Indo-European languages tend to have a higher number of confusable letters in the lower than in the upper part. We manipulated the letter ambiguity of the upper and lower parts of words in two experiments in which we asked participants to report the presentation color of the upper and lower parts of color words and noncolor words, and in a baseline condition, of strings of &s (Stroop task). In Experiment 1, the lower part of noncolor words was more ambiguous than the upper part (upward-unbalanced words), whereas in Experiment 2, the ambiguities of the two parts of the noncolor words were similar (balanced words). For the upward-unbalanced noncolor words, the magnitude of lexical interference (relative to the baseline condition) was greater for the upper than for the lower part. Critically, the differences vanished when this factor was controlled (i.e., balanced words; Exp. 2). Thus, the apparent bias in favor of the upper

part of words can be parsimoniously described as an idiosyncratic feature of the words' component letters.

**Keywords** Visual-word recognition · Lexical interference · Computational models of visual-word recognition · Stroop

In his influential book on reading, Huey (1908; reissued in 1968) claimed that “the upper half of a word or letter is obviously more important for perception than is the lower half” (p. 65) and “the attention concerns itself most with the upper half of the word, and the letters projecting below are not so important” (p. 60). To provide support for this assertion, Huey presented three short passages: (1) with intact words, (2) with words that only kept their upper half, and (3) with words that only kept their lower half. Although no empirical data were presented, the texts were manifestly easier to read when the words only kept their upper half than when they only kept their lower half (see Perea, 2012, for evidence gathered while recording the participants' eye movements; see also Chou, 1930, and Tsao & Wang, 1983, for evidence in Chinese). Courses on speed reading claim that “the top of letters alone are usually all you need to recognize the word at hand” (Wechsler & Bell, 2005, p. 65). The advantage of the upper part of words during visual-word recognition has recently been reported in various languages and paradigms, using degraded stimuli and the Bubbles technique (Blais et al., 2009), presenting primes that only kept their lower or upper halves in a masked-priming technique (Perea et al. 2012b), and presenting stimuli that only kept their lower or upper halves in a delayed-segment technique (Perea et al. 2012a).

What is the reason for the apparent bias toward the upper part of words? After all, none of the computational models of visual-word recognition and reading make any claims concerning a differential role of the upper/lower part of a word's constituent letters (see, e.g., Davis, 2010, for a review).

---

P. Tejero  
Departamento de Psicología Básica and ERI-Lectura, Universitat de València, Valencia, Spain

M. Perea · M. Jiménez  
Departamento de Metodología and ERI-Lectura, Universitat de València, Valencia, Spain

M. Perea  
Basque Center on Brain, Language, and Cognition, San Sebastián, Spain

P. Tejero (✉)  
Facultad de Psicología, Departamento de Psicología Básica,  
Av. Blasco Ibáñez, 21, 46010 Valencia, Spain  
e-mail: pilar.tejero@uv.es

Thus, a genuine advantage in the processing of the upper part of words would require an amendment of our theoretical models. One explanation is that this bias parallels the upward biases that occur in the visual perception of features of objects and faces (see Caldara et al., 2006, and Thomas & Elias, 2011, for recent evidence regarding these biases). As Huey (1908) indicated, “we habitually find most meanings in the upper part of objects; we ourselves are so placed and so oriented as to bring this about” (p. 65). Although the high-level mechanisms involved in visual-word processing seem to be highly selective (see Dehaene, Cohen, Sigman, & Vinckier, 2005), it has been argued that word processing can be influenced by mechanisms that are engaged in the visual processing of the objects (see Behrmann & Plaut, 2013, and Ishai, Ungerleider, & Haxby, 2000). Thus, this bias for the upper part of objects could, in principle, influence the automatic detection of letter features during visual-word recognition and reading.

Before asserting that the bias for the upper part of words is related to the biases that occur in object/face processing, we will need to discard a simpler explanation, that this phenomenon is simply due to differences between the upper and lower parts in the confusability of the words’ constituent letters of the Roman alphabet. Many letters share the lower part (e.g., f, i, l, and r; a similar case applies to the lower parts of the letters n and h and the combination of these two letters with i, f, l, or r, as well as to the letter m; see Fig. 1). However, a number of letters share the upper part (e.g., the upper part of the letter v can be mistaken for the letter y; the same goes for b and h, g and q, i and j, or n and p; see Fig. 1). The net result may suggest that the upper parts of letters should not have an advantage in a letter recognition task on a letter-by-letter basis, as was actually reported by Fiset et al. (2009) in a study that used the Bubbles technique. The story is more complicated, however, because this technique also produced a bias for the upper parts of words (Blais et al., 2009). To explain this apparent discrepancy, Blais et al. (2009) suggested that “letter representations may be slightly different for isolated letters and for letters in words” (p. 6). We believe that a more straightforward explanation is that the letters that are potentially confusable in their lower parts occur more frequently than the letters that are potentially confusable in their upper parts. Specifically, when weighted by letter frequency, the percentage of lowercase letters that are ambiguous in their lower parts in Indo-European languages is substantially larger than the percentage of letters that are ambiguous in their upper parts (e.g., in Arial font, 68 % vs. 50 % in Spanish; 68 % vs. 51 % in English; and 68 % vs. 50 % in French; these values may vary slightly with other fonts).<sup>1</sup> As a result, words in

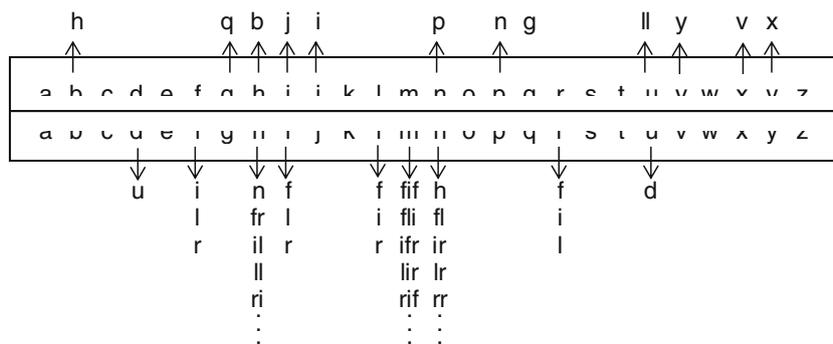
these languages have higher numbers of potentially confusable letters in their lower than in their upper parts, thus providing the upper part of words with an advantage.

A demonstration that letter confusability may play a role in the apparent bias toward the upper part of words was provided by Shimron and Navon (1980). They examined the performance of native English speakers and native Hebrew speakers when reading texts (in English and Hebrew, respectively) in which the lower/upper parts of the words had been removed. The rationale of the manipulation across languages was that, unlike the Roman alphabet, Hebrew letters have critical information distributed along their lower halves (see Shimron & Navon, 1980, for further details). In the English text, Shimron and Navon found that reading was faster (and had fewer errors) when the text only contained the upper part of words than when it only contained the lower part of words. However, the critical finding was that, in the Hebrew text, reading was *slower* (and had more errors) when the text only contained the upper part of words than when it contained the lower part of words (i.e., the *lower* part of words in Hebrew had an advantage).

It is important to stress that the experiments cited above did not control for the potential differences in letter ambiguity across their upper/lower parts. Therefore, texts in English/Spanish/French/Portuguese/Chinese could have a prevalence of words whose upper part was highly informative, whereas the opposite could have been the case in the texts in Hebrew. Obviously, lack of control over letter ambiguity may lead to a bias for the upper part of words in the Roman and Chinese scripts, and a bias for the lower part of words in the Hebrew alphabet. Thus, the question of whether there is a bias for the upper part of words during visual-word recognition remains unanswered. To carefully examine this issue, it will be critical that the letter ambiguity of the upper/lower parts of words *within* a particular language be manipulated. To that end, in Experiment 1, we selected a set of words in which the number of potentially confusable letters in their lower part was greater than the number of letters that were confusable in their upper part (upward-unbalanced words). For instance, the word *molino* (the Spanish for “mill”) has four potentially confusable letters (see Fig. 1) in its lower part (m, l, i, and n: ) but only two in its upper part (i and n: ). In Experiment 2, we selected a set of words with similar numbers of potentially confusable letters in their upper and lower parts (balanced words). For instance, the word *cipres* (the Spanish for “cypress”) has two confusable letters in its lower part (i and r: ) and two confusable letters in its upper part (i and p: ).

In these experiments, we employed a single-word stimulus presentation rather than presenting sentences (or short texts), because that would have made it virtually impossible to control for the ambiguity of the upper/lower parts of the words—as we indicated above, words in Indo-European languages tend to be upward-unbalanced. In particular, we employed a

<sup>1</sup> The Spanish data were obtained using the B-Pal database (Davis & Perea, 2005); the English and French data were taken from the frequency counts available at [http://en.wikipedia.org/wiki/Letter\\_frequency](http://en.wikipedia.org/wiki/Letter_frequency) (retrieved July 2013).



**Fig. 1** In the box: Upper and lower parts of Roman letters in Arial font, as we used in the present experiments. Outside the box: The arrows point at the potentially ambiguous letters when only the upper part of the letter is visible (top) or when only the lower part of the letter is visible (bottom).

subtle manipulation that minimized guessing strategies from the partially presented words: Participants were not required to identify (i.e., reproduce) those words, but rather to name the presentation color of the stimuli—as in the classic Stroop task (Stroop 1935; see MacLeod, 1991, 2005, for a review). The upper and lower parts of four color names, as well as meaningless strings of characters (&&&&), were intermixed among the upper and lower parts of the critical noncolor words (upward-unbalanced in Exp. 1 and balanced in Exp. 2).<sup>2</sup>

In the typical Stroop experiment (i.e., with whole stimuli), the slowest and least accurate responses occur on conflicting trials (e.g., the word “yellow” written in purple), whereas the fastest and most accurate responses occur on congruent trials (e.g., the word “yellow” written in yellow). Of particular relevance here is that performance is worse on trials with noncolor words (e.g., the word “bottle” written in yellow) than on trials with a baseline condition, such as a color patch, a figure in color, a meaningless row of characters in color (e.g., &&&& or XXXXX), or an unpronounceable nonword in color (e.g., “pvztm”) (e.g., see Klein, 1964, and MacLeod, 2005). This shows that the Stroop effect has to do with the effect of congruency between the word meaning and the required response and with lexical interference (Brown, 2011). The magnitude of lexical interference (i.e., the difference between noncolor words vs. the baseline condition) is modulated by different factors, but it typically ranges between 30 and 50 ms (Brown, 2011). The idea is that the higher the degree of lexical activation from noncolor words, the greater the interference. This leads to the key prediction in the present experiments: If there is a genuine bias in favor of the upper parts of words in visual-word recognition (i.e., if the upper parts of words generate more lexical activation than do the lower parts), performance on the trials with noncolor stimuli should be worse (i.e., slower and/or more errors) for the upper than for the lower part. Importantly, this should be the case

regardless of the type of words (i.e., with both upward-unbalanced [Exp. 1] and balanced [Exp. 2] noncolor words). Alternatively, if the favoring of the upper parts of words during visual-word recognition is merely due to the characteristics of the words’ constituent letters of the Roman script, the greater interference for the upper parts of words should only be evident for upward-unbalanced noncolor words (Exp. 1), but not for balanced noncolor words (Exp. 2). (For color words, the stimuli were upward-unbalanced in the two experiments; note that most basic color names in Spanish are upward-unbalanced.) To minimize the effects of perceptual reconstruction of the incomplete words that might occur when enough time was provided (see, e.g., Jordan, Thomas, & Scott-Brown, 1999, for an illusory letter phenomenon), the stimuli were presented very briefly (100 ms) and were immediately masked.

### Experiment 1

#### Method

**Participants** A group of 16 students (11 women, five men) from the Universitat de València voluntarily took part in the experiment. All of them were native Spanish speakers with normal or corrected-to-normal vision.

**Materials** The color names were the Spanish words *rosa* (“pink”), *marron* (“brown”), *morado* (“purple”), and *amarillo* (“yellow”). The mean word frequency per million of these color words was 23.2 in the ESPAL database (Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013). In addition, we selected 24 noncolor words from the ESPAL database—six for every color name. Each noncolor word equated with one of the color words in their initial letter, number of letters, and number of syllables. All the noncolor words were high-frequency words in Spanish (mean word-frequency: 23.8 per million). The list of noncolor words is available in the Appendix. All words—both the color words and the (critical) noncolor words—had more confusable letters in their

<sup>2</sup> Although we could have opted for a purely non-color-word Stroop task (see Burt, 2002; Warren, 1972), for comparison purposes, we thought that it was preferable to include the congruent/incongruent conditions with color words, as in the standard Stroop experiment.

lower than in their upper part; that is, they were upward-unbalanced words. For each word, we created the upper part version (e.g., *arohivn*) and the lower part version (e.g., *arohivv*). The original words were 33 pixels in height, and the width of the part removed was the same in all cases; the stimuli were 18 pixels in height (see Fig. 1). On average, 54 % of the word constituent letters were ambiguous in their lower parts, whereas this percentage was reduced to 18 % in their upper parts [ $F(23)=9.22, p<.001$ ]. Apart from words (presented on 88.9 % of the trials), rows of five ampersands (&&&&&) were also presented (11.1 % of trials) as a baseline condition (see MacLeod, 2005). The stimuli were presented equally often in pink (RGB 255, 17, 136), brown (RGB 91, 49, 17), purple (RGB 70, 18, 74), and yellow (RGB 255, 255, 5) on a black background.

**Procedure** Participants were seated in a dimly lit room. DMDX software (Forster & Forster, 2003) controlled both the presentation of the stimuli on a computer monitor and the recording of the verbal responses. Participants were instructed to name the color of the printed stimulus as rapidly as possible while trying to avoid errors. Each stimulus was presented for 100 ms, was preceded by a fixation point (500 ms), and was followed by a pattern mask (i.e., #)%[#%]#)%[#], which remained on the screen until the participant's response or until 2,000 ms had passed.

**Design** In total, 324 test trials were presented for each participant. Congruent, conflicting, and noncolor trials were included in the same proportion. There were 96 trials with congruent upper/lower color words, 96 with conflicting upper/lower color words, 96 with upper/lower noncolor words. The remaining 36 trials displayed the sequence &&&&&. Each partial color word was presented 24 times in the upper condition and 24 times in the lower condition. In both conditions, half of the trials were congruent and the other half were conflicting. Noncolor words were presented four times per participant, each one in a different ink color; two of them were presented in the upper condition and the other two in the lower condition. All trials were presented in a random order for each participant, with the restriction that no given word or color could be repeated in two consecutive trials.

## Results and discussion

In the experiment, error data and naming times were obtained using CheckVocal (Protopapas, 2007). Error responses and naming times below 200 ms or above 1,300 ms were excluded from the latency analyses (3.5 % of excluded trials). Separate analyses of variance (ANOVAs) were conducted on the mean naming times and percentages of errors, with Type of Trial (congruent color words/conflicting color words/noncolor words) and Word Part (upper part/lower part) as within-subjects factors. To assess the generality of the effects across

items for the set of noncolor words, analyses by items ( $F_2$ ) were also conducted; only analyses by subjects were conducted on the set of color words. (The &&&&& trials were employed as a baseline in subsequent analyses on lexical interference.) The mean naming latencies and percentages of errors across participants for the different conditions are shown in Table 1.

**Naming times** The main effects of type of trial,  $F(2, 30)=96.4$ ,  $MSE=1929$ ,  $p<.001$ ,  $\eta_p^2=.87$ , which reflected a substantial Stroop effect, and word part,  $F(1, 15)=35.21$ ,  $MSE=661$ ,  $p<.001$ ,  $\eta_p^2=.70$ , were significant. More importantly, the interaction between the two factors was also significant,  $F(2, 30)=7.4$ ,  $MSE=855$ ,  $p=.002$ ,  $\eta_p^2=.33$ . This interaction revealed that for conflicting color word trials, responses were also slower for the upper than for the lower parts (697 vs. 644 ms, respectively),  $F(1, 15)=15.7$ ,  $MSE=1,465.9$ ,  $p=.001$ ,  $\eta_p^2=.51$ . In contrast, no differences of word parts occurred for congruent color words (520 vs. 521 ms),  $F<1$ . For noncolor word trials, responses were slower for the upper than for the lower parts (640 vs. 600 ms, respectively),  $F(1, 15)=25.6$ ,  $MSE=500.9$ ,  $p<.001$ ,  $\eta_p^2=.63$ ;  $F(2, 23)=19.19$ ,  $MSE=1,027.4$ ,  $p<.001$ ,  $\eta_p^2=.46$ .

**Error data** The main effects of type of trial,  $F(2, 30)=24.4$ ,  $MSE=11$ ,  $p<.001$ ,  $\eta_p^2=.62$ , and word part,  $F(1, 15)=7.4$ ,  $MSE=8.4$ ,  $p=.016$ ,  $\eta_p^2=.33$ , were significant. Type of trial also interacted with word part,  $F(2, 30)=5.5$ ,  $MSE=7.9$ ,  $p=.009$ ,  $\eta_p^2=.27$ : Upper versus lower differences occurred with conflicting color words,  $F(1, 15)=7.35$ ,  $MSE=20.1$ ,  $p=.016$ ,  $\eta_p^2=.33$ , but not with congruent color words or noncolor words, for which both  $F$ s < 1.

Despite the fact that only the upper/lower parts of the words were displayed, the results revealed the typical Stroop effect: Congruent color words yielded the fastest responses (with fewer errors), whereas conflicting color words yielded the slowest responses (with more errors). Noncolor words were in the middle. More importantly, for the present purposes, naming times were greater for the upper part than for the lower part, thus demonstrating that this paradigm is sensitive to upper–lower differences.<sup>3</sup> This interaction could have been

<sup>3</sup> This upper–lower difference also occurred with the color words (which were upward-unbalanced, as well), but only in the conflicting trials (e.g., *rosa*; i.e., “pink” written in yellow). Although this null effect in congruent trials, which also occurred in Experiment 2, might be surprising at first glance, it is important to stress that previous research has shown that performance on congruent trials is not usually modulated by manipulations that affect performance on conflicting-color trials (e.g., Brown, Gore, & Carr, 2002; Tzelgov, Henik, & Berger, 1992). Indeed, facilitation and interference in the Stroop task are often (very) weakly and inversely correlated (e.g., Brown, 2011; Lindsay & Jacoby, 1994; MacLeod & MacDonald, 2000). Although the specification of facilitation/inhibition mechanisms in the Stroop task goes beyond the scope of the present article, the issue is that not all of the mechanisms that produce facilitation and interference are shared (see Brown, 2011).

**Table 1** Latencies and errors in Experiment 1 (upward-unbalanced noncolor words)

|       | Latencies                           |                                       |                                    | % Errors                            |                                       |                                    |
|-------|-------------------------------------|---------------------------------------|------------------------------------|-------------------------------------|---------------------------------------|------------------------------------|
|       | Congruent<br><i>M</i> ( <i>SD</i> ) | Conflicting<br><i>M</i> ( <i>SD</i> ) | Noncolor<br><i>M</i> ( <i>SD</i> ) | Congruent<br><i>M</i> ( <i>SD</i> ) | Conflicting<br><i>M</i> ( <i>SD</i> ) | Noncolor<br><i>M</i> ( <i>SD</i> ) |
| Upper | 520 (113)                           | 697 (123)                             | 640 (112)                          | 0.8 (1.5)                           | 8.7 (7.1)                             | 4.0 (3.3)                          |
| Lower | 521 (101)                           | 644 (111)                             | 600 (98)                           | 0.8 (1.5)                           | 4.4 (3.3)                             | 3.5 (2.8)                          |

For the baseline condition (&&&&&), the mean naming time was 562 ms (*SD*=103), and the error percentage was 2.0 % (*SD*=3.4).

due to (1) the upper part, but *not* the lower part, producing lexical interference, or (2) the upper part producing *more* lexical interference than the lower part. To examine the nature of this interaction and disentangle these two interpretations, we computed the differences between the response times on the critical trials (i.e., the noncolor words) and the trials with a string of &s (the baseline condition; see MacLeod, 2005, and Brown, 2011). This revealed a 78-ms lexical interference effect for the upper part of words [ $t(15)=8.24$ ,  $p<.001$ ,  $\eta^2_p=.82$ ]; this was reduced to 38 ms for the lower part of words [ $t(15)=4.58$ ,  $p<.001$ ,  $\eta^2_p=.58$ ]. Thus, both the upper and lower parts of the word produced lexical interference, and the amount of interference was greater for the upper than for the lower part.

Once we had demonstrated in a Stroop paradigm that the lexical interference of the upper part of upward-unbalanced words is greater than the effect of the lower parts, the critical question was whether this pattern would also occur with words that have a similar number of potentially confusable letters in their upper and lower parts (i.e., balanced words). This was the aim of Experiment 2.

## Experiment 2

### Method

**Participants** A group of 16 new students from the same population as in Experiment 1 took part in the experiment (13 women and three men).

**Materials** The color words were the same as in Experiment 1. In addition, we selected 24 noncolor words from the ESPAL database (Duchon et al., 2013). These noncolor words had a similar number of confusable letters in their lower and upper parts. On average, 23 % of the word constituent letters in these words were ambiguous in their lower part and 29 % in their upper part [ $t(23)=1.41$ ,  $p>.16$ ]. As in Experiment 1, there were six different noncolor words for every color word, while controlling for the number of letters and number of syllables with their corresponding color words (mean word frequency of the noncolor words: 51.29 per million). The list of noncolor words is available in the Appendix. As in Experiment 1, for

each word, we obtained the upper part version (e.g., *iaاااا*) and the lower part version (e.g., *ااااا*). The remaining details about the stimuli were the same as in Experiment 1.

**Procedure and design** These were the same as in Experiment 1.

### Results and discussion

Table 2 displays the mean naming latencies and percentages of errors across participants in the different conditions. In the latency analyses, 2.81 % of trials were excluded due to values lower than 200 ms or greater than 1,300 ms, or because the responses were incorrect. The statistical analyses were parallel to those conducted in Experiment 1.

**Naming times** The main effects of type of trial,  $F(2, 30)=97.01$ ,  $MSE=1,671.5$ ,  $p<.001$ ,  $\eta^2_p=.87$ , and word part,  $F(1, 15)=6.5$ ,  $MSE=760$ ,  $p=.022$ ,  $\eta^2_p=.3$ , were significant. The interaction between type of trial and word part was also significant,  $F(2, 30)=12.78$ ,  $MSE=655$ ,  $p<.001$ ,  $\eta^2_p=.46$ . This interaction revealed that, as in Experiment 1, the naming times for the conflicting color words were greater when the upper parts of the words were presented than when only the lower parts were presented (663 vs. 611 ms, respectively),  $F(1, 15)=12.84$ ,  $MSE=1,651.7$ ,  $p=.003$ ,  $\eta^2_p=.46$ . The naming times for congruent color words were similar for the upper and lower parts (494 vs. 501 ms, respectively),  $F(1, 15)=2.06$ ,  $MSE=222.27$ ,  $p=.17$ ,  $\eta^2_p=.12$ . More importantly, unlike in Experiment 1, the naming times of noncolor words were virtually the same for the upper and the lower parts (591 vs. 592 ms, respectively),  $F(1, 15)<1$ ,  $F_2(1, 23)<1$ .

**Error data** The ANOVA on the error data revealed main effects of type of trial,  $F(2, 30)=12.39$ ,  $MSE=19.72$ ,  $p=.0001$ ,  $\eta^2_p=.45$ , and word part,  $F(1, 15)=6.05$ ,  $MSE=6.3$ ,  $p=.027$ ,  $\eta^2_p=.29$ . The interaction between type of trial and word part was not significant,  $F(2, 30)=1.5$ ,  $MSE=9.1$ ,  $p=.24$ ,  $\eta^2_p=.09$ .

As expected, the results from the (upward-unbalanced) color words replicated the data from Experiment 1. However, the critical issue at stake was the outcome of the balanced noncolor words. The results were clear cut; we found no trend toward an upper–lower difference in the naming

**Table 2** Latencies and errors in Experiment 2 (balanced noncolor words)

|       | Latencies                           |                                       |                                    | % Errors                            |                                       |                                    |
|-------|-------------------------------------|---------------------------------------|------------------------------------|-------------------------------------|---------------------------------------|------------------------------------|
|       | Congruent<br><i>M</i> ( <i>SD</i> ) | Conflicting<br><i>M</i> ( <i>SD</i> ) | Noncolor<br><i>M</i> ( <i>SD</i> ) | Congruent<br><i>M</i> ( <i>SD</i> ) | Conflicting<br><i>M</i> ( <i>SD</i> ) | Noncolor<br><i>M</i> ( <i>SD</i> ) |
| Upper | 494 (95)                            | 663 (101)                             | 591 (83)                           | 0.8 (1.3)                           | 7.4 (8.5)                             | 2.7 (2.8)                          |
| Lower | 501 (85)                            | 611 (98)                              | 592 (86)                           | 0.5 (0.9)                           | 4.7 (5.1)                             | 2.0 (2.8)                          |

For the baseline condition (&&&&&), the mean naming time was 528 ms ( $SD=71$ ), and the error percentage was 1.2 % ( $SD=2.4$ ).

times (or percentages of errors) with balanced words. As with Experiment 1, it was important to examine whether the effects obtained with noncolor words were also due to lexical interference. To this end, we computed the differences between naming times on trials with noncolor words and trials with strings of &s (i.e., a measure of lexical interference). The amounts of the lexical interference effect were substantial and (unsurprisingly) similar in magnitude for the upper parts [63 ms;  $t(15)=7.92$ ,  $p<.001$ ,  $\eta_p^2=.81$ ] and for the lower parts [64 ms;  $t(15)=9.67$ ,  $p<.001$ ,  $\eta_p^2=.86$ ].

Finally, to assess the reliability of the differences between the noncolor words across experiments (upward-unbalanced words in Exp. 1 and balanced words in Exp. 2), we conducted a joint analysis on the latency data of noncolor words: The factors were Type of Noncolor Words (upward-unbalanced [Exp. 1], balanced [Exp. 2]) and Word Part (upper part, lower part). The ANOVA revealed a significant interaction between the two factors,  $F(1, 30)=19.27$ ,  $MSE=348.1$ ,  $p=.001$ ,  $\eta_p^2=.39$ , and  $F_2(1, 46)=10.27$ ,  $MSE=953.8$ ,  $p=.002$ ,  $\eta_p^2=.18$ . The latencies for the upper and lower parts differed in the case of the noncolor words in Experiment 1 (upward-unbalanced words), but not in Experiment 2 (balanced words).

## General discussion

The main aim of the present study was to examine whether the apparent bias toward the upper part of words reported in previous research on the Roman alphabet was the result of a genuine bias or was simply due to the potential confusability of the upper/lower parts of the words' constituent letters. To that end, we selected upward-unbalanced words in Experiment 1 (i.e., words that contained fewer confusable letters in their upper part than in their lower part; e.g., *molino* vs. *וווווווווו*) and balanced words in Experiment 2 (i.e., words that contained similar numbers of ambiguous letters in their upper and lower parts: e.g., *cinras* vs. *צצצצצצ*). To minimize the potential impact of the participants' guessing strategies, participants were not required to name these partial stimuli, but only to name the presentation color of the stimuli. The difference between the naming times to noncolor words and the baseline (a series of &s) was employed as an index of lexical interference (see Brown, 2011). We

found greater lexical interference for the upper than for the lower part in upward-unbalanced words (78 vs. 38 ms; Exp. 1). However, the amounts of lexical interference were similar for the upper/lower parts of balanced words (63 vs. 64 ms; Exp. 2), and the critical interaction across experiments was significant.

The simplest explanation of the present data is that the apparent bias for the upper part of words reported in previous experiments—from the simple demonstrations by Huey (1908) or Chou (1930) to recent experiments that have employed various techniques (e.g., Blais et al., 2009; Perea, 2012; Perea et al. 2012a; Perea et al. 2012b; Shimron & Navon, 1980; Tsao & Wang, 1983)—can be attributed to the fact that these languages simply include more upward-unbalanced words than balanced or downward-unbalanced words (i.e., English, Spanish, Portuguese, French, or Chinese). Indeed, when the words of a given language contain more downward-unbalanced words than upward-unbalanced words, as is the case with Hebrew, the lower part of words has an advantage when reading a text. Unlike in the Shimron and Navon experiment, we employed a subtle within-language manipulation (i.e., naming the presentation color) that minimized participant strategies to guess/reproduce the words' missing parts; these strategies may occur in reading experiments (see Grainger & Whitney, 2004).

It is important to stress that the important issue is not that the upper parts of isolated letters are more confusable than the lower parts of isolated letters in the Roman script. As we indicated in the introduction, Fiset et al. (2009) failed to find any evidence of such an advantage with isolated letters. The difference occurs because the letters that are ambiguous in their lower parts appear more frequently in languages that employ the Roman alphabet (e.g., English, French, or Spanish) than do letters that are ambiguous in their upper parts. As a result, words in these languages have a higher number of confusable letters in their lower than in their upper parts, and this is consistent with the upper-part advantage reported by Blais et al. (2009; see also Perea, 2012). In other words, we are describing a token rather than a type effect.

The present data are entirely compatible with current models of visual-word recognition and reading (e.g., the spatial-coding model of Davis, 2010, or the SERIOL model of Whitney, 2001, among others), because these models assume that, for a given letter, the effect of a particular letter

feature on the process of visual-word recognition does not depend on its (upper/lower) location in the letter. Thus, these models successfully predict no differences in lexical access when the potential ambiguity of the words' constituent letters is controlled; this occurred in Experiment 2. Furthermore, the interpretation that the upper-part advantage has an early locus at a letter-feature level is consistent with the additive effects found between word part and word frequency in previous experiments that have presented primes/previews and did not control for letter ambiguity (e.g., masked priming: Perea et al. 2012b; delayed-segment: Perea et al. 2012a; see also Perea, 2012, for similar evidence in early fixation measures). In particular, using the additive-factor logic, the effect of word part (lower vs. upper) would affect an earlier processing stage than would the effect of word frequency. Finally, it is important to note that, despite its potential ambiguity, the lower parts of upward-unbalanced words can produce some lexical activation/interference (a 38-ms lexical interference). This is consistent with the presence of reliable masked-priming effects with the lower parts of words that have been reported in previous research (i.e., מוּלִינֹ מוּלִינֹ can still produce some activation on the word *molino* ["mill"]; see Perea et al. 2012a; Perea et al. 2012b, for additional evidence with briefly presented foveal previews). It also reveals that the cognitive system can readily normalize partial information during lexical access (see, e.g., Jordan et al., 1999; Perea, Duñabeitia, & Carreiras, 2008).

It may be important to note that when determining whether letters were confusable in their upper or lower parts, we focused on the individual letters, and not on potential orthotactic constraints. As a reviewer pointed out, it may be argued that a letter that is potentially confusable would not be so if orthotactic constraints only allowed for one of the possible letters (e.g., the ambiguous *g* in *galope* [galope, the Spanish for gallop] would not be completely ambiguous, because the initial sequence *qa* is not legal in Spanish). The idea is that even though the participants did not have to name the words, this could influence automatic word recognition, and thus produce interference effects. If we had computed letter confusability by also taking into account orthotactic constraints in Spanish, the percentages of ambiguous letters in the words would decrease similarly (around 8%–10%) in all conditions (i.e., the categorization of the words in upward-unbalanced vs. balanced in the present experiments was not altered when taking into account orthotactic constraints). Clearly, the dissociation between lower- versus higher-level (orthotactic) constraints in the perception of potentially ambiguous letters embedded in words may be a relevant issue for further experimentation.

In sum, the present experiments have demonstrated that the apparent upper-part advantage is not a universal component in visual-word recognition and reading, but rather is an idiosyncratic feature of the words' constituent letters in languages that employ the Roman alphabet. Words in these languages tend to

have a higher number of confusable letters in their lower than in their upper part. When this factor is controlled, the apparent bias for the upper part of words vanishes completely, as we revealed in Experiment 2. Hence, the bias for the upper part of objects that occurs in the visual perception of particular features in faces (e.g., Caldara & Seghier, 2009) and objects (Thomas & Elias, 2011) does not seem to affect the automatic detection of letter features during visual-word recognition.

**Author note** The research reported in this article was partially supported by Grant No. PSI2011-26924 from the Spanish Ministry of Economy and Competitiveness. We thank Derek Besner, Tracy Brown, and an anonymous reviewer for helpful comments on an earlier version of this article.

## Appendix

### Noncolor words employed in Experiment: 1 (upward-unbalanced words):

alimento, amistad, anciano, anuncio, archivo, artistico, maduro, mancha, mantel, manual, marcar, master, medido, melena, mimado, molino, monton, movido, raiz, rama, raza, rico, ruso, ruta

### Noncolor words employed in Experiment 2 (balanced words):

apodo, cipres, coyote, cuba, espejo, europea, galope, guapa, guia, japonesa, jarabe, jaula, objeto, ocupado, pago, pasajero, peso, playa, quicio, separado, suya, viejo, yeso, zapatero

## References

- Behrmann, M., & Plaut, D. C. (2013). Distributed circuits, not circumscribed centers, mediate visual recognition. *Trends in Cognitive Sciences*, 17, 210–219. doi:10.1016/j.tics.2013.03.007
- Blais, C., Fiset, D., Arguin, M., Jolicœur, P., Bub, D., & Gosselin, F. (2009). Reading between eye saccades. *PLoS ONE*, 4, e6448. doi:10.1371/journal.pone.0006448
- Brown, T. L. (2011). The relationship between Stroop interference and facilitation effects: Statistical artifacts, baselines, and a reassessment. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 85–99. doi:10.1037/a0019252
- Brown, T. L., Gore, C. L., & Carr, T. H. (2002). Visual attention and word recognition in Stroop color naming: Is word recognition "automatic"? *Journal of Experimental Psychology: General*, 131, 220–240. doi:10.1037/0096-3445.131.2.220
- Burt, J. S. (2002). Why do non-color words interfere with color naming? *Journal of Experimental Psychology: Human Perception and Performance*, 28, 1019–1038. doi:10.1037/0096-1523.28.5.1019
- Caldara, R., & Seghier, M. L. (2009). The Fusiform Face Area responds automatically to statistical regularities optimal for face categorization. *Human Brain Mapping*, 30, 1615–1625. doi:10.1002/hbm.20626
- Caldara, R., Seghier, M. L., Rossion, B., Lazeyras, F., Michel, C., & Hauert, C. A. (2006). The fusiform face area is tuned for curvilinear

- patterns with more high-contrasted elements in the upper part. *NeuroImage*, 31, 313–319. doi:10.1016/j.neuroimage.2005.12.011
- Chou, S. K. (1930). Reading and legibility of Chinese characters: II. Reading half-characters. *Journal of Experimental Psychology*, 13, 332–351. doi:10.1037/h0075030
- Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, 117, 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
- 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. doi:10.3758/s13428-013-0326-1
- 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
- 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
- Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a whole? *Trends in Cognitive Sciences*, 8, 58–59. doi:10.1016/j.tics.2003.11.006
- Huey, E. B. (1908). *The psychology and pedagogy of reading*. New York: Macmillan.
- Ishai, A., Ungerleider, L. G., & Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron*, 28, 979–990. doi:10.1016/S0896-6273(00)00168-9
- Jordan, T. R., Thomas, S. M., & Scott-Brown, K. C. (1999). The illusory letters phenomenon: An illustration of graphemic restoration in visual word recognition. *Perception*, 28, 1413–1416. doi:10.1068/p2919
- Klein, G. S. (1964). Semantic power measured through the interference of words with color naming. *American Journal of Psychology*, 77, 576–588. doi:10.2307/1420768
- Lindsay, D., & Jacoby, L. (1994). Stroop process dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 219–234. doi:10.1037/0096-1523.20.2.219
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163–203. doi:10.1037/0033-2909.109.2.163
- MacLeod, C. M. (2005). The Stroop task in cognitive research. In A. Wenzel & D. C. Rubin (Eds.), *Cognitive methods and their application to clinical research* (pp. 17–40). Washington: American Psychological Association. doi:10.1037/10870-002
- MacLeod, C. M., & MacDonald, P. A. (2000). Inter-dimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences*, 4, 383–391.
- Perea, M. (2012). Revisiting Huey: On the importance of the upper part of words during reading. *Psychonomic Bulletin & Review*, 19, 1148–1153. doi:10.3758/s13423-012-0304-0
- Perea, M., Comesaña, M., & Soares, A. P. (2012a). Does the advantage of the upper part of words occur at the lexical level? *Memory & Cognition*, 8, 1257–1265. doi:10.3758/s13421-012-0219-z
- Perea, M., Comesaña, M., Soares, A. P., & Moret-Tatay, C. (2012b). On the role of the upper part of words in lexical access: Evidence with masked priming. *Quarterly Journal of Experimental Psychology*, 65, 911–925. doi:10.1080/17470218.2011.636151
- Perea, M., Duñabeitia, J. A., & Carreiras, M. (2008). R34D1NG WORD5 WITH NUMB3R5. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 237–241. doi:10.1037/0096-1523.34.1.237
- Protopapas, A. (2007). CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, 39, 859–862. doi:10.3758/BF03192979
- Shimron, J., & Navon, D. (1980). The distribution of visual information on the vertical dimension of Roman and Hebrew letters. *Visible Language*, XIV, 5–12.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 12, 643–662. doi:10.1037/h0054651
- Thomas, N. A., & Elias, L. J. (2011). Upper and lower visual field differences in perceptual asymmetries. *Brain Research*, 1387, 108–115. doi:10.1016/j.brainres.2011.02.063
- Tsao, Y.-C., & Wang, T.-G. (1983). Information distribution in Chinese characters. *Visible Language*, 17, 357–364.
- Tzelgov, J., Henik, A., & Berger, J. (1992). Controlling Stroop effects by manipulating expectations for color words. *Memory & Cognition*, 20, 727–735. doi:10.3758/BF03202722
- Warren, R. E. (1972). Stimulus encoding and memory. *Journal of Experimental Psychology*, 94, 90–100. doi:10.1037/h0032786
- Wechsler, H., & Bell, A. (2005). *Speed reading for professionals* (Barron's Educational Series). New York: Barron's.
- Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, 8, 221–243. doi:10.3758/BF03196158