Cognition 132 (2014) 491–500

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Are root letters compulsory for lexical access in Semitic languages? The case of masked form-priming in Arabic

Manuel Perea [a,b,*], Reem Abu Mallouh [b], Manuel Carreiras [b,c,d]

[a] ERI-Lectura and Departamento de Metodología, Universitat de València, Valencia, Spain
[b] Basque Center for Cognition, Brain, and Language, Donostia-San Sebastián, Spain
[c] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
[d] Departamento de Lengua Vasca y Comunicación UPV/EHU, Bilbao, Spain

ARTICLE INFO

Article history:
Received 18 September 2011
Received in revised form 13 January 2014
Accepted 10 May 2014
Available online 20 June 2014

Keywords:
Lexical access
Word recognition
Masked priming

ABSTRACT

Do Semitic and Indo-European languages differ at a qualitative level? Recently, it has been claimed that lexical space in Semitic languages (e.g., Hebrew, Arabic) is mainly determined by morphological constraints, while lexical space in Indo-European languages is mainly determined by orthographic constraints (Frost, Kugler, Deutsch, & Forster, 2005). One of the key findings supporting the qualitative difference between Semitic and Indo-European languages is the absence of masked form priming in Hebrew/Arabic with productive words. Here we examined whether masked form priming occurs in Arabic words when one of the letters from the productive root is replaced in the prime stimulus by another letter. Results showed a significant masked form priming effect with the lexical decision task in three experiments (including yes/no, go/no-go, and sandwich priming), to a similar degree to that reported in previous research with Indo-European languages. These data support the view that the processing of word forms in Semitic vs. Indo-European languages differs more at a quantitative than at a qualitative level.

Introduction

In recent years, there has been increasing interest in the study of how the peculiarities of a language may shape the process of visual-word recognition (see Frost, 2012, for review). Critically for the present purposes, it has been claimed that lexical space in Semitic languages (e.g., Hebrew, Arabic) is primarily determined by morphological constraints (via the root morphemes) whereas lexical space in Indo-European languages is mainly affected by orthographic constraints (Frost, 2009; see also Velan & Frost, 2011). According to Frost, Kugler, Deutsch, and Forster (2005), "the perceptual distance between two words containing different roots [in Semitic languages] would be uncorrelated with their overall orthographic similarity." (p. 1296) This affirmation is based, originally, on one key effect that differs in Indo-European and Semitic languages: While there are a number of reports of facilitative masked form (orthographic) priming with one-letter substitution nonword primes in Indo-European languages (i.e., spuce-SPACE being responded to faster than wudow-SPACE; e.g., *English*: Forster, Davis, Schoknecht, & Carter, 1987; *French*: Ferrand & Grainger, 1992; *Spanish*: Perea & Rosa, 2000; *Dutch*: Brysbaert, 2001), this is absent in Semitic languages (see Frost et al., 2005; Velan & Frost, 2011, for failures to obtain a masked form priming effect in Hebrew and Arabic). Frost (2012) has argued that, unlike Indo-European languages, the "orthographic coding scheme of Hebrew print focuses mainly on the few letters that carry morphological information, whereas the other letters of the word do not serve for lexical access, at

* Corresponding author. Address: Departamento de Metodología, Facultad de Psicología, Av. Blasco Ibáñez, 21, 46010 Valencia, Spain.
E-mail address: mperea@valencia.edu (M. Perea).

http://dx.doi.org/10.1016/j.cognition.2014.05.008
0010-0277/

least not initially" (p. 9). Thus, in this view, sharing all the root letters would be a prerequisite for a facilitative masked priming effect in Semitic languages (Frost, 2009).

Before describing the relevant findings in detail, it is important to mention two key particularities of words in Semitic languages. First, Semitic words can be decomposed into two discontinuous morphemes: (i) a consonantal root (usually composed of three [or four] letters) which provides the core meaning of the word (e.g., the Semitic root k.t.b [writing or marking]; for transcriptions, we employ the Buckwalter transliteration scheme; see Boudelaa & Marslen-Wilson, 2010); and (ii) and a phonological word-pattern, which conveys morphosyntactic and phonological information. (In the following examples, the Cs represent the consonantal pattern.) Each set of root letters, together with a word pattern, can lead to a large number of words: كتاب kitAab = book (word pattern: CiCaaC), كاتب kAtib = writer (word pattern: CaCiC), مكتب maktab = office (word pattern: maCCaC), كتب kutub = books (word pattern: CuCuC), مكتوب maktuub = written (word pattern: maCCuuC) – note that short vowels in Arabic or Hebrew are not typically written down (e.g., the word kitAab [book] would be written as ktAB) so that words convey mostly consonantal information. In addition, in Hebrew and Arabic, there is a small proportion of: (i) words with a non-productive root (i.e., a consonantal root that only appears in that word, accompanied by a word pattern); and (ii) loan words that cannot be decomposed into a root and a word pattern (i.e., they have no internal structure). Second, the percentage of words sharing the same letters (in different order) in Semitic languages is dramatically higher than in Indo-European languages. This is so because the root letters may appear in different combinations forming unrelated morphological families (e.g., s.b.H ["to swim"], s.H.b ["to withdraw"], H.s.b ["to calculate"], or H.b.s ["to imprison"]). Indeed, the transposed-letter effects that can be easily obtained in Indo-European languages when transposing two letters in the lexeme (i.e., "cholocate" being processed as "chocolate"; e.g., see Perea & Lupker, 2004) is noticeably smaller when transposing two letters in the root in Semitic languages (Hebrew: Velan & Frost, 2007; Arabic: Perea, Abu Mallouh, & Carreiras, 2010).

But does the empirical evidence actually support "the qualitative difference in processing base forms" (Velan & Frost, 2011, p. 154) between Semitic and Indo-European languages? Frost et al. (2005; Experiment 2) failed to find significant masked form priming effects (via one-letter substitution primes) in Hebrew both when the word targets were composed of productive roots and when the word targets were composed of non-productive roots (the priming effects were 2 and 4 ms, respectively). This was interpreted as reflecting that, in Hebrew, "even words that are not morphologically complex are not stored according to a purely orthographic code" (p. 1306). However, Velan and Frost (2011, Experiment 4) conducted a parallel experiment and found a significant 11-ms masked form priming effect for word targets composed of non-productive roots. The parallel effect for the word targets composed of productive roots was a nonsignificant 6-ms priming effect. In that same experiment, Velan and Frost also reported a significant 16-ms masked form priming effect when the word targets were loan words (AGRTL אגרטל [a vase]) with no internal root + word pattern structure. The critical interaction between word type (productive, non-productive, loan) and prime condition (related, unrelated) in Velan and Frost's Experiment 4 was not close to significance in the analysis by items ($F2 < 1$). Thus, at the item level, all three types of words in the Velan and Frost experiment (i.e., including the words with productive roots) would be responsible for the main effect of masked form priming. Velan and Frost (2011) concluded, notwithstanding, that while Semitic words are clustered together in the lexical space as a function of their root letters (i.e., thus not showing masked form priming); words whose roots are not productive and words with no internal structure would be processed in a similar way to Indo-European languages (i.e., thus showing masked form priming). However, as Velan and Frost acknowledged, "this duality is not parsimonious." (p. 154).

There are many theorists – from different standpoints – who assume that the differences in word processing between Semitic and Indo-European languages are quantitative rather than qualitative. In their dual-route approach, Grainger and Ziegler (2011) indicated that the findings obtained in Hebrew could be due to a different balance in the priority given to fine-grained versus coarse-grained orthographic information relative to Indo-European languages. While fine-grained orthographic information "optimizes processing via the chunking of frequently co-occurring contiguous letter combinations", coarse-grained orthographic information "optimizes the mapping of orthography to semantics by selecting letter combinations that are the most informative with respect to word identity (diagnosticity), irrespective of letter contiguity" (Grainger & Ziegler, 2011, p. 3). In particular, Grainger and Ziegler (2011) argued that the lack of transposed-letter priming in productive Semitic words occurs because priority is given to fine-grained information when processing productive Semitic words (i.e., a quantitative rather than a qualitative difference among families of languages). Likewise, Davis (2012) argued that "the same coding and processing mechanisms as the spatial coding model" (p. 21) could be successfully applied to Semitic languages – for instance, Davis argued that the inhibitory transposed-letter priming with word–word pairs reported by Velan and Frost (2009) and Perea et al. (2010) could be explained in terms of lexical inhibition at the lexical level. Similarly, Whitney (2012) claimed that an open-bigram coding scheme could be well suited to Semitic languages and that any differences between Semitic and Indo-European languages in orthographic processing would reflect quantitative rather than qualitative differences.

One aspect in which the word-forms from Semitic and Indo-European languages may differ is in the internal structure – in particular in the regularities of root + word pattern sequences of Semitic words. It may be reasonably argued that readers of Semitic languages pick up the statistical regularities of these languages and this may alter how the word-forms are processed in comparison to word-forms from Indo-European languages. To examine this issue, Lerner, Amstrong, and Frost (2013) employed a multi-layer neural network that mapped orthographic inputs from five-letter words to semantic outputs via back-propagation. In