

# Morphological decomposition involving non-productive morphemes: ERP evidence

Richard McKinnon, Mark Allen<sup>1</sup> and Lee Osterhout<sup>1,CA</sup>

Departments of Linguistics; <sup>1</sup>Psychology, University of Washington, Seattle, WA 98195, USA

<sup>CA</sup>Corresponding Author: [losterho@u.washington.edu](mailto:losterho@u.washington.edu)

Received 14 January 2003; accepted 14 March 2003

DOI: 10.1097/01.wnr.0000070192.28954.7e

It is generally believed that readers decompose a complex word into its constituent morphemes only when those morphemes participate productively in word formation. Here we recorded event-related potentials (ERPs) to words (e.g. muffler, receive), non-words containing no morphemes (e.g. flermuf), and non-words containing a prefix and a non-productive bound stem (e.g. in-ceive). Prior work has shown that pronounceable non-words elicit larger-amplitude N400 components than words. If readers treat

non-words containing non-productive morphemes as unanalyzed wholes, then these non-words should elicit larger N400s than matched words. We report here, however, that bound-stem non-words elicit a brain response highly similar to that elicited by real words. This finding suggests that morphological decomposition and representation extend to non-productive morphemes. *NeuroReport* 14:883–886 © 2003 Lippincott Williams & Wilkins.

**Key words:** Event-related potentials; Language processing; Morphological decomposition; Morphological processing; Reading

## INTRODUCTION

Many of the words we produce and recognize are composed of more than one morpheme. One question is whether morphologically complex words are represented and accessed as unanalyzed whole-word forms or as constituent morphemes. For example, a derived word like lateness could be represented as the unstructured orthographic string lateness and associated directly with a corresponding stored meaning, or could be computed from the meaning of the stem late, the lexical-syntactic properties of the affix -ness, and the rules for combining them.

One property thought to restrict decomposition is morphological productivity [1,2], which refers to the extent that the morpheme can be used to form new words. For example, the productive English suffix -ness can be used extensively to derive novel nouns from adjectives (e.g. lateness, blackness), whereas the less-productive morpheme -ity (e.g. brevity, specificity) is much more limited in this regard. At the limits of non-productivity in English are word formations involving bound stems (i.e., stems that must be attached to another morpheme, such as -ceive, -mit, -cede). The meaning of a word like conceive is not obviously predictable from the combination of its prefix and stem; nor does a bound stem like -ceive appear to have any particular meaning. Moreover, the spoken and written forms of many bound stems change when they are suffixed (e.g. conceive-conception). For these reasons, it seems likely that complex words with bound stems and other non-productive morphemes are stored and processed as unanalyzed wholes, rather than as decomposed morphemes.

Evidence that readers do decompose letter strings containing non-productive morphemes, and independently represent these morphemes in their mental lexicons, would strikingly demonstrate that the processor uses a decompositional approach, even when the functional benefit for doing so is limited. Prior work using behavioral methods has been inconclusive in this regard [3–5]. Here, we examined this question by recording ERPs while participants read words and word-like non-words. Many studies have shown that the N400 component of the ERP [6–8] is sensitive to the lexical status (word *vs* non-word) and frequency of a letter string [9–11]. Pronounceable non-words elicit larger-amplitude N400s than real words, and less-frequent words elicit larger N400s than more-frequent words.

In the present study, we compared N400s elicited by words (e.g. retain, intrude) and non-words (hereafter bound-stem non-words; e.g. intain, retrude) that are composed of the same set of prefixes and bound stems. The logic is as follows: if readers treat the non-words as unanalyzed wholes, then the non-words should elicit larger-amplitude N400s than the real words. This follows from the fact that whereas the words are relatively frequent in English, the bound-stem non-words (taken as a whole unit) are non-words with zero frequency. Conversely, if readers decompose the words and non-words into their constituent morphemes, and if these morphemes are represented in the mental lexicon, then the words and non-words might elicit similar amplitude N400s; both the words and non-words are made up of morphemes with the same frequency. The

notion that certain types of morphologically complex non-words might elicit word-like brain responses gains support from two recent ERP studies [12,13]. Gross and colleagues, for example, found that when verb stems in Italian were combined with affixes from the wrong verb class, the resulting non-words elicited brain potentials that were indistinguishable from those elicited by completely well-formed inflected verbs [13].

In Experiment 1, we report that bound-stem words and non-words elicit N400 components that do not differ in amplitude. By contrast, word-like non-words that are not made up of morphemes elicit a much larger N400 than do matched words. Experiments 2 and 3 demonstrate that these results are not due to orthographic similarity or prefix repetition, respectively.

## MATERIALS AND METHODS

**Subjects:** The subjects were 36 healthy, right-handed university students (25 women) with normal or corrected-to-normal vision, aged 18–29 years. Twelve subjects participated in each experiment.

**Stimuli:** For Experiment 1, a set of 60 words containing bound stems was chosen (e.g. sub-mit, re-ceive, dis-cern) from which 60 bound-stem non-words were formed by recombining stems and prefixes into non-word forms (e.g. pro-mit, in-cern, ex-ceive). The bound-stem non-words conformed to category-specific requirements of affix-stem combination, i.e. each stem was a legal host for the prefix with which it was combined. In addition, 60 morphologically complex words (e.g. bookmark, muffler) were selected to serve as control words. Control non-words (e.g. moobkark, flermuf) were formed by permuting the letters of each control real-word into a pronounceable non-word with no apparent morphemes. Two experimental lists were constructed such that each contained 60 non-words (30 bound-stem and 30 control) and 60 words (30 bound-stem and 30 control). The lists were counterbalanced so that the stems that occurred in real words in one list occurred in non-words in the other list. Prefixes were balanced across lists such that equal numbers of each prefix appeared in each list. Items in all conditions had a mean length of seven letters. In order to partially control for the effects of orthographic regularity, words and non-words were matched for bigram frequency. A mean bigram frequency was calculated for each form (based on the frequency of each bigram in the word or non-word, divided by the total number of bigrams, including initial and final segments). Frequency norms were generated by counting bigrams from ~1 million English words. Mean bigram frequencies did not differ between words and non-words (bound stem condition,  $F < 1$ ; control condition,  $F < 1$ ).

In Experiment 1, bound-stem words and non-words shared on average an uninterrupted sequence of 4.33 letters; this was not true for the control condition. For Experiment 2, we made letter substitutions to the bound-stem words used in Experiment 1, preserving a sequence of at least four consecutive letters (mean 4.48) in common with the words (but not preserving morphological content). One letter (the initial letter) or two letters (either the initial letter and the final letter, or the initial letter and the penultimate letter)

were substituted with a different letter, always preserving the CV structure of the word. For example, the morphologically complex word *insult* was transformed into the morphologically empty non-word *ensuld* (hereafter orthographic non-words). The stimuli in each category were counterbalanced across two lists such that 30 exemplars of each category were on a list, with matched-pair members on separate lists.

Experiment 3 was designed to examine another potential confound. In Experiment 1, subjects saw prefixes in the bound-stem words and non-words more than once. Word repetition reduces N400 amplitude [14]. Experiment 3 was designed to determine whether affix priming is sufficient to produce word-like responses to non-words. One hundred twenty prefixed and suffixed words (e.g. *ex-port*, *boy-ish*) were chosen to serve as real words. From this set, 120 affixed non-words were formed by permuting the letters in each of the stems (e.g. *ex-torp*, *yob-ish*). Thus, non-words contained a repeated affix and a nonce stem, and real words contained a repeated affix and an English stem. Two counterbalanced lists were formed that included 60 affixed real words and 60 affixed non-words (30 prefixed and 30 suffixed, in each condition).

**Procedure:** Each trial consisted of the following events: a fixation cross appeared for 500 ms, after which a word appeared on the center of the screen for 700 ms. A 1450 ms blank screen interval followed each word, after which a prompt appeared asking participants to indicate if the preceding item was an English word. Participants responded by pressing one of two buttons, which were counterbalanced (left and right) across participants.

**Data acquisition and analysis:** Continuous EEG was recorded from 13 scalp sites using tin electrodes attached to an elastic cap. Electrode placement included International 10–20 system locations [15] over homologous positions over the left and right occipital (O1, O2) and frontal (F7, F8) regions and from frontal (Fz), central (Cz), and parietal (Pz) midline locations. In addition, several non-standard sites over posited language centers were used, including Wernicke's area and its right hemisphere homologue (WL, WR: 30% of the interaural distance lateral to a point 13% of the nasion-inion distance posterior to Cz), posterior temporal (PTL, PTR: 33% of the interaural distance lateral to Cz), and anterior temporal (ATL, ATR: one-half the distance between F7/F8 and T3/T4). Vertical eye movements and blinks were monitored by means of two electrodes, one placed beneath the left eye and one placed to the right of the right eye. The above 15 channels were referenced to an electrode placed over the left mastoid bone and were amplified with a bandpass of 0.01–100 Hz (3 dB cutoff) by a Grass Model 12 amplifier system. Activity over the right mastoid was recorded on a sixteenth channel to determine whether there were any effects of the experimental variables on the mastoid recordings. No such effects were observed.

Continuous analog-to-digital conversion of the EEG and stimulus trigger codes was performed by a Data Translation 2801-A board and a computer at a sampling frequency of 200 Hz. Each 1280 ms epoch was comprised of the 100 ms of activity immediately preceding and 1180 ms of activity

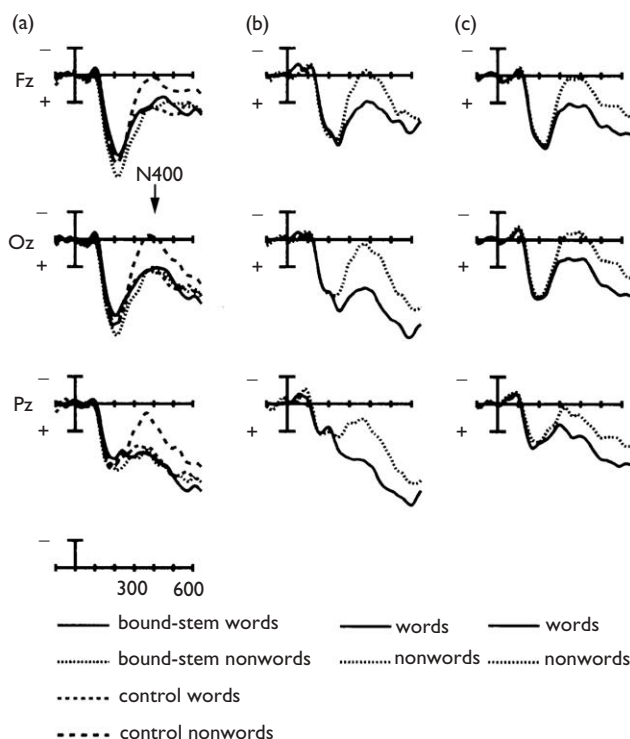
immediately subsequent to the onset of each letter string. Trials contaminated by excessive eye movement or amplifier blocking were removed prior to averaging (< 20% of trials in all experiments and conditions). ERPs were quantified as the mean voltage within three time windows (50–150, 150–300, and 300–500 ms) following onset of each word relative to a baseline (100 ms of activity immediately preceding letter string onset). These windows were chosen because they correspond to the latency ranges of the N1, P2, and N400 ERP components. Data acquired at midline and lateral sites were treated separately to allow for quantitative analysis of hemispheric differences. For Experiment 1, a three-way ANOVA was performed on data acquired over midline sites, with repeated measures on two levels of morphological structure (bound-stem *vs* control), two levels of lexical status (word *vs* non-word), and three levels of electrode site. A four-way ANOVA was performed on data acquired over lateral sites, with repeated measures on bound-stem structure, lexical status, five levels of electrode site, and two levels of hemisphere. In addition, single-df planned comparisons were used to contrast ERPs to words and non-words in the bound-stem and control conditions. These comparisons were tested using a modified Bonferroni procedure [16]. For Experiments 2 and 3, two-way ANOVAs were performed on midline data, with repeated measures on lexical status and electrode site. A three-way ANOVA was performed on lateral data, with repeated measures on lexical status, electrode site, and hemisphere.

## RESULTS

**Experiment 1:** Bound-stem words and non-words were judged to be words on 85% and 17% of the trials, respectively. Control words and non-words were judged to be words on 87% and 1% of the trials, respectively. The difference in errors between bound-stem and control non-words was significant ( $F(1,11) = 6.99, p < 0.03$ ).

All stimuli elicited a negative component (N1) peaking at ~100 ms and a positive component (P2) peaking at ~200 ms (Fig. 1a). These components were followed by a negative wave (N400) that was largest over posterior midline sites. ERPs to the bound-stem non-words were slightly more positive at some electrode sites within the 150–300 ms window (bound-stem/control  $\times$  real-word/non-word  $\times$  electrode site,  $F(2,22) = 4.27, p < 0.05$ ), reflecting a slightly larger P2 in this condition. Within the critical N400 window, the control non-words ( $F(1,11) = 39.33, p < 0.001$ ), but not the bound-stem non-words ( $F(1,11) = 0.75, p > 0.4$ ), elicited a larger-amplitude N400 than the two types of words (bound stem/control  $\times$  word/non-word: midline,  $F(1,11) = 16.66, p < 0.001$ ; lateral,  $F(1,11) = 13.84, p < 0.01$ ).

**Experiment 2:** Participants judged the bound-stem words and orthographic non-words to be words on 88% and 3% of the trials, respectively. Orthographic non-words elicited larger-amplitude N400s than did the words (Fig. 1b; 300–500 ms: midline,  $F(1,11) = 17.27, p = 0.01$ ; lateral,  $F(1,11) = 9.68, p < 0.01$ ). No reliable differences between conditions were found between 50 and 300 ms (all  $F < 2$ ).



**Fig. 1.** Grand-average event-related potentials (averaged over subjects and items), recorded at frontal (Fz), parietal (Pz), and occipital (Oz) midline sites. Word onset is at 0 ms. Each mark on the horizontal axis represents 100 ms. Negative voltage is plotted up. The vertical calibration bar represents 3  $\mu$ V. (a) Experiment 1: ERPs to bound-stem real words (e.g. insult), bound-stem non-words (e.g. exceive), control words (e.g. muffler), and control non-words (e.g. flermuf). (b) Experiment 2: ERPs to words (e.g. insult) and orthographic non-words (e.g. ensuld). (c) Experiment 3: ERPs to affixed words (e.g. drinkable) and non-words (e.g. krindable).

**Experiment 3:** Participants judged 91% and 3% of the words and non-words to be words, respectively. Non-words elicited larger-amplitude N400s than did words (Fig. 1c; 300–500 ms: midline,  $F(1,11) = 14.73, p < 0.01$ ; lateral,  $F(1,11) = 10.88, p < 0.01$ ). The prefix/suffix factor did not interact with any other factor ( $F < 1$  in all analyses). No reliable differences between conditions were observed between 50 and 300 ms (all  $F < 2$ ).

## DISCUSSION

Current theories of word processing claim that morphological decomposition occurs only for complex words that contain productive morphemes, or for words whose meanings are transparently a sum of the meanings of their constituent morphemes [3,17–21]. We report here, however, that bound-stem non-words (e.g. exceive) elicit a brain response remarkably similar to that elicited by bound-stem real words (e.g. receive), especially throughout the part of the ERP waveform known to be sensitive to lexicality (Experiment 1). This was true even though these non-words are composed of morphemes (e.g. ex- and -ceive) that are not used productively in word formation and that have little, if any, meaning. By contrast, word-like non-words

that are not fully made up of morphemes elicit much larger N400s than words. Importantly, our results cannot be accounted for in terms of orthographic similarity between the words and non-words (Experiment 2) or affix priming (Experiment 3). According to current word processing theories, morphological decomposition should be least likely under the conditions of our experiment. Therefore, it is reasonable to conclude on the basis of our results that morphological decomposition is quite ubiquitous during reading, and might occur for the majority of complex words.

Our results are consistent with a model in which a morphologically complex letter string is parsed into its morpheme components, even when these morphemes are non-productive and semantically impoverished. Presumably, morphemes are mapped onto stored entries in the input lexicon, which in turn engage the semantic and syntactic features associated with them. A complex word can be comprehended on-line by using rules of lexical morphology to combine the individual bits of information associated with each morpheme into a coherent linguistic interpretation [3,17,20]. The ERP evidence presented here suggests that a non-productive stem like -mit is stored in the lexicon along with the lexical-syntactic (and any semantic) information that it encodes. During reading, the stem -mit is parsed and its lexical content is retrieved. However, the full meaning of words that contain bound stems cannot be derived straightforwardly by applying rules of morphological/semantic combination. Thus, even if the lexicon includes representations for semantically impoverished morphemes, it would also need to include whole-word representations for each familiar combination of these morphemes as well. This type of redundant system has considerable empirical and theoretical support [4,5,20,22,23].

Our results also bear on hypotheses about the cognitive substrates of the N400 component itself. Prior work has indicated that, for letter strings presented in isolation, N400 amplitude is a function of the lexical status of the string and word frequency. However, in the present study, N400 amplitude was entirely a function of the morphological content of the letter string (see also [12,13]). Indeed, the bound-stem non-words presented here had zero frequency; yet, these letter strings elicited a brain response that was nearly identical to that elicited by frequently occurring words. This was true even though the participants were required to make lexical decisions. This introduces a compelling paradox: At the same time that participants were categorizing the stimuli as word or non-word (with

reasonable success), their brains were seemingly categorizing these stimuli into the quite different categories of string can be fully parsed into morpheme components and string cannot be fully parsed. Apparently, N400 amplitude is sensitive to the presence or absence of morphemes in the string, but not to whether the morphemes combine to form a word.

## CONCLUSION

Our findings suggest that morphological decomposition and representation extend to non-productive and semantically impoverished morphemes. This conclusion follows from the observation that words and non-words composed of a prefix and a bound stem elicit similar brain responses. By contrast, non-words that are not entirely made up of real morphemes, but that are word-like in other ways, elicit much larger N400s than do matched words.

## REFERENCES

1. Baayen H. *Lang Cogn Proc* **6**, 447–469 (1994).
2. Baayen H and Lieber R. *Linguistics* **29**, 801–843 (1991).
3. Marslen-Wilson W, Tyler L, Waksler R et al. *Psychol Rev* **101**, 3–33 (1994).
4. Wurm L. *J Mem Lang* **42**, 255–271 (2000).
5. Forster K and Azuma T. *Lang Cogn Proc* **15**, 539–562 (2000).
6. Kutas M and Hillyard S. *Science* **207**, 203–205 (1980).
7. Kutas M and Hillyard S. *Nature* **307**, 161–163 (1984).
8. Osterhout L, McLaughlin J and Bersick M. *Trends Cogn Sci* **1**, 203–209 (1997).
9. Bentin S, McCarthy G and Wood C. *Electroencephalogr Clin Neurophysiol* **60**, 343–355 (1985).
10. Holcomb P and Neville H. *Lang Cogn Proc* **4**, 281–312 (1990).
11. Chwilla D, Brown C and Hagoort P. *Psychophysiology* **32**, 274–285 (1995).
12. Penke M, Weyerts H, Gross M et al. *Brain Cogn Brain Res* **6**, 37–52 (1997).
13. Gross M, Say T, Kleingens M et al. *Neurosci Lett* **241**, 83–86 (1998).
14. Rugg M. *Q J Exp Psychol (A)* **39**, 123–148 (1987).
15. Jasper H. *Electroencephalogr Clin Neurophysiol* **10**, 371–375 (1958).
16. Keppel G. *Design and Analysis: A Researcher's Handbook*. Englewood Cliffs, NJ: Prentice-Hall; 1982.
17. Allen M and Badecker W. Morphology: the internal structure of words. In: Rapp B (ed.). *The Handbook of Cognitive Neuropsychology: What Deficits Reveal About the Human Mind*. New York: Psychology Press; 2000, p. 211.
18. Bertram R, Schreuder R and Baayen H. *J Exp Psychol Learn Mem Cogn* **26**, 489–511 (2000).
19. Laudanna A and Burani C. *Linguistics* **23**, 775–792 (1985).
20. Schreuder R and Baayen H. *Modeling morphological processing*. In: Feldman L (ed.). *Morphological Aspects of Language Processing*. Hillsdale, NJ: Lawrence Erlbaum; 1995, p. 131.
21. Wurm L and Ross S. *J Mem Lang* **45**, 39–57 (2001).
22. Caramazza A, Laudanna A and Romani C. *Cognition* **28**, 297–332 (1988).
23. Taft M. *Lang Cogn Proc* **9**, 271–294 (1994).

Acknowledgements: This research was supported by Grant R01 DC01947 from the National Institute on Deafness and Communication Disorders, National Institutes of Health.