Are prefixed units processed and represented like suffixed ones? Towards a hybrid model of morphological processing

Hélène Giraudo, Madeleine Voga

Are prefixed units processed and represented like suffixed ones? Toward a hybrid model of morphological processing.

Hélène GIRAUDO (1) & Madeleine VOGA (2)

(1) CNRS & Université de Toulouse
(2) Université Paul-Valéry Montpellier III

The fact that in most languages affixed words are present in a very high proportion leads to the conclusion that morphology constitutes an important variable in word processing. Thirty years of investigation permitted to confirm that morphology intervenes in automatic processes that operate during the very early stages of lexical access, suggesting that morphemes were independently coded or stored somewhere in the mental lexicon. The masked priming paradigm (Forster & Davis, 1984) is the privileged technique used by the psycholinguists to examine the early processes of word recognition. The principle governing this paradigm lies indeed on the transfer of activation from a first processed stimulus (the prime) on the recognition latency of a second stimulus (the target). This activation transfer is admitted to operate on the basis of the shared representations (orthographic/phonological/morphological/semantic) by prime-target pairs. Moreover, given that the prime is presented very briefly (SOAs under 60 ms) and is generally masked (by a string of hash marks), any effect of the prime is considered to be the result of unconscious processes. In the precise case of morphology, many studies manipulated morphologically related words as well as pseudo-words and found systematically very robust positive priming effects: two morphologically related words prime each other across different languages (e.g., Boudelaa & Marslen-Wilson, 2005 in Arabic; Duntildeabeitia, Laka, Perea, & Carreiras, 2009 in Basque; Drews and Zwitserlood, 1995, in both German and Dutch; Frost, Deutsch & Forster, 1997 in Hebrew; Giraudo & Grainger, 2000 in French; Rastle, Davis, Marslen-Wilson & Tyler, 2000 in English) and in experimental settings that include multiple control priming conditions (unrelated but also orthographic/phonological and semantic controls in order to neutralize any interference effect). This general result being established, the question of the nature of morphemic units represented in long-term memory and their precise role within the lexicon remains unanswered.

Two possible hypotheses of representation have been proposed: either morphemic units stand as access units to word representations, or they organize word representations in morphological families. According to the first hypothesis, morphemic units correspond to concrete pieces of words (i.e., stems and affixes, even letter patterns resembling to morphemes but not functioning as such). Complex words are therefore processed by a decomposition mechanism that stripes off the affix in order to isolate the stem. The morphemic nature of the remaining letters is then checked out by the system in order to eliminate any procedural error. Access to word representations (i.e., word forms coded in the orthographic lexicon) can then operate via the pre-activation of the constituent morphemes. This mechanism explains why two morphologically related words prime each other, and this view is broadly shared by numerous authors interpreting their data within a sublexical approach (initially developed by Taft in 1994) that integrates morphemic representations as access units.

According to the second hypothesis, morphemic units are stored at an upper level of processing, at the interface of word and semantic representations. These intermediate units organize the lexicon in morphological families. Subsequently, each time a complex word is encountered, its recognition triggers the activation of all the word forms that can match with it. A competition is then engaged between the pre-activated forms until the right lexical unit reaches its recognition threshold (determined by its surface frequency). However, during the competition phase, competitors send positive activation to their respective base morpheme that in turn, send back positive activation to them. Two morphologically related words prime each other thanks to this

mechanism of co-activation[1]. Following this supralexical theory (Giraudo & Grainger, 2001), morphologically complex words are not "decomposed" in the proper sense (*viz*. following the same procedure described by the sublexical theory) but can trigger the activation of their constituent morphemes.

Regardless of the differences between sublexical and supralexical approaches of morphological processing, they both agree with the idea that separate morphemic units are responsible for priming effects. It is the precise location of these specific units within the architecture of the mental lexicon that specifies their role in word processing (access units vs. organizing units) as well as their own nature.

According to the sublexical view, morphemic units play the role of access units since they correspond to concrete letter clusters (i.e., bound stems, free stems and affixes) that constitute words, insensitive to any grammatical or semantic characteristic of words (i.e., transparency vs. opacity) or to their lexical environment (in terms of orthographic neighbourhood or family size). On the other hand, the supralexical view locates these units above the word forms and before the semantic units. These intermediate units are then supposed to be more abstract than those contained in the words because they have to tolerate form variations induced by the processes of derivation and inflexion (i.e., allomorphy, suppletion, phonological/morphological truncation, haplology). As a consequence, a morphemic unit does not need to exist in the real world in order to be coded in long-term memory but its existence/emergence depends on the interactions between the word and the semantic levels. Such a position also implies that all morphemes of a given language are not necessarily represented within the mental lexicon.

Recent studies explored these issues in order to test the decomposition hypothesis. Using the masked priming paradigm, it was shown that pseudo-derived words (e.g., *corner*) as well as pseudo-derived nonword (e.g., *corning*) primes composed of two existing morphemes were able to produce significant priming effects on the recognition times of their base (e.g., *corn*). Moreover it appears that the quality as well as the magnitude of these priming effects is comparable to the priming effects produced by genuine derived words (e.g., *banker-bank*). Finally, in order to separate pure morphological effects from form overlap effects, these studies used systematically orthographic control primes (i.e., morphologically simple forms whose only one part mimics a stem morpheme; such as *brothel* in which *-el* never functions as a suffix in English). Globally, the results demonstrated that the priming effects induced by derived as well as pseudo-derived primes differed significantly from these controls, suggesting that these effects resulted exclusively from the surface morphological structure of the primes. For instance, Longtin, Segui and Hallé (2003) demonstrated using French materials that a pseudo-derived word such as *baguette* (*'stick'*) (composed with the fragments *bagu-* and *–ette* that correspond to existing morphemes) facilitated the recognition of the target *bague* (*ring*) while at the same time a comparable orthographic control such as the word *abricot* (*'apricot'* in which only the fragment *abri* can be assimilated to an existing morpheme) did not facilitate the recognition of its pseudo-base *abri* (*'refuge'*). These results were replicated by Rastle, Davis and New (2004) who found a strong *corner-corn* priming effect using English materials but no priming effects with the *brothel-broth* prime-target pairs. Then, Longtin and Meunier (2005) explored the "pseudoderivation effect" using pseudo-words in order to test the resistance of early morphological decomposition face to the manipulation of the lexicality of the primes. In their masked priming study, morphologically complex pseudo-words (non existing possible words created with two existing morphemes, for instance, the base *sport-* + the suffix *-ation* produce *sportation*) were used as primes. The data revealed that pseudo-derived pseudowords (i.e., *sportation*) facilitated the recognition latencies of their base (e.g., *sport*) and did not differ from the facilitation effects obtained using transparent primes (e.g., *sportif* which is legal and

[1] It's interesting to note that under certain circumstances, a morphologically related word would not be able to facilitate or could even slow down the recognition latency of the target: when lateral inhibition is equal or stronger than excitation sent by the morphemic unit on its family members. This would be the case with prime words characterized by a high number of orthographic neighbors, a small morphological family and a weak root (in terms of its surface frequency).

semantically transparent derivation of the base *sport*). More recently, McCormick, Rastle and Davis (2008) manipulated a new category of derived stimuli, those that cannot be segmented perfectly into their morphemic components (e.g., *dropper-drop* in which there's a duplicated consonant) in order to test the flexibility of the morpho-orthographic segmentation process described by morpheme-based models. Their results demonstrated the robustness of this segmentation process in the case of various orthographic alterations in semantically related (e.g., *adorable-adore*) as well as in unrelated prime-target pairs (e.g., *fetish-fete*).

Taken together these data strongly support the robustness of a morphological decomposition effect across languages, stimuli and sensorial modalities. A complete review of the literature related to this question was made by Rastle and Davis (2008) and summarized perfectly this result in claiming: "morphological decomposition is a process that is applied to *all* morphologically structured stimuli, irrespective of their lexical, semantic, or syntactic characteristics" (p. 949). This conclusion seemed to deliver the *coup de grace* to any approach (the supralexical model in particular) that would postulate intermediate lexematic units situated above word units.

Nevertheless, the very recent study conducted by Crepaldi, Rastle, Coltheart, & Nickels (2010) opened a breach in this wall of certainty. A series of masked priming experiments were carried out on English irregularly inflected forms (viz. allomorphs). Interestingly enough and in total contradiction to their starting hypothesis, the authors found that allomorphs (e.g., *fell*) whose construction disables decomposition, primed their verbal base (e.g., *fall*) more than did orthographically-matched (e.g., *fill*) and unrelated control words (e.g., *hope*). This result had been already found by Pastizzo & Feldman (2002), and discussed enough by morphologists, but it had not been attributed the right importance by the tenants of the sublexical approach because of minor pitfalls in the control conditions (that do not have any incidence on the results, as the comparison between the data of Pastizzo & Feldman and those of Crepaldi et al. demonstrates). The authors conceded the "existence of a second higher-level source of masked morphological priming" and proposed a lemma-level composed of inflected words acting "at an interface between the orthographic lexicon and the semantic system".

However, this double source of morphological priming leads us to differentiate the nature of the coded morphemes. If we turn back to the locus issue that, according to us, determines the content of the units reflecting (and explaining) morphological effects, it is important to highlight that more than 90% of the experimental studies manipulated suffixed words or pseudowords. However, prefixed and suffixed words show many differences in terms of (1) their position relative to the stem, (2) relative number of suffixes and prefixes, (3) their grammatical properties (Montermini, 2008; Stump, 2001). To our knowledge, only few experimental studies were consecrated to affix processing representation. Two experimental papers (Colé, Beauvillain, & Segui (1989); Meunier & Segui, 1999) presented data obtained through naming and lexical decision tasks suggesting that the processing of prefixes and suffixes might differ. But masked priming studies conducted on one hand in French (Giraudo & Grainger, 2003) and on the other hand in Spanish (Duñabeitia, Perea, & Carreiras, 2008) presented contradictory results. While Giraudo and Grainger found that only prefixed primes – but not suffixed ones - produced morphological facilitation on target recognition latencies (e.g., *prénom-préface*), Duñabeitia and coll. get suffix priming (using a different experimental design[2]). We have nevertheless to observe that when it comes to the test of the decomposition hypothesis ALL the studies were conducted using suffixed words.

[2] Giraudo & Grainger examined affix priming effects using two types of affixed words, prefixed and suffixed, and three priming conditions: (1) an affix condition (e.g., *prénom-préface*), a pseudo-affix condition (e.g., *préfet-préface*) and a unrelated baseline condition (e.g., *guitare-préface*). Only prefixed prime-target pairs produced facilitation that differed significantly relative to pseudo-affixed and unrelated primes, suggesting a genuine morphological effect. Duñabeitia and coll. compared suffix priming using two kinds of words: polymorphemic (e.g., *igualdad*) vs monomorphemic (e.g., *certamen*) and two priming conditions: related vs unrelated. While they found facilitation effects for polymorphemic words (e.g., *brevedad-igualdad* faster than *plumaje-igualdad*), these effects did not occur for monomorphemic words (e.g., *volumen-certamen* equivalent to *topacio-certamen*).

The present paper attempts to bring new elements relatively to two related but unanswered questions: are pseudo-derivation effects observed using prefixed primes and are prefixes represented in long term memory? Four masked priming experiments were conducted using French materials. In these experiments we selected either morphologically simple targets (e.g., *nom*) or morphologically complex targets (e.g., *surnom*) and we systematically manipulated three priming conditions: a morphologically related condition (M+), an orthographic condition (O+) and an unrelated condition (M-O-). While the M+ condition used prefixed word primes (e.g., *prénom*), the O+ condition used either pseudo-prefixed nonwords in Exp.1 and 2 (e.g., *dénom*), nonwords containing a related stem in Exp.3 (e.g., *danom*) or nonwords containing a prefix but no stem in Exp. 4 (e.g., *prénam*). Globally, the results seem to indicate that prefixed and pseudoprefixed primes produce equivalent facilitation effects on both simple and complex target recognition. Moreover, they highlight the strong dependency of these effects on the presence of two existing morphemes within the prime, since nonword primes containing a related stem were not able to produce priming (and, in fact, this condition did not differ from the unrelated baseline condition). Taken together, these results are in line with those found using suffixed and pseudo-suffixed words (Longtin and coll., 2003, 2005; Rastle and coll., 2004; 2008). In order to integrate pseudo-derivation effects as well as affix effects (restricted to prefixes) within the same lexical architecture, while keeping in mind that for certain morphologically complex words and in particular those that cannot be decomposed into morphemes, there is a need to represent morphology at a higher level of processing (as suggested by Crepaldi and coll. 2010), we present a new architecture composed of four levels (Figure 1):

(1) Submorphemic units that only correspond to surface morphemes.
(2) Word units defining a separate level of processing which constitutes the orthographic lexicon.
(3) Base lexemes connected at the upper level with their family members.
(4) Concept units connected to both word and base lexeme units.

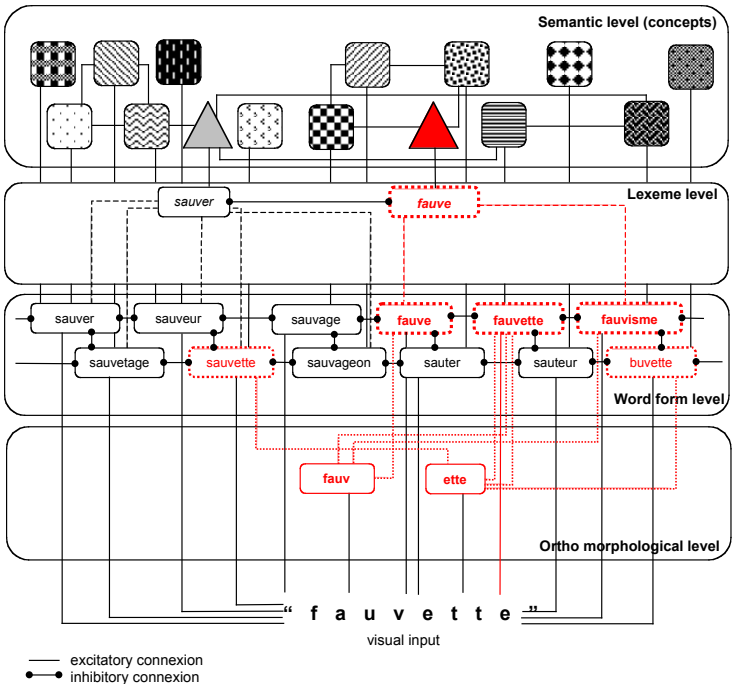

Figure 1: Hybrid model of morphological processing. The visual input *fauvette* triggers at the same time the activation of the ortho-morphological level (i.e., the morphemes *fauv-* and *–ette* are positively activated and send excitation to the related word forms ***fauve***, ***fauvette***, ***fauv***isme, sauv***ette***, buv***ette***, *etc.*) and the word level (i.e., orthographic neighbours compete with each other via inhibitory connexions). Then, the lexeme *fauve* is activated but this activation is indirect because it's not triggered by the input *fauvette* but by its related word forms *fauve* and *fauvisme*.

Morphological information contained within words is then coded according to two dimensions, their surface form and their internal structure. The first level captures the perceptive regularity and the saliency of morphemes within the language. It contains the stems and affixes that can be extracted from words according to a simple segmentation process. At this level of coding, morphologically complex words, pseudo-derived words and nonwords whose surface structure can be divided into (at least two) distinct morphemes are similarly processed. As a consequence, this level cannot be considered as a morphological level in a proper sense, but rather as an ortho-morphological level. Contrarily, the second level deals with the internal structure of words, how they're formed according to morphological rules. This level contains lexemes (nouns, verbs and adjectives) abstract enough to tolerate orthographic and phonological variations produced by derivation and inflection. Lexeme representations are connected to morphologically related word representations and the connections are determined by the degree of semantic transparency between the word forms and the lexeme. Morphologically complex words that are semantically transparent, are connected with both their constituent lexemes and morphemes. However, words whose the morphological structure is semantically opaque (e.g., *fauvette 'warbler'* that is not related anymore to its free-standing stem *fauve 'tawny'*) or illusory (e.g., *baguette 'stick'* in which *bagu-* is not a stem and has nothing to do with *bague 'ring'*) are not connected with their lexeme. Both types of items are nevertheless connected with their constituent morphemes situated at the ortho-morphological level.

References

Colé, P., Beauvillain, C., & Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and Language, 28*, 1-13.

Crepaldi, D., Rastle, K, Coltheart, M., & Nickels, L. (2010). Early morpho-orthographic segmentation and masked irregular priming. *Journal of Memory and Language, to appear*.

Diependaele, K., Sandra, D. & Grainger, J. (2005). Masked cross-modal morphological priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes, 20*, 75-114

Duñabeitia, J.A., Laka, I., Perea, M., & Carreiras, M. (2009). Is Milkman a superhero like Batman? Constituent morphological priming in compound words. *European Journal of Cognitive Psychology, 21(4),* 615-640.

Duñabeitia, J.A., Perea, M., & Carreiras, M. (2008). Does darkness lead to happiness? Masked suffix priming effects. *Language and Cognitive Processes, 23*, 1002-1020.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 10*, 680-698.

Giraudo, H., & Grainger, J. (2001). Priming complex words: Evidence for supralexical representation of morphology. *Psychonomic Bulletin and Review, 8*, 96-101.

Giraudo, H., & Grainger, J. (2003). On the role of derivational affixes in recognising complex words: Evidence from masked priming. In H. Baayen (Ed.), *Aspects of morphological processing*. New York: Mouton de Gruyter.

Longtin, C. M., Segui, J., & Hallé, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes, 18(3)*, 313-334.

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

Meunier, F., & Longtin, C. M. (2007). Morphological decomposition and semantic integration in word processing. *Journal and Memory and Language, 56*, 457-471.

Meunier, F., & Segui, J. (1999). Frequency effects in auditory word recognition: The case of suffixed words. *Journal of Memory and Language, 41*, 327-344.

McCormick, SF., Rastle, K., & Davis, MH (2008), "Is there a 'fete' in 'fetish'?: Effects of orthographic opacity on morpho-orthographic segmentation in visual word recognition. *Journal of Memory and Language, 58*, 307-326.

Montermini, F. (2008). *Il lato sinistro della morfologia. La prefissazione in italiano e nelle lingue del mondo*. Milano: FrancoAngeli.

Pastizzo, M. J., & Feldman, L. B. (2002). Discrepancies between orthographic and unrelated baselines in masked priming undermine a decompositional account of morphological facilitation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*, 244-249.

Rastle, K, Davis, M.H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes, 23*, 942-971.
Rastle K, Davis M H, New B (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin and Review, 11,* 1090-1098.
Stump, G. (2001). Affix position, in Haspelmath M., König H., Oesterreicher W., Raible W. (a c. di) 2001, *Language Typology and Language Universals*, Berlin – New York, de Gruyter.708-714.
Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes, 9*, 271-294.