Are all Semitic languages immune to letter transpositions?

The case of Maltese

Manuel Perea [1 2], Albert Gatt [3], Carmen Moret-Tatay [4], and Ray Fabri [3]

[1] ERI-Lectura, Universitat de València, Valencia, Spain

[2] Departamento de Metodología, Universitat de València, Valencia, Spain

[3] Institute of Linguistics, L-Università ta' Malta, Msida, Malta

[4] Departamento de Metodología, Universidad Católica de Valencia, Valencia, Spain

Word count: 3929 (main text, footnotes, references, and author's notes)

Running head: Transpositions in Maltese

Address for correspondence:

Manuel Perea
Departamento de Metodología
Facultad de Psicología
Av. Blasco Ibáñez, 21
46010-Valencia (Spain)
FAX: +34 96 3864697
Email: mperea@valencia.edu

Abstract

Recent research using the Rapid Serial Visual Presentation (RSVP) paradigm with English sentences including words with letter transpositions (e.g., jugde) has shown that participants can readily reproduce the correctly spelled sentences with little cost; in contrast, there is a dramatic reading cost with root-derived Hebrew words (Velan & Frost, 2007, 2011). This divergence could be due to: i) the processing of root-derived words in Semitic languages, or ii) the peculiarities of the transitional probabilities in root-derived Hebrew words. Unlike Hebrew, Maltese is a Semitic language which does not omit vowel information in print and whose morphology also has a significant non-Semitic (mostly Romance) morphology. Here we employed the same RSVP technique as Velan and Frost (2007, 2011), this time with Maltese (and English) sentences. Results showed that Maltese-English bilinguals were able to reproduce the Malterse words –regardless of whether they were misspelled (involving the transposition of two letters from the consonantal root) or not, with no reading cost—just like in English. The apparent divergences between the RSVP data with Hebrew vs. Maltese sentences are likely to be due to the combination of the characteristics of the Hebrew orthographic system with the Semitic morphology.

As White, Johnson, Liversedge, and Rayner (2008) reported, raeding wrods with jubmled lettres involves some cognitive cost in terms of reading time, but participants can read the words correctly. Indeed, it was more than fifty years ago that Bruner and O'Dowd (1958) found that nonwords created by letter transpositions (e.g., jugde) resembled their base words to a large degree (see also O'Connor & Forster, 1981, for early evidence). The transposed-letter effect has been obtained across a variety of paradigms, not only in other Indo-European languages (e.g., *Spanish*: Perea & Lupker, 2004; *French*: Schoonbaert & Grainger, 2004) but also in other families of languages (e.g., *Basque*: Perea & Carreiras, 2006; *Japanese Kana*: Perea & Pérez, 2009; *Thai*: Winskel, Perea, & Ratitamkul, in press; *Korean Hangul*: Lee & Taft, 2009).

The above-cited cross-linguistic evidence may be taken to indicate that letter position coding is an inherent characteristic of any orthographic system –and as such is considered in the current input coding schemes of visual-word recognition (e.g., *spatial coding model*: Davis, 2010; *SERIOL model*: Whitney, 2001; *open-bigram model*: Grainger & van Heuven, 2003; *overlap model*: Gomez, Ratcliff, & Perea, 2008; *noisy Bayesian Reader model*: Norris, Kinoshita, & van Casteren, 2010). However, in a recent study, Velan and Frost (2007) argued that letter transpositions may merely "reflect the specificities of the lexical structure of a given language" (p. 914) rather than being a

universal phenomenon (see also Frost, in press). The evidence for their argument comes from an experiment using a rapid serial visual presentation (RSVP) of sentences –each word was presented for 200 ms. The experiment was conducted both with English sentences and with Hebrew sentences on Hebrew-English bilinguals; in half of the cases, the sentences were correctly written (e.g., “My sister accused me of lifting some skirts from her closet”), whereas in the other half, three words were misspelled via adjacent transpositions of internal consonants (e.g., “My sitser accused me of litfing some skitrs from her closet”). The participants’ task was to reproduce the whole (correctly spelled) sentence without replicating the actual transpositions. For the English sentences, participants were able to produce the target words around 85% of the time, regardless of whether several words were presented jumbled or intact (84 vs. 86%, respectively); similarly, participants were able to reproduce around 81.5% of the words in the sentences, regardless of whether the sentences had jumbled words or were presented intact (81 vs. 82%, respectively). That is, there was very little cost produced by the jumbled words in English, consistent with the accuracy scores in eye movement experiments that employ transposed-letter words embedded in sentences (White et al., 2008). But the critical finding was that, for the Hebrew sentences, recollection of the target words was around 25% lower when three of the words had letter transpositions in

the consonantal root than when the sentences were correctly spelled (59 vs. 84% for the words in sentences with jumbled letters and intact sentences, respectively); likewise, the percentage of recollection of words in the sentences was also substantially higher when the sentences were presented intact than when the sentences had jumbled words (81 vs. 62%, respectively). That is, Hebrew readers had some difficulty reconstructing the actual transpositions in the Hebrew sentences. Velan and Frost (2007) concluded that "effects of letter transposition probably reflect the principles of defining lexical space and lexical organization, and do not emerge from the peripheral registering of letters in alphabetic orthographies" (p. 916). Their view is that lexical space in Semitic languages would be structured according to the morphological roots, whereas Indo-European languages would be structured according to full orthographic/phonological forms (see Frost, Kugler, Deutsch, & Forster, 2005; Perea, Abu Mallouh, & Carreiras, 2010, for parallel evidence from another Semitic language, Arabic). Thus, the transposition of two letters of the consonantal root in Semitic languages would make it difficult to access the actual word.

Velan and Frost (2011) successfully replicated their 2007 RSVP experiment with a new set of English sentences and a new set of Hebrew sentences with root-derived words. Importantly, they also included a third set of Hebrew sentences in

which the transposition occurred in morphologically simple words –from a non-Semitic origin (e.g., אגרטל AGRTL [a vase]). In this latter case, there was only a small cost (around 4%), which was similar to that found with English sentences. Velan and Frost (2011) explained this very small reading cost as indicating that "Hebrew words that are morphologically simple, do not contain a root, and do not have any internal structure, are lexically organized by orthographic neighbourhoods just like base-words in English" (p. 152). In this respect, it is important to note that root-derived words in Hebrew have a "well-defined set of conditional probabilities that rigidly determine few open slots for the consonants of the root only" (Velan & Frost, 2011, p. 153), which implies a "high saliency of the root morpheme" (see Frost, in press, for further discussion). Indeed, this saliency may apparently be picked up by learners of Hebrew as an L2. Frost et al. (2005; Experiment 3) reported that native English speakers whose L2 was Hebrew showed the same pattern of masked priming effects as native speakers of Hebrew (i.e., identity but not form priming).

Here we examine whether the special status of root-derived words in Hebrew (or Arabic) can be generalized to another Semitic language, Maltese. There are two possible reasons why Maltese is a potentially interesting testbed for the results found in other Semitic languages. First, unlike Hebrew and Arabic, Maltese represents both

consonants and vowels in the orthography, which employs the Latin alphabet. Second, Maltese differs from Hebrew or Arabic in that it has a very productive non-Semitic (mostly Romance) morphology in addition to the Semitic component (see Mifsud, 1995a for a thorough descriptive characterization).

As in other Semitic languages, many verbs and (some) nouns in Maltese can be represented as a root and a word pattern. For instance, the Maltese word miktub ["written"] would be formed by the Semitic root k.t.b (with the meaning of "marking, writing") and the word pattern mi--u-. Experimental evidence suggests that, similarly to Hebrew and Arabic, the consonantal root in Maltese plays an important role in lexical access (e.g., see Twist, 2006; Ussishkin & Twist, 2009). However, there are reasons to believe that Maltese root-and-pattern morphology is no longer truly productive. For instance, as Hoberman and Aronoff (2003) note, the most productive derivation rule in the verbal component of Maltese, accounting for many novel formations, is based on the Italian *–are* and operates on borrowings. Hence, from the borrowed (English) word *monitor*, via initial germination and suffixation of *–ja*, we obtain the verb *immoniterja* ('to monitor'). While this is far from the only case, it is illustrative of a general tendency in Maltese morphology to exhibit productive concatenative, rather than root-and-pattern, word formation processes.

Thus, despite the evident psychological importance of the root (Twist, 2006), words formed out of Semitic roots appear to constitute a fixed list, with no novel formations (Mifsud, 1995b). As shown in a recent study, based on an exhaustive survey of standard lexicographic resources (Spagnol, 2011), roots in the verbal component occur on average in only two out of ten possible *binyanim* (range: 1 to 5). There is furthermore a marked asymmetry among the *binyanim* in their frequency of occurrence, with over 75% of all verbs being in binyan I, II and V. This need not imply that Semitic words are *infrequent* – indeed, the materials used in our study include a number of very frequent formations, as we show below. However, frequency does not imply productivity (e.g., Baayen, 2009).

Given the strong evidence for the hybrid nature of the Maltese morphological system (Fabri, 2009), as well as the differences in the writing system compared to other Semitic languages, the question arises whether the consonantal root plays as crucial a role in reading as it does in Hebrew and Arabic. In the present experiment, we employed the same RSVP technique and procedure as Velan and Frost (2007, 2011). The only difference is that instead of presenting sentences in English vs. sentences in Hebrew, we presented sentences in English vs. sentences in Maltese. The letter transposition in the Maltese sentences always involved the transposition of two consonants of the root in

words of Semitic origin –as in the Velan and Frost (2007, 2011) experiments. For comparison purposes, we employed the same set of 20 English sentences employed by Velan and Frost (2011) –for these sentences we expect little/no differences between the sentences correctly spelled and the sentences with two transposed-letter words.

If the pattern of data in Maltese sentences is similar to that of Hebrew sentences (i.e., a substantial reading cost in sentences with jumbled root-derived words), this would reinforce that view that Semitic morphology governs the way words are processed, regardless of the presence/absence of vowel information in the orthography and/or the unproductivity of root-based morphological processes.

Alternatively, if the pattern of data in Maltese is similar to that found in English and with morphologically simple words in Hebrew (i.e., little or no reading cost in sentences with jumbled root-derived words in Maltese), this would raise at least three possibilities. One is that the Semitic root in Maltese, despite evidence for its role in lexical access, simply does not have the saliency that it has for Hebrew readers (Velan and Frost 2011). A second possibility, clearly related to the first, is that Maltese does not evince the same rigid letter position coding of Hebrew and other Semitic languages, that is, the root has limited utility in determining transitional probabilities in Maltese words. Finally, orthographic differences (the presence of vowels) may also play a role.

Method

Participants. Twenty students from the University of Malta took part in the experiment. They were recruited from Linguistics courses on Maltese at the undergraduate/graduate level. All of them reported having Maltese as their mother tongue and using Maltese on a daily basis. They also were highly fluent in English –which is the official second language (after Maltese) in Malta –all students at the University of Malta have their instruction in English, unless the instruction is part of a course related to a specific language (Maltese, French, Spanish, etc). Most students reading for a degree in Maltese will also have a second subject, for which instruction is likely to be in English. Apart from an early exposure to English on a day-to-day basis being likely for a large proportion of Maltese children, the language forms part of their formal education from the first year of primary education (age 5) in both State-run and independent (private) schools.

Materials. We created 20 sentences in Maltese such as the one shown below.

Ilbieraħ Marija libset id- dublett l- iswed li kienet xtrat Londra

Yesterday Marija wear-3SgF.PERF DEF-skirt DEF-black that AUX.3SGF.PAST buy.3SGF.PERF London

"Yesterday Maria put on the black skirt she had bought in London"

In each sentence, we transposed two adjacent middle consonants from the root letters of two of the words (e.g., libset →lisbet and iswed→iwsed; the roots are l.b.s and s.w.d

respectively). The average number of letters of the target words was 5.8 (range: 5-8), the mean word frequency of the root used in the target word (i.e. the number of times the root occurred in any morphological derivation) was 28.27 per million tokens (range: 0.14 - 480), and the mean number of orthographic neighbors of the target words was 3.8 (range: 1-13). Neighbourhood density and frequency were calculated on the basis of a wordlist compiled from the 99.2 million word corpus of the Maltese Language Resource Server (MLRS, which is available at http://mlrs.research.um.edu.mt). As in the Hebrew experiments, words with transposed letters were never consecutive. The individual sentences, as well as a spreadsheet containing frequency characteristics of the target words in Maltese, are available at http://staff.um.edu.mt/albert.gatt/pubs/rsvp2012.html. In addition, we employed the 20 sentences in English from Experiment 1 of Velan and Frost (2011). For both the Maltese and the English sentences, we created two counterbalanced lists of 10 intact sentences and 10 sentences with jumbled words (i.e., if a given sentence was presented intact in List 1, it would be presented with jumbled words in List 2, and vice versa). Thus, each participant was presented 20 Maltese sentences in one block (10 intact and 10 with jumbled words), and 20 English sentences in the other block (10 intact and 10 with jumbled words). There were two blocks in the experiment: half of the participants were initially tested with the Maltese sentences, and

then the English sentences, while the other half were tested in the opposite order.

Procedure. Participants were tested in groups of three to four in a quiet room. Presentation of the stimuli was controlled by Windows-based computers using DMDX (Forster & Forster, 2003). Each trial began when the participant pressed the space bar. Then, each word of the sentence was presented for 200 ms on the center of the screen. As in the typical RSVP experiment (see Forster, 1970), participants were instructed to write down the sentence after each trial, or at least the words they could see in the order that they were presented. Participants wrote sentences down on paper. They had been alerted that some of the sentences could involve jumbled words and that they had to produce the sentences with correctly spelled words. The instructions and examples in the Maltese block were given in Maltese, while the examples in the English block were given in English. Sentences within each block (English or Maltese) were presented in randomized order for each participant.

Results

As in the Velan and Frost (2007, 2011) RSVP experiments, we computed, for each participant in each language, two dependent variables: on the one hand, we calculated the overall percentage of correct report of words, both in intact sentences and in the sentences containing jumbled words; on the other hand, we computed the

percentage of target words produced in each language (20 target words in each language) which were presented intact or jumbled in the RSVP sentences. The averages per participant in each language for intact sentences and for sentences with jumbled words are shown in Table 1.

Please_insert_Table_1_around_here

Percent report of all words. The percent report of all words was very similar for intact sentences and for sentences with jumbled words, both in English (88.7 vs. 85.6%, respectively, $F<1$) and in Maltese (88.5 vs. 88.8, respectively, $F<1$). Note that the percent report of all words for the intact sentences (i.e., the "baseline") was also similar for English and Maltese sentences, $F<1$.

Percent report of target words. Participants were able to reproduce the target words, in a similar proportion of cases, when these words were presented intact and when these words were presented jumbled, both in English (78.0 vs. 76.5, respectively, $F<1$) and in Maltese (81.0 vs. 81.5, respectively, $F<1$). There were no significant differences between the percentage of report of target words for the intact sentences in English and Maltese ($F<1$).[1]

[1] We conducted some post hoc analyses to examine whether the pattern of data changed depending on whether the letter transposition produced an existing (57% of target words) or a non-existing root (43% of target words), but the pattern of data was essentially the same as that reported here. Note that all words produced by transposing root letters were non-words.

Discussion

Unlike other Semitic languages, Maltese uses the Latin script, it does not omit vowel information in print, and it also has productive non-Semitic morphology. This allowed us to examine, in a writing system similar to that of Indo-European languages, whether Maltese-English bilinguals have difficulty producing the target words in RSVP in Maltese sentences when two root letters from a Semitic Maltese word were transposed. The findings were very clear: participants were able to reproduce these words –regardless of whether they were misspelled or not. This pattern of data closely resembled that of English sentences –thus successfully replicating the results of Velan and Frost (2007, 2011) in English. Therefore, unlike the Hebrew data with reported by Velan and Frost (2007, 2011), in which there was "a dramatic drop in performance" when the sentences included root-derived jumbled words, here we found no signs of a parallel effect with Maltese root-derived words (see Table 1). Thus, readers of Maltese, a Semitic language, do not have much trouble in reading transposed words of Semitic origin –or rather do not have trouble reconstructing sentences containing jumbled root-derived words (e.g., libset→lisbet).

In the Introduction, we suggested that such a finding might have three possible

explanations. One is that differences in the orthographic systems of Maltese and other Semitic languages – especially the presence of vowels – may help to explain the results. However, the mere omission of (short) vowels in Hebrew orthography cannot be the only cause of the difficulties in reading transposed text, because –similarly to the present experiment in Maltese– Velan and Frost (2001) demonstrated that Hebrew readers can easily reconstruct transposed-letter words in an RSVP task when these are morphologically simple. This suggests that part of the explanation must lie in the differences between the morphological systems of the languages in question. Here, we considered two related factors.

One possibility is that the lack of productivity of root-and-pattern morphology in Maltese may result in lower saliency of the root for Maltese readers. Support for this claim comes from the observation that Maltese has a highly productive non-Semitic concatenative morphology (e.g., Mifsud, 1995a, 1995b; Spagnol, 2011) in addition to the (diachronically prior) root-and-pattern morphology characteristic of Semitic languages. However, this is unlikely to be the whole story, given previous experimental work suggesting that the root plays a role in lexical access (e.g. Twist, 2006; Ussishkin and Twist, 2009).

Perhaps more importantly, the centrality of the root in Hebrew and Arabic has

been cited as a crucial factor in determining transitional probabilities between letters. To explain the difficulty of reproducing jumbled root-derived words, Frost (in press) indicated that, because of the distribution of transitional probabilities between letter sequences, the Hebrew reading system focuses “on a specific subset of letters and is rigid regarding their position” –the consonantal root, in particular.

We believe that the apparent divergences between the RSVP data with Hebrew vs. Maltese sentences with morphologically complex target words are due to a combination of the characteristics of the Hebrew writing system with aspects of the Semitic morphology of Hebrew versus Maltese. In particular, the presence of a more flexible process of letter position coding with root-derived words in Maltese than in Hebrew (i.e., two Semitic languages) provides some empirical support to the claim that the flatter the distribution of transitional probabilities of letter sequences, the more the orthographic code will be flexible in letter position (see Frost, in press). Future computational/empirical studies should examine in detail the transitional probabilities in different languages and how they affect the process of letter position coding across languages.

In sum, we believe that further research in Maltese, a Semitic language in which both consonants and vowels are written down (as in Indo-European languages)

using the standard Latin alphabet and which, in addition, has a mixed morphology, opens up a window of opportunity to examine the intricacies of Semitic morphology and its similarities/differences with respect to Indo-European morphology.

References

Baayen, R.H. (2009) Corpus linguistics in morphology: morphological productivity. In A. Luedeling and M. Kyto (Eds.), *Corpus Linguistics. An international handbook*. Berlin: Mouton De Gruyter.

Bruner, J.S., & O'Dowd, D. (1958). A note on the informativeness of parts of words. *Language and Speech, 1,* 98-101.

Fabri, R. (2009). Stem allomorphy in the Maltese verb. *Ilsienna – Our Language, 1*, 1-20.

Forster, K.I. (1970). Visual perception of rapidly presented text segments of varying complexity. *Perception and Psychophysics, 8*, 215-221.

Forster, K.I., & Forster, J.C. (2003). DMDX: A Windows program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 16–124.

Forster, K.I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *Quarterly Journal of Experimental Psychology, 39A,* 211-251.

Frost, R. (in press). Towards a universal model of reading. *Behavioral and Brain Sciences*.

Frost, R., Kugler, T., Deutsch, A., & Forster, K.I. (2005). Orthographic structure versus morphological structure: principles of lexical organization in a given language. *Journal of Experimental Psychology: Learning, Memory & Cognition, 31*, 1293-1326.

Gómez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review, 115*, 577-601.

Hoberman, R.D., & Aronoff, M. (2003). The verbal morphology of Maltese: From

Semitic to Romance. In Shomron, J. (Ed.), *Language processing and acquisition in languages of Semitic, root-based morphology,* 61-78. Amsterdam: John Benjamins.

Lee, C.H., & Taft, M. (2009). Are onsets and codas important in processing letter position? A comparison of TL effects in English and Korean. *Journal of Memory and Language, 60*, 530-542.

Mifsud, M. (1995a). *Loan verbs in Maltese: A descriptive and comparative study*. New York: Brill.

Mifsud, M. (1995b). The productivity of Arabic in Maltese. In J. Cremona, C. Holes and G. Khan (Eds.), *Proceedings of the 2nd Conference of the International Association for Arabic Dialectology (AIDA).* Cambridge: University Publications Centre.

O'Connor, R.E., & Forster, K.I. (1981). Criterion bias and search sequence bias in word recognition. *Memory and Cognition, 9*, 78-92.

Perea, M., Abu Mallouh, R., & Carreiras, M. (2010). The search of an input coding scheme: Transposed-letter priming in Arabic. *Psychonomic Bulletin and Review, 17*, 375-380.

Perea, M., & Lupker, S.J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language, 51*, 231-246.

Perea, M., & Carreiras, M. (2006). Do transposed-letter effects occur across lexeme boundaries? *Psychonomic Bulletin and Review, 13*, 418-422.

Perea, M., & Pérez, E. (2009). Beyond alphabetic orthographies: The role of form and

phonology in transposition effects in Katakana. *Language and Cognitive Processes, 24*, 67-88.

Schoonbaert, S., & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes, 19*, 333-367.

Spagnol, M. (2011). *A tale of two morphologies: Verb structure and argument alternations in Maltese*. PhD Dissertation, University of Konstanz.

Twist, A. (2006). *A psycholinguistic investigation of the verbal morphology of Maltese*. PhD Dissertation, Department of Linguistics, University of Arizona, Tucson.

Ussishkin, A., & Twist, A. (2009). Auditory and visual lexical decision in Maltese. In B. Comrie, R. Fabri, E. Hume, M. Mifsud, & M. Vanhove (Eds.), *Introducing Maltese linguistics* (pp. 233-249). Amsterdam: John Benjamins.

Velan, H. & Frost, R. (2007). Cambridge University Vs. Hebrew University: The impact of letter transposition on reading English and Hebrew. *Psychonomic Bulletin & Review, 14*, 913-918.

Velan, H., & Frost, R. (2011). Words with and without internal structure: what determines the nature of orthographic and morphological processing? *Cognition, 118*, 141-156.

White, S.J., Johnson, R.L., Liversedge, S.P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of word-beginning letters. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 1261-1276.

Winskel, H., Perea, M., & Ratitamkul, T. (in press). On the flexibility of letter position coding during lexical processing: Evidence from eye movements when reading

Thai. *Quarterly Journal of Experimental Psychology*. DOI: 10.1080/17470218.2012.658409

Table 1. Percent report of all words (top) and percent report of target words (bottom) in Maltese and in English with intact sentences and sentences with jumbled words. Standard deviations are shown in parentheses.

	Intact sentences	Sentences with jumbled words
All words		
Maltese	88.5 (11.9)	88.8 (7.0)
English	88.7 (9.5)	85.6 (19.4)
Target words		
Maltese	81.0 (14.7)	81.5 (11.3)
English	78.0 (15.3)	76.5 (20.7)

Author's notes

The research reported in this article has been partially supported by Grants PSI2011-26924 and CONSOLIDER-INGENIO2010 CSD2008-00048 from the Spanish Ministry of Education and Science and by Grant X14-31-514 from the University of Malta. We thank Ram Frost, Jon Andoni Duñabeitia, and Adam Ussishkin for helpful comments on an earlier version of the paper.

Correspondence concerning this article should be addressed to: Manuel Perea, Departamento de Metodología, Facultad de Psicología, Av. Blasco Ibáñez, 21, 46010-Valencia, Spain (e-mail: mperea@valencia.edu).