# Words matter more than morphemes: Evidence from masked priming with bound-stem stimuli

**Hélène Giraudo**
Laboratoire CLLE (Equipe ERSS)
CNRS & Université Toulouse Jean Jaurès
giraudo@univ-tlse2.fr

**Madeleine Voga**
University Paul-Valéry Montpellier III
madeleine.voga@univ-montp3.fr



## Abstract

Five masked priming studies were carried in order to shed light on the processing of bound-stem words (e.g., *terr-* in *terrible*). Both orthographic (e.g., *termite*) and unrelated (e.g., *montagne* 'montain') conditions stand as baselines for controlling morphological effects. The results of the experiments using unrelated word controls suggest that in the particular case of bound-stem words, only genuinely derived word primes (*terrible*) produce positive effects differing from formal overlap effects. Morphological effects are interpreted as resulting from both "morceme" and "base-lexeme" activations.

## 1 Introduction

As is broadly admitted, morphologically related words prime each other in various languages (Arabic: Boudelaa & Marslen-Wilson, 2001; English: Rastle, Davis, Marslen-Wilson & Tyler, 2000; French: Giraudo & Grainger, 2000; German and Dutch: Drews and Zwitserlood, 1995; Hebrew: Frost, Deutsch & Forster, 1997) thus suggesting the existence of a morphological level of processing. This kind of study has used different types of materials, words or pseudowords, as well as multiple settings: for the masked priming technique (Forster & Forster, 2003), widely used to shed light on morphological processes as well as in this study, the distinction can be made between designs using only unrelated controls and those using both unrelated and orthographic/phonological controls, as suggested by Giraudo & Grainger (2001) or Pastizzo & Feldman (2002).

Even though the existence of a morphological level of processing is unanimously acknowledged, the exact nature, locus and the role of these representations within the mental lexicon is a matter of ongoing research. Two hypotheses can be drawn: according to the first, morphemic units correspond to concrete pieces of words (i.e., stems and affixes). Complex words are therefore processed through a decomposition mechanism that strips off the affix in order to isolate the stem. The morphemic nature of the remaining letters is then verified by the system and access to word representations (i.e., word forms coded in the orthographic lexicon) operates via the pre-activation of their constituent morphemes, i.e., morphemic representations stand as access units. This mechanism is exemplified by Taft's model (1994), the basic principles of which are followed by many psycholinguistic studies (e.g., Crepaldi, Rastle & Davis, 2010). Morphemic units are situated between the level of letters/syllables and the word level; consequently, they can only be matched to concrete letter clusters (i.e., bound-stems, free-stems and affixes) that constitute words. This decompositional mechanism is also insensitive to any semantic characteristics of words (i.e., transparent vs. opaque morphological formation) or to their lexical environment (in terms of orthographic neighborhood or family size). One of the strong predictions of the decompositional approach is that morphological priming effects should vary following the ease with which constituent morphemes can be identified/extracted.

According to the second hypothesis, morphology is coded at the interface of word and semantic representations and corresponds rather to lexemes (Aronoff, 1994). Lexeme units are coded at the interface of the word and the semantic level, organizing the lexicon in terms of morphological families. The recognition of any complex word triggers first the activation of all word forms that can match with it; a competition is then engaged between the pre-activated forms (forms matching the input, i.e., those who are morphologically


In Vito Pirrelli, Claudia Marzi, Marcello Ferro (eds.): *Word Structure and Word Usage*. Proceedings of the NetWordS Final Conference, Pisa, March 30-April 1, 2015, published at http://ceur-ws.org

related but also those who are only orthographically related) until the right lexical representation reaches its recognition threshold (determined by its surface frequency). During this competition phase, morphologically related words send positive activation to their respective base lexeme, feeding back activation to them. Morphological priming effects result from this mechanism of co-activation. Following this supralexical approach (Giraudo & Grainger, 2000; 2001), complex words are not "decomposed", but are able to trigger the activation of their constituent morphemes. In this kind of architecture, lexeme units are supposed to be abstract enough to tolerate variation induced by derivation and inflection (i.e., allomorphy, suppletion, phonological/ morphological truncation, haplology, verb-noun conversion). In other words, a morphological unit does not necessarily need to surface in the real world in order to be coded in long-term memory. This organisation, compatible with recent neuroimaging data (Lévy, Hagoort, Démonet, 2014), also implies that all morphemes of a given language are not necessarily represented within the mental lexicon: units such as neologisms, hapaxes and nonce words are not necessarily directly connected with existing morphological units; bound-stem words could be such a case.

## 2 The study

The present paper focuses on the processing of bound-stem words by opposition to free-stem words. For ex., on one hand, the word *viral* composed of the bound-stem *vir-,* also present in *virus*, *virulent*, *virulence*, *virology* and *virologist* and, on the other hand the word *singer* composed with the free-stem *sing* that forms *singing*, *song*, etc. Both are defined as being morphologically complex but while it is evident for the standard speaker/reader that the complex word *singer* derives from the root *sing*, it is less evident to say from which root the complex word *viral* derives. The morpheme *vir-*, which does not have any clear meaning in English, can be considered as a bound-stem whereas *sing-* in *singer* is a free-stem. From a processing point of view, the *vir-viral* example can be viewed as a case where the lexical unit is not directly connected to the morphological unit, by virtue of its twofold handicap: the first aspect is semantic interpretability, i.e., derivations composed with a bound-stem could be less interpretable than those with a free-stem. Psycholinguists tested this difference and found that processing for free and bound-stems may differ but both produce significant priming effects (Forster & Azuma, 2000; Järvikivi & Niemi, 2002; Marslen-Wilson, Tyler, Waksler & Older, 1994; Pastizzo & Feldman, 2004). Of great importance to our study, Pastizzo and Feldman (2004) observed that the magnitude of facilitation varied following the baseline used in the experiments: equivalent magnitudes of priming for free and bound-stems were obtained relative to an unrelated baseline; with an orthographic control however, free-stems produced systematically greater priming than bound-stems. The interpretation of this line of research suggests that morphological priming effects are not directly constrained by semantic similarity between prime and target. The second handicap, in terms of surface analysis, consists in the difficulty in segmenting the word forms into morphemes. At this point, the two different models presented above give rise to different predictions: according to the morpheme-based approach all complex forms (free-stem as well as bound-stem words) are first analyzed in morpheme fragments and then access word representation, in other words, the lexicality of the base doesn't matter. This approach predicts morphological priming between derivations (e.g., *virus-viral*) as well as between the base and its derivation (e.g., *vir-viral*).

According to the supralexical approach, the members of a morphological family are linked together by virtue of their common base at the lexeme level; however, the base of bound-stem words is not represented at the word level. In this case, priming effects between related derived words (e.g., *virus-viral*) are expected but no effect should be observed using their bound-stems as primes, the access to the base lexeme being conditioned by the prior activation of a word form at the word level.

Taft and Kougious (2004) investigated this issue in English through a masked priming experiment. They compared both semantically and orthographically related words (e.g., *virus-viral*) to merely orthographically related words (e.g., *future-futile*) and, unsurprisingly, found facilitation in the former case but not in the latter. Nevertheless, the design of this study is not very informative with respect to the decomposition issue, given that the critical condition examining the effect of the bound-stem on its derivations has not been considered.

Our study aims to fill this gap through five visual masked priming experiments with native French speakers. In this kind of protocol, subjects are unaware of the presence of the prime, which allows minimizing strategy use and examining automatic processing during the early stages of word identification: all five experiments use a within-priming (Latin square) design, in which we directly compare the effects of different primes on the same target. A 57ms prime duration) was used and the task was lexical decision.

Exp. 1 examined morphological effects induced by words sharing the same bound-stem, e.g., *terrible – terreur 'terrible-terror'* relative to an orthographic control baseline, e.g., *termite – terreur*' *'termite-terror'* (where *'termite'* is a monomorphemic word), as well as an unrelated baseline (*montagne – terreur* 'mountain-terror'). Results show that only truly derived word primes produce facilitation, relative to unrelated (36ms of facilitation) as well as orthographic controls (35ms). However, this first result does not inform us about how derived words constructed with a bound-stem are processed: are they analyzed in terms of stem + affix or are they globally processed? Exp. 2 examined the extent to which the facilitation we take as morphological could be due to formal overlap: this is done by using non-existent orthographic controls, sharing all but one letter with the 'true' bound-stem, e.g., for the target *terreur,* the first possible prime is the true bound-stem *terr-* presented in isolation (e.g., *terr* - *terreur*); the second priming condition is the non-existing bound-stem condition (orthographic control) *tarr-* (e.g., *tarr* – *terreur*); the third condition is an unrelated baseline (e.g., *montag* – *terreur*). Although only true bound-stems induced significant facilitation relative to the unrelated baseline (28ms), the non-existing stem condition (e.g., *tarr-*) exhibited reaction times (RTs) that did not differ significantly from those of the true bound-stem condition. This result highlights the fact, already pointed out by Forster (1999), that there is an influence of formal factors in this kind of protocol, as well as the need to include orthographic controls in the design. Experiment 3 directly compared the effects of complex word primes to those of bound-stem primes: the targets were the same as in Exp. 1 and the three levels of the prime type factor were the following: a morphologically related suffixed word sharing the same bound-stem, e.g., *terrible* – *terreur 'terrible-terror'*; its bound-stem, e.g., *terr* - *terreur*; an unrelated control, e.g., *montagne* – *terreur*. Results showed that only complex word primes (e.g., *terrible*) produced significant priming effects (33ms), though these conditions did not significantly differ from the bound-stem condition (e.g., *terr-*) whose effect (18ms) did not manage to reach significance.

Exp. 4 was designed to see if the advantage for the complex word sharing the same bound-stem found in exp. 3 holds up to the comparison with non-word primes constructed with the same bound-stem and an existing suffix. The three priming conditions were the following: the morphologically related word sharing its bound-stem with the target, e.g., *terrible – terreur*; a non-word made of the same bound-stem and a suffix different to that of the target, e.g., *terrage* – *terreur* (where *-age* corresponds to an existing morpheme); an unrelated control, e.g., *montagne* – *terreur*. The statistical analysis of the results revealed that only related word primes (e.g., *terrible*) produced significant morphological priming (40ms) relative to the unrelated controls. Even if the non-word prime condition (e.g., *terrage*) led to quicker reaction times compared to the unrelated baseline (688 vs 703ms), it didn't differ significantly from it. More importantly, the 25ms difference between the word prime condition and the non-word one is statistically significant. This suggests that it takes a real word to induce morphological priming, independently and above orthographic low-level perceptual influences, to which the masked priming technique is known to be sensitive. Our results show that the presence of an existing bound-stem in a non-word does not suffice to produce morphological priming, a finding which contradicts those published by Longtin and Meunier (2005) as we shall see in the discussion. Experiment 5 examined the extent to which the morphological facilitation found in exp. 4 could be due to formal factors: in order to test this, we replaced the morphologically related word primes by non-words constructed with a bound-stem and a final letter sequence that does not correspond to any suffix in French. The following three prime conditions defined the three levels of the prime type factor: a complex non-word formed by a bound-stem and a suffix, e.g., *terrage* – *terreur* (where *terr-* and *–age* correspond to existing morphemes); a simplex non-word formed by a bound-stem and a non-existing ending, e.g., *terryme* – *terreur,* in which *–yme* is not a suffix; finally, an unrelated non-word, e.g., *moitagne* – *terreur*. The statistical analysis of the results revealed that both complex and simplex non-word primes produced shorter RTs than unrelated primes (31 and 27ms of effect respectively): both types of prime are able to facilitate target recognition and produce thus morphological-like facilitation. Nevertheless, the fact that the effects produced by complex primes (e.g., *terrage*) did not differ from those produced by simplex non-word primes (e.g., *terryme*) leads us to reject any interpretation based on pre-lexical morphological decomposition. We suggest interpreting this pattern of results on the basis of formal criteria: for real words it takes a real word to facilitate processing (exp. 4), but for non-words, given the absence of representation in the word-level, morphological-like priming does nothing but reflect low-level perceptual similarities, such as between the two non-words (both complex and simplex, *terrage/terryme*) and the target *terreur.* Besides the role attributed to formal factors, the point that should be stressed in the interpretation of exp. 4 and 5 is that while in exp. 4 the nonword made up from an existing bound-stem and an existing suffix (*terrage*) seems to interfere with processing of the target (*terreur*) by virtue of its morphological structure, in exp. 5 this interference disappears. The fact that, in the '*terrage'* condition (exp. 4) we observe RTs that are not significantly quicker than the unrelated condition, despite the existence of a formal overlap combined with morphological-like structure (*terr-age/terr-eur*), can only be due to some kind of interference, otherwise we should observe at least a small formal effect. This interference nevertheless disappears in exp. 5, since both types of non-words (with existing suffix, e.g., *terrage,* as well as well as non-existing suffix, i.e., simplex non-words such as *terryme*) lead to significant facilitation. We therefore obtain a different pattern of priming for words (exp. 4) and for non-words (exp. 5) which leads us towards an approach where lexicality of the prime does matter in the overall pattern of results. Even if the processing system can take advantage of orthographic similarities between prime and target (and will not prevent itself from doing so, as exp. 2 showed) this does not tell the whole story, and it certainly does not tell a morphological story: it is just another demonstration of a fact that researchers working with masked priming are familiar with, namely that this technique is sensitive to formal factors (Forster, Mohan & Hector, 2003). The experiments presented here provide evidence that we can use this valuable technique in order to shed light on truly morphological effects, as opposed to morphological-like effects.

Taken together, the results of the experiments using unrelated word controls (exp. 1, 3 and 4) suggest that in the particular case of bound-stem words, only genuinely derived word primes (*terrible*) produce positive effects differing from formal overlap effects. This is true with the exception of exp. 3, where the effect of genuinely derived word primes did not differ from bound-stem primes (*terr-*); note however that in this experiment, the bound-stem condition did not differ from the unrelated condition, while the derived word condition did. This is a demonstration of the fact that "nonwords would be always better form-primes than words, even when masked. The reason is simply because a related word prime will compete more vigorously with the target than a related nonword prime" (Forster, 1999: 8). These results are not in accordance with those found by Longtin and Meunier (2005) using roughly the same priming conditions. In their study, derived non-word primes (e.g., *garagité*) systematically produced significant priming effects on target recognition relative to unrelated word controls, while non-morphological non-word primes (e.g., *rapiduit*) yielded a 29 ms non-significant effect. Two factors can explain these contradictory results: a) the type of unrelated controls: contrary to Longtin and Meunier, we examined priming effects relative to unrelated non-word primes when the prime conditions included non-words and word primes when the prime conditions included words; b) the type of word targets: given that our study focuses on bound-stem words, our targets are mandatorily complex words, and not bare-bases, as in the Longtin & Meunier study. Bare-bases are by definition more frequent, and, subsequently, easier to activate because of their lower activation threshold (due to their residual activation; for a discussion on this point based on McClelland & Rumelhart 1981, see Voga & Giraudo, 2009; Giraudo & Voga, 2014).

## 3 Discussion: On the representation of bound-stem words

On the basis of the above results, we can conclude that recognition of complex words benefits from two springs of facilitation: a bottom-up excitation from a sublexical level and a top-down facilitation from a supralexical level. The idea of a double representation for morphology was recently expressed by Diependaele, Sandra & Grainger (2005), suggesting that the morphological level should be situated both above and below the word-form level. Subsequently, morphological representations would be either defined as morphologically constrained orthographic representations (depending on frequencies) or as morphologically constrained semantic representations (coded in terms of regularities in the mapping of word forms onto semantics). In the same line, Crepaldi et al. (2010) proposed an extension of Taft's (1994) sublexical model integrating a lemma level comprised between an orthographic lexicon and the semantic system. However, these two models consider the two morphological levels equivalent, given that they both contain units corresponding to concrete morphemes. One may nevertheless assume that different locations imply different contents: the hybrid model we propose (Giraudo & Voga, 2014) is based exactly on this assumption. Within this model, morphological complex words are coded according to two dimensions, their surface form and their internal structure. The first level captures the statistical regularities of morphemes translated in terms of perceptual saliency in the language. At this level, morphologically complex and pseudo-derived words as well as non-words whose surface structure can be divided into distinct morphemes, are equally processed. This level is not a morphological level but rather a sub-orthographic level containing "morcemes". The second level, i.e., the morphological level is paradigmatically oriented, it deals with the construction of words according to morphological rules (Booij, 2005; Corbin, 1987/1991); it contains "base-lexemes", units abstract enough to tolerate orthographic and phonological variations produced by derivation and inflection processes and connected to their related word forms on the basis of semantic transparency.

## References


Aronoff, M. (1994). *Morphology by itself.* Cambridge: MIT Press.

Boudelaa, S., & Marslen-Wilson, W. D. 2001. Morphological units in the Arabic mental lexicon. *Cognition*, 81:65-92.

Booij, Geert. 2005. Compounding and derivation: evidence for Construction Morphology. In W.U. Dressler, D. Kastovsky and F. Rainer (eds) *Demarcation in Morphology*. Amsterdam / Philadelphia: John Benjamins. 109-132.

Corbin, D. (1987/1991). *Morphologie derivationnelle et structuration du lexique*, vol. 2. Tubingen/Villeneuve d ' Ascq: Max Niemeyer Verlag / Presses Universitaires de Lille.

Crepaldi, D., Rastle, K., Davis, C. 2010. Morphemes in their place: Evidence for position-specific identification of suffixes. *Memory & Cognition,* 38:312-321.

Drews, E. & Zwitserlood, P. (1995). Morphological and orthographic similarity in visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance,* 21:1098-1116.

Forster, K. I. (1999). The microgenesis of priming effects in lexical access. *Brain and Language*, 68:5–15.

Forster, K.I. and Azuma, T. 2000 Masked priming for prefixed words with bound-stems: Does submit prime permit? *Language and Cognitive Processes,* 15(4-5):539-561.

Forster, K.I., & Forster, J.C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavioral Research Methods: Instruments & Computers,* 35:116-124.

Frost, R., Deutch, A., & Forster, K.I. 2000. Decomposing morphologically complex words in a nonlinear morphology. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 26:751-765.

Giraudo, H. & Grainger, J. 2000. Prime word frequency in masked morphological and orthographic priming. *Language and Cognitive Processes,* 15:421-444.

Giraudo, H., & Grainger, J. 2001. Priming complex words: Evidence for supralexical representation of morphology. *Psychonomic Bulletin & Review,* 8(1):127-131.

Giraudo, H. & Voga, M. (2014). Measuring Morphology: the tip of the iceberg ? A retrospective on 10 years of morphological processing. *Carnets de Grammaire*, 22:136-167.

Järvikivi, J., & Niemi, J. 2002. Form-based representation in the mental lexicon: Priming (with) bound stem allomorphs in Finnish. *Brain & Language,* 81:412-423.

Lévy, J., Hagoort, P., & Démonet, J.F. 2014. A neuronal gamma oscillatory signature during morphological unification in the left occipitotemporal junction. *Human Brain Mapping*, 35(12):5847-60.

Longtin, M.-C. & Meunier, F. 2005. Morphological decomposition in early visual word processing, *Journal of Memory and Language,* 53(1), 26-41.

McClelland, J. L., D.E. Rumelhart. 1981. An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review,* 88: 375-407.

Marslen-Wilson, W. D., Tyler, L. K., Waksler, R., & Older, L. 1994. Morphology and meaning in the English mental lexicon. *Psychological Review,* 101:3-33.

Pastizzo, M. J. and Feldman, L. B. 2004. Morphological processing: A comparison between free and bound-stem facilitation. *Brain & Language,* 90:31-39.

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. 2000. Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes,* 15:507-537.

Taft, M. 1994. Interactive activation as a framework for understanding morphological processing. *Language & Cognitive Processes,* 9:271-294.

Taft, M., & Kougious, P. 2004. The processing of morpheme-like units in monomorphemic words. *Brain & Language,* 90:9-16.

Voga, M. & H. Giraudo. 2009. Pseudo-family size influences processing of French inflections: evidence in favor of a supralexical account in F. Montermini, G. Boyé, J. Tseng (eds), *Selected Proceedings of the 6th Décembrettes: Morphology in Bordeaux.* Somerville, MA: Cascadilla Proceedings Project, 148-155.