Discussion

Linguistic theory and psychological reality: a reply to Boudelaa & Marslen-Wilson

Shlomo Bentin*, Ram Frost

Department of Psychology, Hebrew University of Jerusalem, Mount Scopus, Jerusalem 91905, Israel

In a very interesting paper, Boudelaa and Marslen-Wilson (in press) revive a 19th century theory of Semitic language morphology suggesting that the major morphological unit that conveys the core meaning of words (mainly verbs) in Semitic languages is a bi-consonantal structure labeled "etymon" (Gesenius, 1817, edited and enlarged by Kautzsch).[1] From a historical perspective, this idea has been raised several times during the last century, eliciting wide discussions and disagreement among linguists (see Kurylowicz, 1973). The strongest argument in favor of the existence of such a morphemic unit came from the analysis of bi-consonantal ancient nouns in Arabic and Hebrew (Nöldeke, 1910; for more recent references see Bohas, 1977; Bohas & Darfouf, 1993; McCarthy, 1979).

The concept of the etymon as the cornerstone of Semitic morphology is in sharp contrast to the well-accepted view that considers the three-consonantal root to be the kernel and major carrier of meaning in Hebrew and Arabic words. Although linguistically plausible, the etymon theory has never gained wide support, and probably cannot be unequivocally proven through linguistic investigation. Arguments against the etymon as a theoretical construct usually rely on the verbal system of Semitic languages, demonstrating that without exception, verbs which were derived from ancient two-consonantal nouns have always expended to include another consonant to create the classic three-consonantal structure (e.g. Blau, 1971). More importantly, the major weakness of the etymon as a theoretical construct is the lack of constrained rules for determining which two consonants of a given word form the etymon (Y. Blau, pers. commun.).

Boudelaa and Marslen-Wilson (B&MW) were not concerned with the linguistic status of the etymon theory, but focused on the psychological reality of the etymon.

They report a set of experiments exploring cross-modality and masked priming effects between words sharing the etymon, with or without a semantic relation between primes and targets. The effect of sharing the etymon on lexical decision performance was assessed in comparison with pairs of words in which the targets and the primes had a similar surface (phonological) structure, or were completely unrelated. The results revealed that lexical decisions were considerably faster (60–70 ms in cross-modal priming and 20–30 ms in masked priming) if the targets and the primes shared a common etymon than if they had different etymons. Semantic overlap did not interact with morphological priming in any of the tasks. Prima facie, these results strongly support the psychological validity of the etymon in word recognition. Indeed, on the basis of this outcome, B&MW embraced an approach to the morphological structure of Semitic languages in which the etymon (and not the root) is the primary organizing unit in the lexicon. By their account, the third letter of the traditional root is added in order to provide the skeletal morpheme imposed by the surface structure of the absolute majority of words in Semitic languages. In contrast, the etymon is a higher abstract morphological unit, part of the lexical entry of the word.

The results of this study are seemingly compelling. However, we would like to raise some theoretical caveats about the B&MW interpretation. Although the historical perspective regarding the rise and fall of the etymon as a linguistic construct is interesting, this aspect of the etymon controversy is beyond the scope of our reply (for a critical review, see however Kurylowicz, 1973). In the present reply, like B&MW, we will limit ourselves to the issue of its psychological reality. We begin by outlining the complications that the concept of the etymon poses for any model of morphological representation. Then we raise some queries about the results obtained by B&MW and end by suggesting a possible alternative interpretation of their findings.

The idea that two letters are the kernel morpheme carrying the core meaning of the words in Semitic languages is difficult to maintain because of several reasons. First, there are no a priori and clearly defined rules for morphological decomposition that would result in the unequivocal stripping or isolation of the etymon letters or phonemes. In contrast to the root, which is conspicuous in the word's structure and is derived by considering the nominal or verbal patterns of words (Feldman, Frost, & Pnini, 1995; Frost, Deutsch, & Forster, 2000), finding the etymon of a word seems to be a very difficult task. Morphological decomposition into etymons is especially complicated because, in contrast to the root, there are no constraints on the order by which the two letters of the etymon should appear within the word. Indeed, there are many examples in the stimuli provided by B&MW in which the two letters of the etymon appear in one order in the prime and in another order in the target. Apparently, one could make the claim that "b" and "l" are the etymon of [mu**b**ta**l**lun] and [waa**bil**un] but not of [**b**a**l**iidun] because the first two are semantically related. But as B&MW admit, a semantic relationship is not a necessary condition for determining the etymon. Since processes of morphological decomposition are determined by the interrelation of form transparency and distributional properties (e.g. Laudanna & Burani, 1995; Schreuder & Baayen, 1995), a theory of decomposition into etymons

does not seem very compelling. Whether Arabic speakers are sensitive to the etymon and can extract it from printed words is, however, an empirical question. We addressed this question through a short paper-and-pencil experiment.

Because the concept of etymon is not taught in schools (at least not in the Arab schooling system in Israel), we did not expect our participants to have a priori knowledge of the etymon and be able to overtly identify it without explicit instruction. We assumed, however, that if the etymon is a psychologically real morpheme, native Arabic speakers should be able to identify it once it is defined. To test this assumption, we selected at random 20 prime–target pairs from those used by B&MW in their study. These pairs were presented in random order in a list of 40 words to 20 Arab speakers, students at the Hebrew University who were instructed as following: "You will be presented with a list of words. In each word, some letters have a greater importance in conveying the origin or the meaning of the word. For each of the 40 words in the list, please mark the two most important letters that convey its meaning or origin in the best way." Obviously, as speakers of a Semitic language, we expected the participants to focus on the letters of the root. We were intrigued, however, whether they would be accurate in identifying the two letters of the etymon among the three letters of root.

For each participant we calculated accuracy separately among the words used by B&MW as primes, and among the words used by B&MW as targets. More importantly, we counted the number of pairs in which the participants accurately identified the etymon in *both* the prime and its matched target. Note that the extraction of the same etymon in the prime and in the target is the basis for expecting etymon-based priming effects. The results showed that, on the average, our students correctly identified the etymon of only 27.2 words (68%) out of the 40 words in the list, equally distributed among the primes and among the targets. This level of accuracy is lower than one would expect if the etymon would be a psychologically real morpheme, particularly noticing that there are only three possible two-letter combinations in a given three-letter root. The weakness of this performance is even more conspicuous considering that native speakers of Semitic languages can easily report without error what letters of the word belong to the root, even at the first grades of primary school. Even more problematic was the fact that in only 50% of the pairs (10.1 out of 20) was the same etymon correctly identified for *both* the prime and the target. This outcome is particularly disturbing because any claims regarding priming effects rely on the implicit assumption that the same etymon is perceived for the prime and the target. Our present test revealed that this assumption might be incorrect.

Although, in general, native speakers acquire meta-awareness of morphemic units, which may account for morphological priming, the results above suggest that explicit knowledge of Arabic speakers about etymons could not have accounted for the large priming effects observed by B&MW. The alternative explanation is that these effects reflected knowledge accumulated implicitly by exposure to the statistical regularity that exists between orthographic and phonological sublexical units and semantic features (see, for example, Plaut & Shallice, 1993; Seidenberg, 1987; Seidenberg & McClelland, 1989; Van Orden, Pennington, & Stone, 1990).

However, an account based on implicit learning is also difficult to maintain. As we mentioned above, in the absence of a priori rules for morphological decomposition, without minimal constraints on the order of the letters, and with only weak, or no correlations between letters and semantic meaning, it is almost impossible to build up statistical regularities that are the basis for implicit learning.

In addition to the theoretical and empirical complications that emerge assuming that etymons are primary morphemic units in Arabic, we were also puzzled by some aspects of the data reported by B&MW. The first puzzle concerns the almost identical priming effect in the [+Etym +Sem] condition and the [+Etym −Sem] condition in the cross-modality paradigm. Even assuming that the morphological priming effect is independent of semantics (and perhaps particularly assuming independence) larger priming effects should have been found in the [+Sem] condition than in the [−Sem] condition. Indeed, Frost, Deutsch, Gilboa, Tannenbaum, and Marslen-Wilson (in press) have shown that morphological priming effects in Hebrew increase with semantic similarity in the cross-modal paradigm. Similarly, in the study of Bentin and Feldman (1990), morphological and semantic priming effects (within modality) were found to be additive. As both of these studies examined Hebrew, a Semitic language, they provide an unexplained contrast with the findings of B&MW. The second puzzle concerns the size of the morphological masked priming effects, which is unusually large. In the last few years Frost and his colleagues have reported results from an extensive series of masked priming experiments in Hebrew, which focused on the effect of shared morphemic units between targets and primes on word recognition (Deutsch, Frost, & Forster, 1998; Frost et al., 2000, in press; Frost, Forster, & Deutsch, 1997). A meta-analysis of their results across studies revealed a surprisingly stable size of the root priming effect, which is about 13 ms (±2 ms). Indeed, the facilitation induced by the shared two letters of the etymon was sometimes larger than the effects reported by Frost and his colleagues in the full identity priming conditions. Considering that the root is a much more conspicuous morphological unit than the etymon, the much larger effects reported by B&MW for the etymon are hard to account for. Furthermore, the two etymon letters do not necessarily appear in the same order within the prime and the target, and masked priming is sensitive only to form and not to semantic overlap.

What alternative factors might account for the priming effects in the [+Etym] conditions reported by B&MW? One factor to be considered is the baseline from which the priming effects were computed, that is, the "unrelated" condition. Although B&MW went a long way towards controlling for form overlap, some discrepancies in the phonological similarity between primes and targets in the [+Etym] and [−Etym] conditions may still remain. B&MW argue perhaps convincingly why form overlap could not be calculated linearly given the non-concatenative morphology of Arabic. Nevertheless, at least in some of their stimuli the primes in the [+Etym] conditions were systematically more similar in form to the targets than the primes in the [−Etym] conditions. Taking the first stimulus of the appendix as an example, FAASILUN seems to be more phonologically similar to MUNTA-SAFUN than NADIMUN is to MUNTASAFUN. This discrepancy in similarity is particularly salient if we focus on the root letters. Whereas the prime in the [+Etym]

condition shares two root letters with the target, the prime in the [−Etym] condition shares only one. We admit that we cannot unequivocally determine whether some hidden phonological factors indeed affected the results reported by B&MW, and we may be proven wrong. However, because the theory offered by B&MW has such far-reaching implications, we believe that additional experiments and conducted replications are required before it is unquestionably embraced.

Acknowledgements

References

Bentin, S., & Feldman, L. B. (1990). The contribution of morphological and semantic relatedness to the repetition effect at long and short lags: evidence from Hebrew. *Quarterly Journal of Experimental Psychology*, *42A*, 693–711.

Blau, Y. (1971). *Phonology and morphology (in Hebrew)*. Tel Aviv: Hakibbutz Hameuchad Press.

Bohas, G. (1977). *Matrices, étymons, racines, elements d'une théorie lexicologique du vocabulaire arabe. Orbis Supplementa*, Louvain-Paris: Peeters.

Bohas, G., & Darfouf, N. (1993). Contribution à la reorganization du lexique de l'arabe, les etymons nonordonnés. *Linguistica Comunicatio*, *1*, 55–103.

Boudelaa, S., & Marslen-Wilson, W. D. (in press). Morphological units in the arabic mental lexicon. *Cognition*, (*this issue*).

Deutsch, A., Frost, R., & Forster, K. I. (1998). Verbs and nouns are organized and accessed differently in the mental lexicon: evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1238–1255.

Feldman, L. B., Frost, R., & Pnini, T. (1995). Decomposing words into their constituent morphemes: evidence from English and Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 947–960.

Fleisch, H. (1961). *Traite de philologie arabe* (p. 256). Beyrouth: Impr. Catholique.

Frost, R., Deutsch, A., & Forster, K. I. (2000). Decomposing morphologically complex words in a nonlinear morphology. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 751–756.

Frost, R., Deutsch, A., Gilboa, O., Tannenbaum, M., & Marslen-Wilson, W. D. (in press) Morphological priming. Dissociation of phonological, semantic, and morphological factors. *Memory and Cognition*.

Frost, R., Forster, K. I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew: a masked priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 829–856.

Gesenius, W. (1817). *Gesenius' Hebrew grammar*. Edited and enlarged by E. Kautzsch (1909). English edition by A.E. Cowley (1970). Oxford: Clarendon Press.

Kurylowicz, J. (1973). *Studies in Semitic grammar and metrics*. London: Curzon Press.

Laudanna, A., & Burani, C. (1995). Distributional properties of derivational affixes: implications for processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 345–364). Hillsdale, NJ: Lawrence Erlbaum.

McCarthy, J. J. (1979). *Formal problems in Semitic phonology and morphology*. Unpublished doctoral dissertation. MIT, Cambridge, MA.

Nöldeke, T. (1910). *Neue zeibrage fur semitischen sprachwissenschaft*. Strassburg: K.J. Trubner Verlag.

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: a case study of connectionist neuropsychology. *Cognitive Neuropsychology*, *10*, 377–500.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing*. Hillsdale, NJ: Lawrence Erlbaum.

Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: access units or orthographic redundancy? In M. Coltheart (Ed.), *Attention and performance XII* (pp. 244–263). Hillsdale, NJ: Lawrence Erlbaum.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, *97*, 488–522.