



HAL
open science

Morphological processing without semantics: An ERP study with spoken words

Elisabeth Beyersmann, Deirdre Bolger, Chotiga Pattamadilok, Boris New, Jonathan Grainger, Johannes Ziegler

► **To cite this version:**

Elisabeth Beyersmann, Deirdre Bolger, Chotiga Pattamadilok, Boris New, Jonathan Grainger, et al.. Morphological processing without semantics: An ERP study with spoken words. *Cortex*, 2019, 10.1016/j.cortex.2019.02.008 . hal-02047828

HAL Id: hal-02047828

<https://hal.science/hal-02047828v1>

Submitted on 25 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

Morphological processing without semantics: An ERP study with spoken words

Elisabeth Beyersmann, Deirdre Bolger, Chotiga Pattamadilok, Boris New, Jonathan Grainger, Johannes C. Ziegler



PII: S0010-9452(19)30060-7

DOI: <https://doi.org/10.1016/j.cortex.2019.02.008>

Reference: CORTEX 2565

To appear in: *Cortex*

Received Date: 26 November 2017

Revised Date: 23 April 2018

Accepted Date: 8 February 2019

Please cite this article as: Beyersmann E, Bolger D, Pattamadilok C, New B, Grainger J, Ziegler JC, Morphological processing without semantics: An ERP study with spoken words, *CORTEX*, <https://doi.org/10.1016/j.cortex.2019.02.008>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Morphological processing without semantics: An ERP study with spoken words

Elisabeth Beyersmann¹, Deirdre Bolger², Chotiga Pattamadilok³, Boris New⁴, Jonathan Grainger⁵ & Johannes C. Ziegler⁵

¹ *Department of Cognitive Science and ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney*

² *Aix-Marseille Univ, ILCB, Aix-en-Provence, France*

³ *Aix-Marseille Univ, CNRS, LPL, Aix-en-Provence, France*

⁴ *Univ. Savoie Mont Blanc, CNRS, LPNC, Chambéry, France*

⁵ *Aix-Marseille Univ, CNRS, LPC, Marseille, France*

Address for correspondence:

Elisabeth Beyersmann,

Department of Cognitive Science

Australian Hearing Hub

16 University Avenue

Macquarie University NSW 2109

Tel: +61 406 249 916

Email: lisi.beyersmann@gmail.com

Short title: spoken complex word recognition

Declarations of interest: none.

The time-course of morphological processing during spoken word recognition was investigated using event-related brain potentials (ERPs) in an auditory lexical decision task. We compared three different types of French words: truly suffixed (e.g., *pochette* ‘little pocket’ = *poche* ‘pocket’ + diminutive suffix *-ette*), pseudo-suffixed (e.g., *mouette* ‘seagull’ = *mou* ‘soft’ + pseudo-suffix *-ette*) and non-suffixed target words (e.g., *fortune* ‘fortune’ = *fort* ‘strong’ + non-suffix *-une*). Suffixed (e.g., *mouesse* = *mou* + suffix *-esse*) and non-suffixed nonwords (e.g., *mouipe* = *mou* + non-suffix *-ipe*) were also tested. The behavioural results showed that participants responded more slowly to non-suffixed words than to truly suffixed and pseudo-suffixed words, but there was no difference between these two suffixed conditions. Moreover, participants made more errors rejecting pseudo-suffixed nonwords than non-suffixed nonwords. **In the ERP analyses, T0 was shifted to the end of the embedded stem or pseudo-stem.** The ERP results revealed enhanced N400 amplitudes for non-suffixed words compared to truly suffixed and pseudo-suffixed words. Again, there was no difference between the truly and pseudo-suffixed conditions. In addition, we found an increased N400 amplitude for both pseudo-suffixed and non-suffixed nonwords than for words. The latency of the onset of this N400 effect varied between the three experimental conditions: the word-nonword difference occurred earliest in the truly suffixed condition, slightly later in the pseudo-suffixed condition and latest in the non-suffixed condition. Both behavioural and EEG data jointly suggest that spoken words with a genuine morphological structure and words with a pseudo-morphological structure are decomposed into morphemic sub-units. Moreover, the earlier appearance of the N400 effects in the truly suffixed condition indicates that morphological information is more readily available in words with a semantically transparent morphological structure.

ACCEPTED MANUSCRIPT
Keywords: morphological processing, spoken word recognition, auditory lexical decision,
event-related potentials (ERPs)

ACCEPTED MANUSCRIPT

This research, carried out within the Labex BLRI (ANR-11-LABX-0036) and the Institute of Convergence ILCB (ANR-16-CONV-0002), has benefited from support from the French government, managed by the French National Agency for Research (ANR) and the Excellence Initiative of Aix-Marseille University A*MIDEX (ANR-11-IDEX-0001-02). This work was further supported by an ANR grant to JG (ANR-15-FRAL-0003-01, MORPHEME) and a Macquarie University Research Fellowship (MQRF) to EB.

1. Introduction

Several decades of research have examined automatic parsing mechanisms involved in the visual recognition of morphologically complex words. However, comparatively little work has focussed on the automatic mechanisms involved in decomposing spoken complex words. The goal of the present study was to address this research gap by examining the early stages of morphological processing during spoken word recognition, using both behavioural data and event-related brain potentials (ERPs).

1.1. Morphological processing during visual word recognition

One of the key findings in the masked priming literature is that real suffixed (e.g. *hunter*) and pseudo-suffixed primes (e.g. *corner*) both facilitate the recognition of the embedded target word (e.g. *hunter-hunt*; *corner-corn*), relative to an orthographic control (e.g. *cashew-cash*) (for reviews, see Amenta & Crepaldi, 2012; Rastle & Davis, 2008). A common interpretation of these widely replicated findings is that both real suffixed words and pseudo-suffixed words are rapidly decomposed into their real or pseudo-morphemic subunits (*hunt + er*; *corn + er*) during the initial stages of visual word recognition (e.g., Beyersmann, Ziegler, et al., 2016; Longtin, Segui, & Hallé, 2003; Rastle, Davis, & New, 2004), whereas no such decomposition mechanism applies to words consisting of an embedded word and a non-morphemic ending (*cashew*, where *ew* is not an affix). The extent to which individual participants rely on morphological parsing during lexical processing can be modulated by individual differences in spelling and vocabulary (e.g., Andrews & Lo, 2013; Beyersmann, Casalis, Ziegler, & Grainger, 2015). However, what continues to be a matter of debate is how early or late morphological processing is influenced by semantics (e.g., Cavalli et al., 2016; Feldman, Milin, Cho, Moscoso Del Prado Martin, & O'Connor, 2015; Feldman, O'Connor, & Moscoso del Prado Martin, 2009). Some studies have reported equal magnitudes of priming for truly and pseudo-suffixed words, suggesting that the initial stages of morphological processing are

ACCEPTED MANUSCRIPT

semantically 'blind' (e.g., Beyersmann, Ziegler, et al., 2016; Longtin et al., 2003; Rastle & Davis, 2008; Rastle et al., 2004). Other studies have reported significantly increased magnitudes of priming in the truly suffixed compared to the pseudo-suffixed condition, suggesting that semantics modulate the very initial stages of visual word recognition (e.g., **Feldman et al., 2015; Feldman et al., 2009; Jared, Jouravlev, & Joanisse, 2017; Schmidtke, Matsuki, & Kuperman, 2017**). The latter view is consistent with parallel distributed processing theories (e.g., Gonnerman, Seidenberg, & Andersen, 2007; Plaut & Gonnerman, 2000) according to which the reading system picks up on statistical regularities, such as the consistency with which the letters of a morpheme are mapped onto semantic representations.

Whether or not the early stages of morphological processing are modulated by semantics may also depend on the nature and productivity of the morphological parsing system within specific languages. For instance, Semitic languages such as Hebrew and Arabic consist of divergent internal word structures, where stems and affixes are intertwined rather than concatenated linearly. Morphological priming effects in these languages are particularly robust compared to other Indo-European languages (e.g., Frost, Kugler, Deutsch, & Forster, 2005; Velan & Frost, 2011), suggesting that the degree of semantic influence on early morphemic parsing might vary between different languages.

Very useful in the context of this debate have been studies combining masked priming and high-temporal resolution recordings of event-related brain potentials (ERPs) to investigate the influences of orthography and semantics during morphological processing (e.g., Beyersmann, Iakimova, Ziegler, & Colé, 2014; Dominguez, De Vega, & Barber, 2004; Jared et al., 2017; Lavric, Rastle, & Clapp, 2011; Morris, Frank, Grainger, & Holcomb, 2007; Morris, Grainger, & Holcomb, 2008, 2013; Morris, Holcomb, & Grainger, 2008; Morris, Porter, Grainger, & Holcomb, 2011; Royle, Drury, Bourguignon, & Steinhauer, 2012). The majority of the reported ERP results have typically shown no difference between true

ACCEPTED MANUSCRIPT

morphological and pseudo-morphological priming in earlier time windows (for converging evidence from MEG, see Lehtonen, Monahan, & Poeppel, 2011; Lewis, Solomyak, & Marantz, 2011; Solomyak & Marantz, 2009; 2010; but see Jared et al., 2017). However, when visible primes were used (e.g., Lavric et al., 2011), morphological priming effects continued to be significant, whereas pseudo-morphological priming effects tended to be absent or reduced in the later time windows. Concurrent evidence is also reported by Lavric, Elchlepp, and Rastle (2012), who carried out an unprimed lexical decision study, contrasting morphological (*hunter*), pseudo-morphological (*corner*), and non-morphological (*cashew*) target words. Results revealed a difference at about 190 ms between the non-morphological condition and the two morphological conditions (which showed no differentiation), reflecting greater negative amplitudes in central and posterior electrodes for the non-morphological condition. In addition, a difference was observed between the pseudo-morphological condition and the other two conditions at about 250 ms (which showed no differentiation), reflecting greater positive amplitudes in central and parietal electrodes in the pseudo-morphological condition. In sum, these findings appear to suggest that complex printed words are initially decomposed based on a purely structural form of morphological analysis, whereas semantic constraints are only taken into account at later stages during morphological processing.

The sensitivity of the N400 component to morphological processes is also evidenced by studies contrasting inflected relative to simplex words. For instance, two Finnish unprimed lexical decision studies reported larger N400 amplitudes for inflected words compared to matched mono-morphemic words (Lehtonen et al., 2007; Leinonen et al., 2009; for related evidence from MEG, see Vartiainen et al., 2009).

ACCEPTED MANUSCRIPT
1.2. Morphological processing during spoken word recognition

Unlike the rich literature on the morphological processing of printed words, we know comparatively little about its role in spoken word recognition. A number of studies have found that listeners show sensitivity to morphological information during spoken word recognition (e.g., Balling & Baayen, 2008; Emmorey, 1989; Gwilliams, Monahan, & Samuel, 2015; Marslen-Wilson, Tyler, Waksler, & Older, 1994; Meunier & Segui, 1999; Wurm, 2000). For instance, Emmorey (1989) carried out an auditory priming task showing that pseudo-prefixed words with bound stems produce priming (e.g. *submit – permit*) whereas phonologically related words do not (e.g. *balloon – saloon*). However, these findings were not replicated with suffixed words. A cross-modal priming study by Marslen-Wilson et al. (1994) revealed priming effects with truly suffixed (e.g. *punishment-punish*), but not pseudo-suffixed forms (e.g. *casualty-casual*), suggesting that morphological decomposition only applies to words with a genuine morphological structure (see also Meunier & Longtin, 2007, for converging evidence from spoken pseudoword processing). Consistent with these findings, it has been shown that the way listeners process prefixed words is affected by the transparency of the semantic relationship between the word constituents (e.g., Wurm, 1997, 2000; Wurm & Ross, 2001). Hence, these findings from spoken word recognition seem to provide very little support in favour of semantically "blind" morphological segmentation, a pattern typically seen in visual word recognition.

Of course, auditory word studies differ from printed word studies in one important aspect. As discussed earlier, semantically blind processing of printed words is typically observed in masked priming paradigms that tap into the early, initial stages of word recognition. In contrast, auditorily presented stimuli are very difficult to mask. As a result, evidence for the morphological processing of spoken words mainly comes from lexical decision tasks using overtly presented primes or no primes, whose behavioural outcome may reflect both early and later, more strategic, stages of word recognition, and therefore provide

little insight into the automatic, online processes of word recognition. One solution to this problem is the recording of ERPs to more closely monitor how morphological processes unfold over time.

Auditorily presented inflected words typically elicit larger N400 amplitudes than mono-morphemic words (Leinonen et al., 2009; Leminen et al., 2011), thus replicating what is seen in the visual modality (e.g., Lehtonen et al., 2007; Leinonen et al., 2009; Vartiainen et al., 2009). Moreover, illegal derived nonwords elicit a larger wide-spread negativity than existing derived words and legal derived nonwords at ~300 ms after suffix onset (Leminen, Leminen, & Krause, 2010). This larger negativity (resembling the N400) in the illegal nonword condition was interpreted as reflecting a more difficult lexical-semantic integration of the morpheme constituents. A comparable pattern of ERP results was described by Leminen, Leminen, Kujala, and Shtyrov (2013), showing that derived words produced an enhanced mismatch negativity (MMN) 130-170 ms after suffix onset compared to their derived nonword counterparts (see also Ettinger, Linzen, & Marantz, 2014, for related evidence from MEG). However, although these studies clearly demonstrate that morphological information influences spoken word recognition, they do not address the important question of whether the processing of spoken complex words is largely determined by semantic transparency, as several behavioural studies would suggest (e.g., Marslen-Wilson et al., 1994; Wurm, 1997, 2000), or if form-based morphological segmentation mechanisms operate independently of semantics, as repeatedly reported in the visual domain (e.g., Amenta & Crepaldi, 2012; Beyersmann, Ziegler, et al., 2016; Dominguez et al., 2004; Rastle & Davis, 2008; Rastle et al., 2004; Royle et al., 2012; Stockall & Marantz, 2006). The goal of the present study was to address this question using an auditory lexical decision task and ERP recordings to uncover how auditory morphological processes unfold over time.

Our study was conducted in French, comparing three different item types: truly suffixed words (e.g., *pochette* ‘little pocket’, consisting of the stem *poche* ‘pocket’ + diminutive suffix *-ette*), pseudo-suffixed words (e.g., *mouette* ‘seagull’, consisting of the pseudo-stem *mou* ‘soft’, where the pseudo-suffix *-ette* is homophonous to the real diminutive suffix) and non-suffixed words (e.g., *fortune* ‘fortune’, containing the embedded word *fort* ‘strong’ followed by a non-suffix ending ‘une’). To our knowledge, this is the first ERP study contrasting the auditory processing of truly suffixed and pseudo-suffixed words.

If it is indeed the case that morphological decomposition in the auditory modality is largely determined by semantics, we would expect to see (i) a reduction of the N400 amplitude for truly suffixed words relative to the pseudo-suffixed and non-suffixed conditions, and (ii) a greater “word-advantage” for truly suffixed words (i.e. shorter response latencies, lower error rates, and reduced N400 amplitudes relative to their corresponding nonwords) than for pseudo-suffixed and non-suffixed words. Behaviorally, this pattern should be reflected in faster and more accurate responses in the truly suffixed condition compared to the pseudo-suffixed and non-suffixed conditions, because access to the affix and the stem typically “bolsters” word responses (Gwilliams et al., 2015).

If however morphological processing operates *independently* of semantics (as Lavric et al., 2012's findings suggest), we would expect the N400 amplitudes for truly and pseudo-suffixed words to pattern together (e.g., *pochette* and *mouette*) and differ from the non-suffixed condition (e.g., *fortune*). In other words, the presence of an affix in the truly suffixed and pseudo-suffixed word conditions should lead to a reduction of the N400 amplitude, relative to the non-suffixed control. Moreover, a word-advantage (i.e., a reduction in N400 amplitude) should be evident for both truly suffixed and pseudo-suffixed words relative to their corresponding nonword counterparts, whereas no such difference would arise between non-suffixed words and their corresponding nonword controls. Given that the behavioural

ACCEPTED MANUSCRIPT

responses in an unprimed lexical decision paradigm represent the end point of the lexical decision process, it is impossible to clearly tease apart early form-based vs. later semantic influences on the observed response latencies or error rates (which is precisely why we combined behavioural measures with the recordings of ERPs in this study). That is, the behavioural component of the morpheme facilitation effect may emerge for both truly suffixed and pseudo-suffixed words (Lavric et al., 2012), or alternatively only for truly suffixed words (Rastle, Davis, Marslen-Wilson, & Tyler, 2000).

In addition to the word trials, two sets of pseudo-suffixed and non-suffixed nonwords were included in the lexical decision task. For each word (e.g. *mouette*), we created a corresponding pseudo-suffixed nonword (by replacing the ending with a suffix, e.g. *mouesse*, where *-esse* is a real French affix) and a corresponding non-suffixed nonword (by replacing the ending with a non-suffix, e.g. *mouipe*, where *-ipe* is not an affix). As such, we obtained item triplets with an identical word stem (e.g., *mou*) at word onset. If auditorily presented complex nonwords are decomposed into morphemic subunits just like words are, we would expect to see a “morpheme interference effect” (Caramazza, Laudanna, & Romani, 1988; Taft & Forster, 1975) for pseudo-suffixed compared to non-suffixed nonwords. That is, nonwords consisting of stems and suffixes (e.g., *mouesse*) should be harder to reject and therefore elicit an enhanced N400 amplitude compared to nonwords consisting of stems and non-suffixes (e.g., *mouipe*).

Generally, the N400 effect for morphologically complex words occurs later in the auditory modality than in the visual modality, and response latencies are overall slower than in the visual modality (for a direct comparison of suffixed inflected word processing in the visual and auditory modalities, see Leinonen et al., 2009). **Moreover, the N400 in the auditory modality is more spread out than the N400 in the visual modality (e.g., see Figure 1 in Perre, Midgley, & Ziegler, 2009).** Given that all our stimuli were presented in

the auditory modality, we expected to observe a relatively late emergence of the N400 effect and overall slower response times than standardly reported in the visual modality.

2. Method

2.1. Participants

Twenty-eight healthy, native French-speaking volunteers (16 females, 12 males) participated in this study for monetary reimbursement. Their ages ranged from 18 to 46 years (mean age = 23.8, SD = 6.8). All were right-handed, with normal or corrected or corrected-to-normal vision and none reported any neurological or language impairment. All participants were university students and gave written consent. Ethics approval was obtained from the Institutional Review Board of Aix-Marseille University.

2.2. Materials

Fifty-one French pseudo-suffixed target words (e.g. *mouette*), were selected from the Lexique database (New, Pallier, Brysbaert, & Ferrand, 2004). The pseudo-suffixed target words were selected such that the whole word never shared any semantic relationship with the embedded pseudo-stem (*mou*). In addition, we chose 51 truly suffixed target words (e.g. *pochette*), where whole word and stem (*poche*) always shared a semantic relationship. The truly and pseudo-suffixed target words shared the exact same suffixes (and thus the exact same final string of phonemes). Finally, a third set of 51 non-suffixed items was selected from Lexique (e.g. *fortune*), which consisted of an embedded word (*fort*) and a meaningless non-morphemic ending (*-une*). Given that the items were presented auditorily, we ensured that the pronunciation of the stem did not change in the embedded context. All 153 target words were nouns.

ACCEPTED MANUSCRIPT

The three sets of target words were matched as closely as possible on written word frequency, spoken word frequency, number of letters, number of phonemes, number of syllables, orthographic neighbourhood (*orthographic N*), phonological neighbourhood (*phonological N*), orthographic Levenshtein distance 20 (*old20*), phonological Levenshtein distance 20 (*pld20*), ending length and uniqueness point (*UP*). Semantic relatedness values between whole words and their embedded words (e.g. *mouette* and *mou*) were extracted using the Latent Semantic Analysis Web facility (<http://lsa.colorado.edu>; Landauer & Dumais, 1997). This analysis revealed that semantic relatedness values in the truly suffixed condition (0.24) were significantly higher than those in both the pseudo-suffixed (0.11) and the non-suffixed conditions (0.11), but the pseudo-suffixed and non-suffixed conditions did not differ (see Figure 1). Crucially, a close inspection of the 20 nearest semantic neighbours of our target words showed that the results were highly inaccurate, presumably due to a bug in the French corpus analysis. We therefore applied the Latent Semantic Analysis model (Landauer & Dumais, 1997) to a lemmatised corpus of 1.2 Go of French books, which replicated the results of the Latent Semantic Analysis Web facility (i.e. the semantic similarity between words and their embedded stems was significantly higher for truly suffixed words [0.54] than for pseudo-suffixed words [0.29], and 0.28 for non-suffixed words [0.28]). The mean item characteristics for each condition and *p*-values for the critical comparisons are reported in Appendix A.

- Figure 1 -

For each word target, a pseudo-suffixed and a non-suffixed target nonword were created (306 nonwords in total). Pseudo-suffixed nonwords included the same stem but different affix, such that the whole letter string was not a word (e.g., for *mouette*, we selected the pseudo-suffixed nonword *mouesse*). The suffixes of the word targets were 'recycled' in the nonword targets by changing suffixes between different stems. Non-suffixed nonwords included the same stem with a non-morphemic ending (e.g. for *mouette*, we selected the non-

ACCEPTED MANUSCRIPT
suffixed nonword *mouipe*). The non-morphemic endings of the nonwords were identical to the non-morphemic endings of the non-suffixed words. Only a few phonemes had to be changed in the nonwords to maintain the pronounceability of the letter string, and to match syllable structure. Three counterbalanced experimental lists were created using a Latin square design, such that every embedded word only occurred once in each list (e.g. *mouette*, *mouesse* and *mouipe* all occurred in different experimental lists, in order to avoid repetition of the embedded word *mou* within lists and therefore within participants).

Auditory targets were produced with the OS X Speech Synthesizer, using the French male voice 'Thomas'. The naturalness of the synthesized files was checked by two independent French native speakers. The speaking rate was set to 180 words per minute. All stimuli had a bit rate of 705 kbps. Auditory files were edited to ensure that any silence at the beginning and end of each item was removed. The mean stimulus duration across all items was 628ms. The mean durations per word-type and condition are presented in Table 1. A list of all items is presented in Appendix B.

- Table 1 -

2.3. Procedure

Stimuli were presented using experimental software EPrime 2.0 (Psychology Software Tools, Pittsburgh, PA). Participants were tested individually in a Faraday cage. Each trial consisted of a fixation cross which appeared in the centre of an LCD computer screen for 1000 ms, followed by the auditory target. The inter-trial interval was 1000 ms. If participants did not respond after 3 seconds had elapsed, the experiment proceeded automatically to the next trial. The auditory target words were presented via headphones binaurally. Participants were instructed to decide as quickly and accurately as possible if the presented items were real French words or not. Participants responded by pressing one of two different response buttons. The right hand was used to respond YES and the left hand was used to respond NO. Stimuli were presented in randomized order. All participants completed the three

experimental lists, but in randomised order (i.e. six possible list orders were created, which were assigned to 4 or 5 participants each). The ERP cap installation time (prior to testing) took between 20-30 minutes per participant. Each experimental list took approximately 10 minutes to complete (30 minutes in total).

3. Results

3.1. ERP recording and Pre-processing

ERPs were recorded using the Biosemi Active2 system at a sampling rate of 2048Hz. Sixty-four electrodes were arranged on the participants' scalps using the 10-20 placement system. Four additional electrodes were used to record vertical and horizontal eye movements (vEOGs and hEOGs, respectively). Two electrodes were positioned on the right and left mastoids; the left mastoid served as reference during recording. Throughout EEG acquisition, electrode impedance was kept below 20k Ω .

Processing of EEG data was carried out using the EEGLAB toolbox (Delorme & Makeig, 2004). The acquired EEG was down-sampled to 512Hz offline and a second-order Butterworth band-pass filter (0.1Hz – 40Hz) was applied. The data were re-referenced to the average of the right and left mastoids. Noisy electrodes were detected in a semi-automatic manner by observing the electrode spectra and by calculating the kurtosis for each channel; those channels with a kurtosis value exceeding 5 (z-score) were considered for rejection. Ocular movements were corrected with Independent Component Analysis (ICA) by calculating the infomax ICA algorithm (Bell & Sejnowski, 1995) on the 64 scalp electrodes. To facilitate the calculation of clean ICA components, intervals of signal presenting very large noise exceeding 75 μ V were detected automatically and removed from the continuous data before ICA calculation. Those components corresponding to eye artefacts were identified via component topography, spectra and time course and only those components corresponding

ACCEPTED MANUSCRIPT

to eye artefacts were removed; this generally corresponded to the first component. After ICA correction, those electrodes rejected due to noise were interpolated using spherical spline interpolation. The continuous data were then segmented into individual trials using the end of the embedded stems or pseudo-stems (e.g. the end of *mou* in *mouette*) as T0¹. For each trial a pre-stimulus interval of 200ms was defined to ensure the same baseline activity for all word types. With T0 at the offset of the embedded stems, a post-stimulus interval of 700ms was defined. The pre-stimulus interval served as baseline, and baseline correction was carried out for all trials. Those trials in which the participants' reaction times fell outside the lower and upper limits of 200ms and 3000ms, respectively, were automatically rejected. Noisy trials were detected semi-automatically. Firstly, those epochs with activity exceeding a limit of $\pm 75\mu\text{V}$ were removed. Linear drift was assessed and instances of drift exceeding $10\mu\text{V}$ were marked after which the remaining epochs were assessed using kurtosis, applying a limit of 5 (z-scores).

Five participants were excluded from the study due to the large number of trials rejected because of high error rates, extreme reaction times, or noisy EEG data. In addition, twelve words in the non-suffixed condition were incorrectly classified as non-morphological as they consisted of pseudo-stem + pseudo-suffix (highlighted with an asterisk in Appendix B) and were therefore excluded from behavioural and ERP analyses. Table 2 summarizes the total number of epochs retained after data cleaning as well the average proportion of epochs rejected across participants.

- Table 2 -

3.2. Statistical analyses

3.2.1. Behavioural analyses

¹ In an earlier version of our manuscript the continuous data were segmented using the onset of the auditory target as T0, while using the same baseline. These earlier analyses led to similar but later effects, which are reported in the supplementary materials.

Lexical decisions to word and nonword targets were analysed as follows. Reaction times and error rates were analysed separately. Incorrect responses were removed from the reaction time (RT) analysis (7.7% of all data). Inverse RTs ($-1000/\text{RT}$) were calculated for each participant to correct for RT distribution skew and used throughout the analyses (Kliegl, Masson, & Richter, 2010). RTs and error rates are presented in Tables 3 and 4 (see below) and were analysed for each participant.

We used linear mixed-effect modelling to perform the main analyses (Baayen, 2008; Baayen, Davidson, & Bates, 2008). Fixed effects, random effects, and random slopes were only included if they significantly improved the model's fit in a backward stepwise model selection procedure. Models were selected using chi-squared log-likelihood ratio tests with regular maximum likelihood parameter estimation. The model was refitted after excluding data-points whose standardised residuals were larger than 2.5 in absolute value (see Baayen, 2008), which led to the removal of 1.6% of the nonword data and 2.4% of the word data. Trial order was included to control for longitudinal task effects such as fatigue or habituation. Experimental list order was included as a covariate, in order to examine whether or not the observed effects would be modulated by number of exposures to the embedded word. In addition, we included word properties (i.e., subtitle word frequency, number of phonemes, number of syllables, phonological Levenshtein distance 20 (PLD20), uniqueness point (UP), and semantic relatedness proportion (LSA)) for both whole words (i.e. target words) and embedded words as covariates in the word analyses, to control for item specific differences across target words. The subtitle word frequencies were extracted from the film subtitle corpus in Lexique based on 52 million French words (New, Brysbaert, Veronis, & Pallier, 2007), coming from a variety of films, and then transformed using the Zipf scale (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014). In addition, the duration of the embedded stem or pseudo-stem (in ms) was added as a covariate in the analyses to control for

ACCEPTED MANUSCRIPT
differences in stimulus length. All continuous variables were centered and the lmer default coding for treatment contrasts (i.e., in alphabetical order) was used for the item type variable.

Word targets and nonword targets were first analysed separately, followed by a combined analysis. In the word data, factor item type was a 3-level factor (truly suffixed, pseudo-suffixed, non-suffixed), whereas in the nonword data, factor item type was a 2-level factor (suffixed, non-suffixed). Linear mixed-effects model as implemented in the lme4 package (Bates, Maechler, Bolker, & Walker, 2014) in the statistical software R (Version 3.0.3; RDevelopmentCoreTeam, 2008) were fitted using the above described model selection procedure. P-values were determined using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2014).

Error analyses followed the same logic as the RT analyses. We applied a binomial variance assumption to the trial-level binary data using the function *glmer* as part of the R-package *lme4*.

3.2.1.1. Words.

In the reaction time analyses, the final linear mixed-effect model included five fixed effects factors (item type, list order, whole-word frequency, embedded word frequency, embedded word duration), random intercepts for participants and items, and random slopes for list order by participants. A significant effect of item type showed that participants responded more slowly to non-suffixed words than to truly suffixed words ($t = 3.89, p < .001$) and pseudo-suffixed words ($t = 3.39, p < .001$). The difference between the truly suffixed and the pseudo-suffixed condition was not significant ($t = 0.53, p = .598$). There was also a marginal main effect of list order ($X^2(1) = 3.28, p = .070$) showing that participants responded gradually more slowly in the second and third experimental list compared to the first list. In addition, the analyses revealed a significant main effect of whole-word frequency ($X^2(1) = 22.54, p < .001$) showing that participants responded faster to high frequency words, as well

ACCEPTED MANUSCRIPT

as a significant main effect of embedded word frequency ($X^2(1) = 6.85, p = .009$), showing that participants responded slower to targets with embedded high frequency words than to targets with embedded low frequency words (see Figure 2). The interactions between whole-word frequency and item type as well as the interaction between embedded word frequency and item type were not significant ($X^2(2) = 2.24, p = .326$; $X^2(2) = 0.96, p = .619$). To examine whether or not semantics influenced the activation of the embedded word, we tested the interaction between embedded word frequency and LSA semantic overlap, which turned out to be non-significant ($X^2(1) = 2.05, p = .152$). Finally, there was a significant main effect of embedded word duration ($X^2(1) = 42.23, p < .001$), showing that response times increased with increasing stimulus durations. No other effects were significant.

-Figure 2-

In the error analyses, the final linear mixed-effect model included two fixed effects factors (whole-word frequency, number of phonemes of the whole word), and random intercepts for participants and items. The results showed that participants made less errors responding to high frequency than low frequency words ($z = 3.61, p < .001$) and less errors responding to shorter than longer words ($z = 2.56, p = .011$).

-Table 3-

3.2.1.2. Nonwords.

In the reaction time analyses, the final linear mixed-effect model included one fixed effects factor (list order), random intercepts for participants and items, and random slopes for list order by participants. The results revealed a marginal main effect of list order ($t = 1.72, p = .099$), suggesting that participants were gradually able to more rapidly reject nonwords in the second and third experimental list compared to the first list. There were no other significant effects.

In the error analyses, the final linear mixed-effect model included two fixed effects factors (item type; list order), random intercepts for participants and items, and random slopes

ACCEPTED MANUSCRIPT
for item type by participants. There was a significant effect of item type ($z = 7.51, p < .001$), showing that participants made more errors rejecting suffixed nonwords than non-suffixed nonwords. There was also a significant main effect of list order ($z = 4.84, p < .001$), indicating that participants gradually made less errors in the second and third experimental list compared to the first list, which is consistent with the training effect seen in the reaction time data.

-Table 4-

3.2.1.3. Words vs. nonwords.

In the reaction time analyses, the final model included two fixed effects factors (lexicality; item type), their interaction, random intercepts for participants and items, and random slopes for lexicality by participants. Reaction time analyses revealed a marginal main effect of lexicality ($X^2(1) = 3.57, p = .059$), suggesting that participants were on average faster at responding to words than to nonwords. There was also a significant interaction between lexicality and item type ($X^2(2) = 37.44, p < .001$), suggesting that the "word-advantage" was greater in the truly suffixed and pseudo-suffixed conditions than in the non-suffixed condition ($t = 6.47, p < .001; t = 6.51, p < .001$), whereas the word-advantage was equally large in the suffixed and pseudo-suffixed conditions ($t = 0.01, p < .992$). No other effects were significant.

In the error analyses, the final model included two fixed effects factors (lexicality; list order), random intercepts for participants and items, and random slopes for lexicality by participants. The main effect of lexicality ($z = 6.12, p < .001$) showed that overall participants made more errors responding to words than to nonwords, which indicates that the lexicality main effect in the RT data is due to an accuracy-speed trade off. There was also a significant effect of list order ($z = 2.08, p = .038$), indicating that participants gradually made less errors in the second and third experimental lists compared to the first list.

3.2.2. ERP analyses

Response contingent analyses were performed on ERP data. **Throughout the ERP analyses, T0 was shifted to the end of the embedded stem or pseudo-stem.**

3.2.2.1. Words.

Figure 3 presents the scalp maps of the activity over time, from -200 to 700ms for each word condition; each scalp map presents the mean activity in a 100ms time window. In addition, a plot of the Global Field Power (GFP) of each condition was plotted from -200 to 700ms. The GFP is the spatial root mean square across all electrodes and provides a global measure of the electric activity at the level of scalp (Skrandies, 1990). One of the main advantages of GFP is that it yields a general estimate of electric activity that does not suffer from spatial bias, from the variations in the latencies of peaks activity that can be observed across different electrodes (Michel et al., 2004). Here, the GFP provides a clear picture of the difference and similarity in the time-course of activity for the truly suffixed (TS), pseudo-suffixed (PS) and non-suffixed (NS) conditions. Both the scalp maps and the GFP plots of the three word conditions demonstrate the similarity between the truly suffixed and pseudo-suffixed conditions over the entire trial, as well as the divergence of the NS condition activity at a later time window spanning 400 to 600ms.

- Figure 3 -

To determine when and over which brain regions differences emerge, without having to define temporal windows or electrodes of interest a priori, the participant-level grand-average ERP data were subjected to a permutation test with cluster-based correction (Maris & Oostenveld, 2007) for each pairwise comparison (TS vs. NS, PS vs. NS, TS vs. PS). This analysis was carried out using the Matlab toolbox, FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). The cluster-corrected permutation test simplifies the resolution of the multiple comparison problem by correcting at the level of clusters formed on the basis of spatio-temporal adjacency. To calculate the permutation distribution, 2000 random partitions were computed and only those samples with a permutation p-value below the critical cluster

ACCEPTED MANUSCRIPT
alpha-level ($p \leq .05$) were selected. Clusters were formed from these samples based on an adjacency criterion at the spatio-temporal level. Spatio-temporal adjacency was established on the basis of a minimum of 3 electrodes, and neighbouring electrodes were defined using the triangulation algorithm implemented in FieldTrip. Finally, those clusters with a Monte-Carlo p-value less than .025 (two-tailed test) were retained; a two-tailed test was carried out as we were interested in both negative and positive directions. The cluster-corrected permutation test was carried out for all time points (2ms time windows), however to facilitate visualisation, the results of the test are presented as topographies over time in 100ms time steps.

A cluster-corrected permutation test revealed that the NS-TS difference was statistically significant from 300-500ms and the NS-PS difference was statistically significant from 400-500ms, and that these differences were concentrated over central and parietal regions (Figure 4). The TS vs. PS comparison did not reveal any significant differences according to the permutation test.

-Figure 4 -

The effect of item type on the grand average ERP activity is further highlighted in Figures 5a, 5b, and 5c. Significant differences between the NS vs. TS and NS vs. PS comparisons emerged over central and posterior electrodes of the right and left hemispheres. Figures 5a, 5b and 5c present a subset of these electrodes (C3, Cz, C4, P3, Pz, P4, PO3, POz, PO4). Taken together, the results of the spatio-temporal analysis (cluster-corrected permutation test) and the analysis of the ERP data suggest that non-suffixed words elicited a significantly greater negativity than both truly suffixed and pseudo-suffixed words over central and parietal electrodes from 300-500ms following stimulus onset. There were no significant differences between the truly suffixed and pseudo-suffixed conditions (Figure 5c).

- Figures 5a, 5b and 5c -

ACCEPTED MANUSCRIPT

In addition, we investigated the possible variation of ERP amplitude as a function of word frequency and uniqueness point, which have been found to influence ERP amplitude and in particular N400 amplitude (e.g., Dufour, Brunellière, & Frauenfelder, 2013; O'Rourke & Holcomb, 2002). We applied an ERP-image visualisation method developed by Delorme, Miyakoshia, Jung, and Makeiga (2015), which revealed no significant variation of ERP-activity as a function of spoken word frequency or uniqueness point.

3.2.2.2. Words vs. nonwords.

Grand-average ERPs at all electrodes comparing words and nonwords for each word type were calculated. **As in the word analyses, T0 was shifted to the end of the embedded stem or pseudo-stem.** To determine time windows presenting significant differences, three pairwise comparisons were carried out for each word type by applying a permutation test with FDR correction ($p \leq .05$) on the grand-average ERP data: non-suffixed nonwords (nonword-NS) vs. suffixed nonwords (nonword-S), words vs. non-suffixed nonwords and words vs. suffixed nonwords (see Figures 6, 7, and 8). While the permutation test did not reveal any statistically significant differences between the two types of nonwords (nonword-NS vs. nonword-S), it did reveal a significant difference between words and nonwords. The word-nonword difference emerged between 300-700ms in the truly suffixed and pseudo-suffixed conditions (Figures 6 and 7) and between 400-700ms in the non-suffixed condition (Figure 8). In addition, the statistical analysis revealed that in the non-suffixed and pseudo-suffixed conditions the difference between non-suffixed nonwords and words (e.g. *fortaque* vs. *fortune*) emerged slightly earlier than the difference between suffixed nonwords and words (e.g. *forteur* vs. *fortune*).

-Figures 6, 7, and 8 -

To obtain insight into the spatial and temporal distribution of the word-nonword differences observed in the grand-average ERPs, a cluster-corrected permutation test ($p \leq .025$,

two-tailed) was carried out on the participant-level grand-average data (see Figure 9).

Consistent with the ERP results in Figures 6, 7, and 8, the results revealed an enhanced negativity for nonwords compared to words across all word types, in a time window between 300 – 700ms. This N400 effect was widely distributed, spanning frontal, central and parietal regions, for the truly-suffixed and pseudo-suffixed conditions in particular. Interestingly, the spatial-temporal analysis revealed that the latency of this late effect was not uniform across the three conditions (see Figure 9): its onset occurred earliest in the truly-suffixed condition (300ms after stimulus onset), second in the pseudo-suffixed condition (300-400ms after stimulus onset) and latest in the non-suffixed condition (400-500ms after stimulus onset). Nonwords showed a sustained negative deflection across all item types (see Figures 6, 7, and 8).

- Figure 9 -

It is noteworthy, however, that the earlier difference observed for the non-suffixed condition over the 0-200ms time window did not reach statistical significance within the two-tailed cluster-corrected permutation test, given the parameters set for its calculation (in particular, the requirement of a minimum of 3 electrodes to establish spatial adjacency). Only when we carried out statistical analysis for individual electrodes ($p \leq .05$, *fdr* corrected) did we find a significant effect for a limited number of electrodes, which was concentrated over the left posterior region (see Figure 10).

- Figure 10 -

4. Discussion

The aim of the present study was to examine morphological processing during spoken word recognition using ERP recordings in combination with an auditory lexical decision task. To this end, truly suffixed words were compared to pseudo-suffixed and non-suffixed words, and pseudo-suffixed to non-suffixed nonwords. **T0 was shifted to the end of**

the embedded stem or pseudo-stem throughout the ERP analyses. The results can be summarised as two key findings, which we discuss below.

First, both EEG and behavioural results clearly dissociate the two morphological conditions from the non-morphological condition, thus providing evidence for a robust morpheme facilitation effect. Participants responded more slowly to non-suffixed words than to truly suffixed and pseudo-suffixed words, but no difference was found between the two suffixed conditions. This is in line with previous evidence from spoken word recognition, demonstrating that auditorily presented morphologically complex words are easier to classify than non-suffixed words, because hypothetically, access to the affix and the stem facilitates word recognition (Gwilliams et al., 2015). In line with the behavioural findings, the ERP waveforms of the two suffixed conditions did not significantly differ from each other (see Figure 5c). The presence of an affix in the truly suffixed and pseudo-suffixed words led to a reduction in N400 amplitude relative to the non-suffixed control condition (see Figures 5a and 5b), thus providing evidence for morphological processing operating *independently* of semantics (Lavric et al., 2012).

One explanation for this pattern of results is that the spoken word recognition system benefits from *the principle of full decomposition* combined with *the principle of edge-aligned embedded word activation* (Grainger & Beyersmann, 2017), suggesting that word recognition is facilitated when the whole letter string can be completely divided into potential constituent morphemes. For example, *hunter* can be **parsed** into *hunt* and *er*, *corner* can be **parsed** into *corn* and *er*, but for non-suffixed words like *cashew* the principle of full decomposition fails, because the edge-aligned embedded word (*cash*) cannot be combined with another morpheme (*cash* + ?) to create an exhaustive decomposition of the full word. The principle of full decomposition has been previously described as a mechanism underlying visual word recognition and can account for a wide range of findings from visual lexical decision and masked priming (e.g., Grainger & Beyersmann, 2017; Münte, Say, Clahsen, Schiltz, & Kutas,

1999; Stockall & Marantz, 2006; Taft & Forster, 1975). Here we propose to extend the principle of full decomposition to spoken word recognition, as it unambiguously explains why the morpheme facilitation effect is observed in the truly suffixed and pseudo-suffixed conditions, but not in the non-suffixed control condition (for a related proposal on auditory prefixed word processing, see Wurm, 1997).

The second key finding, which is complementary to the first one, is that the reduction of the N400 amplitude in the truly suffixed relative to the non-suffixed condition emerged slightly earlier (from 300ms onwards; see Figure 5a) than in the pseudo-suffixed condition (from 400ms onwards; Figure 5b). This result is further enhanced by the fact that we also found an enhanced and sustained N400 amplitude for nonwords compared to words which varied between the three experimental conditions: the word-nonword difference occurred earliest in the truly suffixed condition (e.g., *pochette* vs. *pochure/pochique*; Figure 6), slightly later in the pseudo-suffixed condition (e.g., *mouette* vs. *mouesse/mouipe*; Figure 7) and latest in the non-suffixed condition (e.g., *fortune* vs. *forteur/fortaque*; Figure 8). Thus, although the difference between the truly and pseudo-suffixed conditions was not significant in the 300-400 ms time window, the permutation and word-nonword analyses appear to provide cumulative evidence for an effect of semantic transparency which is coupled with the semantically blind morphological parsing effect seen in the two morphological conditions with word stimuli. In other words, the here observed N400 appears to be a reflection of not just one, but two additive effects: a semantically blind morpheme facilitation effect and as well as a semantic transparency effect. The reason why these two effects emerge nearly simultaneously can be explained by the interplay between a purely form-based morphological parsing mechanism and the added influence of semantics. We hypothesise that both truly suffixed (e.g., *hunt + er*) and pseudo-suffixed words (e.g., *corn + er*) are exhaustively decomposed into morphemic sub-units (e.g., Rastle & Davis,

2008; Rastle et al., 2004), and that this decomposition process facilitates the activation of the corresponding whole words (*hunter* and *corner*) and the herewith associated lexical decision response. The morpheme facilitation effect therefore explains the processing advantage in the two morphological conditions compared to the non-suffixed condition. The semantic transparency effect on the other hand suggests that the semantic transparency of truly suffixed words provides an additional boost in activation to the whole word representation, which could be the cause of the earlier onset of effects seen with semantically transparent words. While semantically transparent complex words like *hunter* benefit from feedback connections from the semantic level to the lexical level, semantically opaque words like *corner* do not benefit from semantic support (Diependaele, Sandra, & Grainger, 2009; Grainger & Beyersmann, 2017). The spoken word recognition system might use semantic feedback to more efficiently evaluate the morphological segmentation process, which would explain why the N400 effect emerged slightly earlier in the truly suffixed condition, compared to the other two conditions.

This complex interplay between the semantic transparency and morpheme facilitation effects also fits with the renowned complexity of the N400 component, with difficulty in processing and amount of semantic activation being two likely contributions to variations in N400 amplitude. The N400 component is commonly associated with semantic influences on word recognition (Holcomb & Neville, 1990), but is also thought to reflect word form processing difficulty (e.g., Perre et al., 2009; Winsler, Midgley, Grainger, & Holcomb, 2018), with N400 amplitude increasing with an increased difficulty in identifying words and associating the word identities with meaning (see Grainger & Holcomb, 2009, for a review of the evidence). Of course, in sentence comprehension experiments, other factors, such as lexically-based prediction (e.g., Frank & Bod, 2011) and orthographically guided lexical prediction (e.g., Kim & Lai, 2012; Laszlo & Federmeier, 2009) can also play a role.

ACCEPTED MANUSCRIPT
The word-nonword differences in the two suffixed conditions (Figures 6 and 7)

emerged about the same time as the morphological effects in these conditions (Figures 5a and 5b), presumably because of the linearity of the auditory input signal. Since the onset of every auditory stimulus (without exception) was always a word (e.g. *mou* in *mouette*, *mouesse* and *mouipe*), participants had to delay their lexical decisions as well as morphological processing until the final part of the input signal, which is likely why word-nonword differences and the effect of morphology coincided in these data.

With respect to the timing of ERP responses in visual versus auditory lexical decision, our study is only partially consistent with the results of ERP studies that have examined morphological processing in the visual modality. On the one hand, the reduction of N400 amplitude is comparable to results from visual word recognition, where semantically blind morphological segmentation mechanisms are typically found around 200-500 ms after stimulus onset (e.g., Beyersmann, Ziegler, et al., 2016; Lavric et al., 2012; Longtin et al., 2003; Quémart, Casalis, & Colé, 2011; Rastle & Davis, 2008; Rastle et al., 2004). For instance, Beyersmann et al. (2014) reported a reduction of the N400 amplitude between 200-400 ms in the morphological condition but not in the semantic or orthographic control conditions. Lavric et al. (2007) found a reduced N400 amplitude between 340-460 ms in both the morphological and pseudo-morphological conditions, but not in the non-suffixed condition. **On the other hand, our results deviate from visual word recognition studies in the sense that we did not observe any effects on earlier ERP components. For instance, Lavric and colleagues reported a study in which they combined visual lexical decision with ERP recordings, showing an early difference from ~190 ms between suffixed and pseudo-suffixed words compared to non-suffixed words (Lavric et al., 2012). A masked primed lexical decision study by Morris et al. (2007) revealed an N250 effect which was significant in the truly suffixed but not in the pseudo-suffixed condition (see also Lavric, Clapp, & Rastle, 2007). In contrast, we did not witness any ERP components of**

ACCEPTED MANUSCRIPT

morphological processing that preceded the N400 effect. Morphological information may be more readily available in visual compared to auditory word recognition, because the visual system can perform a quick initial form-based analysis of the letter string. In auditory word recognition, the acoustic signal unfolds in time and is more variable (e.g., depending on the speaker's voice and intonation) which may constrain the speed by which the listener is able to morphologically segment any given input signal. In addition, as Rastle (2018) pointed out, while printed morphemic sequences possess a high degree of regularity (e.g. the suffix *-ed* denotes the past tense), the phonetic forms of morphological units are considerably less consistent (e.g. the suffix *-ed* maps onto multiple sound sequences /ɪd/, /ɜd/, and /əd/). The inconsistency in sound-to-morpheme mappings is likely to provide a challenge to the morphological segmentation system in the auditory domain and thus offers an explanation for why the here observed morphological effects emerge comparatively later compared to the visual domain.

An additional aspect of our findings, which provides further insights into the nature of the morphological decomposition process, is that auditory lexical decision latencies decreased with increasing whole-word frequencies, but increased with increasing embedded word frequencies (see Figure 2). The interaction between embedded word frequency and LSA semantic relatedness was not significant, suggesting that embedded words were activated independently of whether or not they shared a semantic relationship with the whole word. The main effect of embedded word frequency is consistent with a growing body of evidence (e.g., Amenta, Marelli, & Crepaldi, 2015; Beyersmann et al., 2015; Beyersmann, Cavalli, Casalis, & Colé, 2016; Marelli & Amenta, 2018; Marelli, Amenta, & Crepaldi, 2015; Taft, Li, & Beyersmann, 2018), suggesting that embedded words are always activated independently of whether they are accompanied by an affix (as in *farm + er* or *corn + er*) or a non-affix (as in *cash + ew*). The activation of embedded words (e.g., *cash*) then generates lexical competition with the whole word (e.g., *cashew*), thus leading to an increase in response times across

conditions (Grainger & Beyersmann, 2017). As the left panel of Figure 2 shows, the greater the frequency of the embedded word, the greater the competition. Crucially, although the embedded word frequency effect was observed across all item types, our combined behavioural and ERP data clearly demonstrate a robust morpheme segmentation effect in the two morphological conditions compared to the non-affixed control. This indicates that differences between (pseudo-)affixed and non-affixed words might depend on factors that are not necessarily **related** to the stem. Instead, it seems to be the absence of a suffix in the non-suffixed items that is driving differences in effect sizes across item types. The presence of a suffix in the truly and pseudo-suffixed words would have provided a boost in activation to the whole stimulus, thus leading to faster response times and reduced N400 amplitudes in these conditions.

Finally, the behavioural results showed that participants made more errors rejecting pseudo-suffixed nonwords than non-suffixed nonwords, thus replicating the classical morpheme interference effect (Caramazza et al., 1988; Taft & Forster, 1975). The presence of the embedded morphemic unit increases the 'word-likeness' of the nonword, which as a result becomes harder to reject. The EEG data did not reveal a statistically significant difference between the pseudo-suffixed and non-suffixed nonword conditions. It is possible that the morpheme interference effects in the behavioural data reflect later post-lexical processing stages, which are not picked up in the ERP signal. In any case, it is worth noting that previous ERP studies have not always confirmed the morpheme interference effect for complex nonwords (e.g., Leinonen et al., 2009), suggesting that this effect may not be as robust as the morpheme facilitation effect that is typically seen in the word data.

In sum, the present data shed new light onto the mechanisms involved in the recognition of morphologically complex spoken words. Both the behavioural and ERP data concurrently suggest that not only words with a true morphological structure, but also words with a pseudo-morphological structure are decomposed into morphemic sub-units during

spoken word recognition. **This indicates that morphemic units are automatically identified, independently of semantics. The earlier onset of the morpheme facilitation effect in the truly suffixed condition suggests, however, that feedback from semantics can provide an additional activation boost in the processing of spoken words with a genuine morphological structure.**

5. References

- Amenta, S., & Crepaldi, D. (2012). Morphological processing as we know it: an analytical review of morphological effects in visual word identification. *Frontiers in Psychology*, 3, 1-12.
- Amenta, S., Marelli, M., & Crepaldi, D. (2015). The fruitless effort of growing a fruitless tree: Early morpho-orthographic and morpho-semantic effects in sentence reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 41(5), 1587-1596. doi:10.1037/xlm0000104
- Andrews, S., & Lo, S. (2013). Is morphological priming stronger for transparent than opaque words? It depends on individual differences in spelling and vocabulary. *Journal of Memory & Language*, 68, 279-296.
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412.
- Balling, L. W., & Baayen, R. H. (2008). Morphological effects in auditory word recognition: Evidence from Danish. *Language and Cognitive Processes*, 23, 1159–1190.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-5. <http://CRAN.R-project.org/package=lme4>. Retrieved from
- Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural computation*, 7(6), 1129-1159.
- Beyersmann, E., Casalis, S., Ziegler, J. C., & Grainger, J. (2015). Language proficiency and morpho-orthographic segmentation. *Psychonomic Bulletin & Review*, 22, 1054–1061.

- Beyersmann, E., Cavalli, E., Casalis, S., & Colé, P. (2016). Embedded Stem Priming Effects in Prefixed and Suffixed Pseudowords. *Scientific Studies of Reading, 20*(3), 220-230. doi:10.1080/10888438.2016.1140769
- Beyersmann, E., Iakimova, G., Ziegler, J. C., & Colé, P. (2014). Semantic processing during morphological priming: An ERP study. *Brain Research, 1579*, 45-55.
- Beyersmann, E., Ziegler, J. C., Castles, A., Coltheart, M., Kezilas, Y., & Grainger, J. (2016). Morpho-orthographic segmentation without semantics. *Psychonomic Bulletin & Review, 23*, 533-539. doi:10.3758/s13423-015-0927-z
- Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition, 28*, 297-332.
- Cavalli, E., Cole, P., Badier, J. M., Zielinski, C., Chanoine, V., & Ziegler, J. C. (2016). Spatiotemporal Dynamics of Morphological Processing in Visual Word Recognition. *Journal of Cognitive Neuroscience, 28*(8), 1228-1242.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9-21.
- Delorme, A., Miyakoshia, M., Jung, T.-P., & Makeiga, S. (2015). Grand average ERP-image plotting and statistics: A method for comparing variability in event-related single-trial EEG activities across subjects and conditions. *Journal of Neuroscience Methods, 250*, 3-6.
- Diependaele, K., Sandra, D., & Grainger, J. (2009). Semantic transparency and masked morphological priming: The case of prefixed words. *Memory & Cognition, 37*, 895-908.
- Dominguez, A., De Vega, M., & Barber, H. (2004). Event-related brain potentials elicited by morphological, homographic, orthographic, and semantic priming. *Journal of Cognitive Neuroscience, 16*(4), 598-608.

- Dufour, S., Brunellière, A., & Frauenfelder, U. H. (2013). Tracking the Time Course of Word-Frequency Effects in Auditory Word Recognition With Event-Related Potentials. *Cognitive Science*, 37(3), 489–507.
- Emmorey, K. D. (1989). Auditory morphological priming in the lexicon. *Language and Cognitive Processes*, 4(2), 73-92.
- Ettinger, A., Linzen, T., & Marantz, A. (2014). The role of morphology in phoneme prediction: Evidence from MEG. *Brain and Language*, 129, 14-23.
- Feldman, L. B., Milin, P., Cho, K. W., Moscoso Del Prado Martin, F., & O'Connor, P. A. (2015). Must analysis of meaning follow analysis of form? A time course analysis. *Front Hum Neurosci*, 9, 111. doi:10.3389/fnhum.2015.00111
- Feldman, L. B., O'Connor, P. A., & Moscoso del Prado Martin, F. (2009). Early morphological processing is morpho-semantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review*, 16(4), 684-691.
- Frank, S. L., & Bod, R. (2011). Insensitivity of the human sentence-processing system to hierarchical structure. *Psychological Science*, 22(6), 829-834. doi:10.1177/0956797611409589
- Frost, R., Kugler, T., Deutsch, A., & Forster, K. I. (2005). Orthographic structure versus morphological structure: Principles of lexical organisation in a given language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 1293–1326.
- Gonnerman, L. M., Seidenberg, M. S., & Andersen, E. S. (2007). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General*, 136, 323-345.

- Grainger, J., & Beyersmann, E. (2017). Edge-aligned embedded word activation initiates morpho-orthographic segmentation. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (pp. 285–317).
- Grainger, J., & Holcomb, P. J. (2009). Watching the word go by: On the time-course of component processes in visual word recognition. *Language and Linguistics Compass*, 3, 128-156.
- Gwilliams, L. E., Monahan, P. J., & Samuel, A. G. (2015). Sensitivity to morphological composition in spoken word recognition: Evidence from grammatical and lexical identification tasks. *Journal of Experimental Psychology: Learning, Memory and Cognition*.
- Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes*, 5(4), 281-312. doi:10.1080/01690969008407065
- Jared, D., Jouravlev, O., & Joanisse, M. F. (2017). The effect of semantic transparency on the processing of morphologically derived words: Evidence from decision latencies and event-related potentials. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 43, 422-450. doi:10.1037/xlm0000316
- Kim, A., & Lai, V. (2012). Rapid interactions between lexical semantic and word form analysis during word recognition in context: Evidence from ERPs. *Journal of Cognitive Neuroscience*, 24(5), 1104–1112.
- Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition*, 18(5), 655-681. doi:10.1080/13506280902986058
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2014). *lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package)*: R package version 2.0-6. <http://CRAN.R-project.org/package=lmerTest>.

- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.
- Laszlo, S., & Federmeier, K. D. (2009). A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory & Language*, *61*(3), 326–338. doi:10.1016/j.jml.2009.06.004
- Lavric, A., Clapp, A., & Rastle, K. (2007). ERP evidence of morphological analysis from orthography: A masked priming study. *Journal of Cognitive Neuroscience*, *19*, 866–877.
- Lavric, A., Elchlepp, H., & Rastle, K. (2012). Tracking hierarchical processing in morphological decomposition with brain potentials. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 811–816.
- Lavric, A., Rastle, K., & Clapp, A. (2011). What do fully visible primes and brain potentials reveal about morphological processing? *Psychophysiology*, *48*, 676–686.
- Lehtonen, M., Cunillera, T., Rodríguez-Fornells, A., Hultén, A., Tuomainen, J., & Laine, M. (2007). Recognition of morphologically complex words in Finnish: Evidence from event-related potentials. *Brain Research*, *1148*, 123–137.
- Lehtonen, M., Monahan, P. J., & Poeppel, D. (2011). Evidence for early morphological decomposition: Combining masked priming with magnetoencephalography. *Journal of Cognitive Neuroscience*, *23*(11), 3366–3379.
- Leinonen, A., Gronholm-Nyman, P., Jarvenpaa, M., Soderholm, C., Lappi, O., Laine, M., & Krause, C. M. (2009). Neurocognitive processing of auditorily and visually presented inflected words and pseudowords: evidence from a morphologically rich language. *Brain Research*, *1275*, 54–66. doi:10.1016/j.brainres.2009.03.057
- Leminen, A., Leminen, M., & Krause, C. M. (2010). Time course of the neural processing of spoken derived words: an event-related potential study. *NeuroReport*, *21*, 948–952.

- Leminen, A., Leminen, M., Kujala, T., & Shtyrov, Y. (2013). Neural dynamics of inflectional and derivational morphology processing in the human brain. *Cortex*, *49*, 2759-2771.
- Leminen, A., Leminen, M., Lehtonen, M., Nevalainen, P., Ylinen, S., Kimppa, L., . . . Kujala, T. (2011). Spatiotemporal dynamics of the processing of spoken inflected and derived words: A combined EEG and MEG study. *Frontiers in Human Neuroscience*, *5*, 66. doi:10.3389/fnhum.2011.00066
- Lewis, G., Solomyak, O., & Marantz, A. (2011). The neural basis of obligatory decomposition of suffixed words. *Brain and Language*, *118*, 118-127.
- Longtin, C. M., Segui, J., & Hallé, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes*, *18*(3), 313-334.
- Marelli, M., & Amenta, S. (2018). A database of orthography-semantics consistency (OSC) estimates for 15,017 English words. *Behavior Research Methods*, *50*(4), 1482-1495. doi:10.3758/s13428-018-1017-8
- Marelli, M., Amenta, S., & Crepaldi, D. (2015). Semantic transparency in free stems: The effect of orthography-semantics consistency on word recognition. *Quarterly Journal of Experimental Psychology*, *68*(8), 1571-1583. doi:10.1080/17470218.2014.959709
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177-190.
- Marslen-Wilson, W., Tyler, L., Waksler, R., & Older, L. (1994). Morphology and Meaning in the English Mental Lexicon. *Psychological Review*, *101*, 3-33.
- Meunier, F., & Longtin, C. M. (2007). Morphological decomposition and semantic integration in word processing. *Journal of 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 & Language*, *41*, 327-344.
- Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology*, *115*, 2195-2222.

- Morris, J., Frank, T., Grainger, J., & Holcomb, P. J. (2007). Semantic transparency and masked morphological priming: an ERP investigation. *Psychophysiology*, *44*, 506-521.
- Morris, J., Grainger, J., & Holcomb, P. J. (2008). An electrophysiological investigation of early effects of masked morphological priming. *Language and Cognitive Processes*, *23*(7/8), 1021-1056.
- Morris, J., Grainger, J., & Holcomb, P. J. (2013). Tracking the consequences of morpho-orthographic decomposition using ERPs. *Brain Research*, *1529*, 92-104.
- Morris, J., Holcomb, P. J., & Grainger, J. (2008). An electrophysiological investigation of early effects of masked morphological priming. *Language and Cognitive Processes*, *23*(7/8), 1021-1056.
- Morris, J., Porter, J. H., Grainger, J., & Holcomb, P. J. (2011). Effects of lexical status and morphological complexity in masked priming: An ERP study. *Language and Cognitive Processes*, *26*(4-6), 558-599.
- Münte, T. F., Say, T., Clahsen, H., Schiltz, K., & Kutas, M. (1999). Decomposition of morphologically complex words in English: Evidence from event-related brain potentials. *Cognitive Brain Research*, *7*, 241-253.
- New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, *28*, 661-677.
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*, 516-524.
- O'Rourke, T., & Holcomb, P. J. (2002). Electrophysiological evidence for the efficiency of spoken word processing. *Biological Psychology*, *60*, 121-150.
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *1*.

- Perre, L., Midgley, K., & Ziegler, J. C. (2009). When beef primes reef more than leaf: Orthographic information affects phonological priming in spoken word recognition. *Psychophysiology*, *46*(4), 739-746.
- Plaut, D. C., & Gonnerman, L. M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes*, *15*, 445-485.
- Quémart, P., Casalis, S., & Colé, P. (2011). The role of form and meaning in the processing of written morphology: A priming study in French developing readers. *Journal of Experimental Child Psychology*, *109*, 478-496.
- Rastle, K. (2018). The place of morphology in learning to read in English. *Cortex*. doi:<https://doi.org/10.1016/j.cortex.2018.02.008>
- Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language & Cognitive Processes*, *23*, 942-971.
- Rastle, K., Davis, M. H., Marslen-Wilson, W., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes*, *15*(4-5), 507-537.
- Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, *11*, 1090-1098.
- RDevelopmentCoreTeam. (2008). *R: A language and environment for statistical computing*: R package version 1.1-5. <http://CRAN.R-project.org/package=lme4>.
- Royle, P., Drury, J. E., Bourguignon, N., & Steinhauer, K. (2012). The temporal dynamics of inflected word recognition: a masked ERP priming study of French verbs. *Neuropsychologia*, *50*(14), 3542-3553. doi:10.1016/j.neuropsychologia.2012.09.007
- Schmidtke, D., Matsuki, K., & Kuperman, V. (2017). Surviving blind decomposition: A distributional analysis of the time-course of complex word recognition. *Journal of*

doi:10.1037/xlm0000411

- Skrandies, W. (1990). Global field power and topographic similarity. *Brain Topography*, 3(1), 137-141.
- Solomyak, O., & Marantz, A. (2009). Evidence for early morphological decomposition in visual word recognition. *Journal of Cognitive Neuroscience*, 22(9), 2042-2057.
- Solomyak, O., & Marantz, A. (2010). MEG evidence for early morphological decomposition in visual word recognition: A single-trial correlational MEG study. *Journal of Cognitive Neuroscience*, 22, 2042-2057.
- Stockall, L., & Marantz, A. (2006). A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon Journal*, 1(1), 85-123.
- Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14, 638-647.
- Taft, M., Li, S., & Beyersmann, E. (2018). What cross-morphemic letter transposition in derived nonwords tells us about lexical processing. *Journal of Cognition*, 1(1). doi:10.5334/joc.39
- Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: a new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, 67(6), 1176-1190. doi:10.1080/17470218.2013.850521
- Vartiainen, J., Aggujaro, S., Lehtonen, M., Hulten, A., Laine, M., & Salmelin, R. (2009). Neural dynamics of reading morphologically complex words. *Neuroimage*, 47(4), 2064-2072. doi:10.1016/j.neuroimage.2009.06.002
- Velan, H., & Frost, R. (2011). Words with and without internal structure: what determines the nature of orthographic and morphological processing? *Cognition*, 118(2), 141-156. doi:10.1016/j.cognition.2010.11.013

- ACCEPTED MANUSCRIPT
- Winsler, K., Midgley, K., Grainger, J., & Holcomb, P. J. (2018). An electrophysiological megastudy of spoken word recognition. *Language, Cognition and Neuroscience*, 33(8), 1063-1082.
- Wurm, L. (1997). Auditory Processing of Prefixed English Words Is Both Continuous and Decompositional. *Journal of Memory & Language*, 37, 438–461.
- Wurm, L. (2000). Auditory processing of polymorphemic pseudowords. *Journal of Memory and Language*, 42, 255–271.
- Wurm, L., & Ross, S. E. (2001). Conditional root uniqueness points: Psychological validity and perceptual consequences. *Journal of Memory and Language*, 45(1), 39-57.
doi:10.1006/jmla.2000.2758
- ACCEPTED MANUSCRIPT

Appendix A

All variables were extracted from the Lexique database (New et al., 2004). Mean word frequencies are given as Zipf values (log10 occurrences per billion). Standard deviations are shown in parentheses. Freq = frequency; N = neighbourhood size; OLD 20 = Orthographic Levenshtein distance; PLD 20 = Phonological Levenshtein distance; TS = truly suffixed; PS = pseudo-suffixed; NS = non-suffixed. The semantic relatedness proportion between whole words and embedded words was extracted from Latent Semantic Analysis Web facility (<http://lsa.colorado.edu>; Landauer & Dumais, 1997), based on the semantic space 'Francais Total'. The *p*-values of the pair-wise comparisons (t-tests) between item types (TS, PS, and NS) are provided in the final three columns.

Properties	TS	PS	NS	TS vs. NS	PS vs. NS	TS vs. PS
Whole words						
Written word frequency	3.54 (0.64)	3.71 (0.61)	3.59 (0.71)	.730	.365	.183
Subtitle word frequency	2.69 (0.71)	2.82 (0.81)	2.84 (0.77)	.321	.878	.419
Number of letters	7.08 (1.06)	7.10 (1.04)	7.02 (0.91)	.763	.686	.925
Number of phonemes	5.06 (0.76)	5.02 (0.99)	5.12 (0.86)	.716	.595	.823
Number of syllables	2.08 (0.39)	2.06 (0.51)	2.12 (0.33)	.584	.487	.827
Orthographic N	1.88 (1.83)	1.65 (1.67)	1.63 (2.02)	.506	.958	.499
Phonological N	6.18 (6.04)	6.59 (5.02)	4.61 (5.52)	.174	.070	.709
OLD 20	1.98 (0.34)	1.91 (0.26)	2.04 (0.41)	.412	.062	.261
PLD 20	1.57 (0.46)	1.56 (0.40)	1.71 (0.44)	.122	.080	.918
Uniqueness point	4.51 (0.95)	4.65 (0.84)	4.82 (0.99)	.106	.336	.441
Ending length	3.12 (0.77)	3.18 (0.82)	3.16 (0.90)	.813	.909	.708
Embedded words						
Written word frequency	4.25 (0.72)	4.08 (0.79)	4.16 (0.92)	.583	.617	.244
Subtitle word frequency	3.59 (0.75)	3.47 (0.83)	3.46 (0.90)	.422	.954	.436
Number of letters	4.67 (0.77)	4.73 (0.90)	4.53 (0.88)	.403	.268	.722
Number of phonemes	3.25 (0.52)	3.24 (0.71)	3.04 (0.69)	.079	.161	.874
Number of syllables	1.06 (0.24)	1.04 (0.20)	1.10 (0.30)	.466	.244	.650
Orthographic N	8.78 (4.65)	9.31 (5.19)	8.65 (5.56)	.893	.532	.589
Phonological N	17.27 (9.26)	18.90 (9.35)	20.80 (9.01)	.054	.298	.379
OLD 20	1.35 (0.25)	1.32 (0.26)	1.37 (0.29)	.726	.397	.588
PLD 20	1.15 (0.22)	1.11 (0.20)	1.08 (0.19)	.113	.557	.312
Uniqueness point	3.24 (0.74)	3.24 (0.71)	3.04 (0.69)	.169	.161	1.00
Semantic relatedness proportions between whole words and embedded words						
LSA	.240 (.213)	.112 (.134)	.142 (.140)	.008	.291	<.001

Appendix B

Truly suffixed (TS) condition:

	suffixed word	suffixed nonword	non-suffixed nonword	embedded stem
1	polaire	polesse	polige	pôle
2	notaire	notier	notine	note
3	lunaire	lunet	lunepe	lune
4	fêtard	fêtede	fêtin	fête
5	caveau	cavon	cavice	cave
6	barreau	barresse	barrime	barre
7	bandeau	bandard	bandaste	bande
8	fronteau	frontet	frontonne	front
9	plateau	plataire	platisse	plat
10	tourteau	tourteur	tourtan	tourte
11	pruneau	prunon	prunaste	prune
12	dentier	dentesse	dentil	dent
13	sagesse	saget	sagule	sage
14	richesse	richon	richie	riche
15	finesse	finette	finine	fine
16	poulet	pouleau	poulise	poule
17	filet	filasse	filige	fil
18	signet	signesse	signise	signe
19	muret	mureau	muque	mur
20	pochette	pochure	pochique	poche
21	mallette	malleau	mallare	malle
22	noisette	noison	noisou	noix
23	bichette	bichier	bichine	biche
24	boulette	boulion	boulue	boule
25	biquette	biquon	biquotte	bique
26	roulette	roulesse	roulache	roule
27	cuvette	cuveur	cuverge	cuve
28	voilette	voileau	voilouse	voile
29	fillette	filleur	fillare	fille
30	skieur	skiette	skiache	ski
31	longueur	longuier	longune	longue
32	largeur	largette	largine	large
33	portier	portesse	portipe	porte
34	laitier	laitette	laitine	lait
35	fermier	fermeau	fermole	ferme
36	casier	caseau	caseme	case
37	pommier	pommeur	pommenne	pomme
38	palmier	palmon	palmune	palme
39	soumission	soumissette	soumissate	soumis
40	tonneau	tonnette	tonniche	tonne
41	logement	logerie	logereuil	loge

42	peuplement	peuplerie	peuplure	peuple
43	raton	rateur	ratille	rat
44	piéton	piétette	piétotte	pié
45	veston	vestette	vestipe	veste
46	ourson	oursier	ourseille	ours
47	cruchon	crucheur	cruchope	cruche
48	fiston	fistet	fistiot	fil
49	blouson	blouseur	blousure	blouse
50	cordon	cordette	cordope	corde
51	épicerie	épiçement	épiçereuil	épice

Pseudo-suffixed (PS) condition:

	suffixed word	suffixed nonword	non-suffixed nonword	embedded stem
1	calcaire	calqueau	calquice	calque
2	libraire	libreau	librache	libre
3	salaire	salette	salige	salle
4	foulard	foulette	foulose	foule
5	couteau	coutier	coutise	coût
6	biseau	biseur	bisule	bise
7	rameau	ramesse	ramil	rame
8	pinceau	pinceur	pincisse	pince
9	poireau	poiron	poirule	poire
10	moineau	moinette	moinan	moine
11	roseau	rosaire	rosonne	rose
12	paresse	pareau	parine	part
13	prouesse	proument	prougle	proue
14	caresse	carette	carige	car
15	palier	palard	palose	pale
16	parquet	parquesse	parquise	parc
17	fouet	fouure	fouie	fou
18	bolet	bolon	bolie	bol
19	criquet	criqueau	criquie	crique
20	étiquette	étiquon	étiquare	étique
21	carpette	carpon	carpique	carpe
22	bavette	baveau	bavou	bave
23	mouette	mouesse	mouipe	mou
24	cassette	casseau	cassotte	casse
25	fauvette	fauvon	fauvine	fauve
26	coquette	coqueur	coquache	coque
27	vignette	vigneau	vignerge	vigne
28	chouette	chouon	chouole	chou
29	reINETte	reinon	reinue	reine
30	couette	courie	counure	cou
31	valeur	valier	valune	val
32	secteur	sectette	secteche	secte

33	terreur	terrette	terrache	terre
34	sanglier	sanglette	sanglille	sangle
35	gravier	gravette	gravole	grave
36	collier	collesse	collenne	colle
37	sommier	sommette	sommile	somme
38	tablier	tablerie	tablatot	table
39	passion	passette	passaque	passé
40	panneau	pannon	pannestre	panne
41	châtiment	chatirie	chatalot	chat
42	département	départerie	départerope	départ
43	mouflon	mouflaire	mouflaque	moufle
44	dragon	dragueau	draguine	drague
45	flacon	flaquette	flacule	flaque
46	guidon	guidure	guidige	guide
47	pompon	pompesse	pompache	pompe
48	prison	prisette	prisipe	prise
49	grillon	grillette	grillate	grille
50	faucon	fauquette	fauquin	faux
51	batterie	batture	battige	batte

Non-suffixed (NS) condition:

	suffixed word	suffixed nonword	non-suffixed nonword	embedded stem
1	*machine	macheau	machipe	mâche
2	vertige	verteau	vertine	vert
3	*loupot	loupesse	loupise	loup
4	*bourrin	bourresse	bourrige	bourre
5	fourmi	fourment	fournure	four
6	*marquise	marqueau	marquin	marque
7	confit	conment	conreuil	con
8	couronne	couresse	courule	cour
9	rotule	roteau	rotonne	rôt
10	boucan	boucaire	boucare	bouc
11	saucisse	sauceau	saucil	sauce
12	*bougie	bougeau	bougule	bouge
13	sourcil	sourresse	sourcisse	source
14	caprice	caprette	caprou	câpre
15	tournure	touraire	tourmil	tour
16	combat	combesse	combouse	combe
17	*garrot	garette	garure	gare
18	ventouse	ventier	ventise	vent
19	*bêtise	bêtette	bêtache	bête
20	boutique	boutet	boutice	bout
21	tartare	tartet	tartipe	tarte
22	caillou	caillon	cailline	caille
23	principe	princet	prinçan	prince

24	mascotte	mascette	mascerge	masque
25	*massue	massier	massule	masse
26	pistache	pistette	pisteille	piste
27	auberge	aubon	aubine	aube
28	*bassine	bassier	bassique	basse
29	bulletin	bullement	bullenure	bulle
30	corneille	corneau	cornotte	corne
31	fortune	forteur	fortaque	fort
32	*patin	pateur	pateuil	pâte
33	bourgeon	bourjette	bourjenne	bourg
34	*matin	matette	matole	mât
35	gondole	gondon	gondaque	gond
36	baleine	ballesse	balline	balle
37	chalut	chalon	chaline	châle
38	gratin	gratette	gratule	gras
39	pirate	pirion	pirique	pire
40	*routine	roution	routate	route
41	écureuil	écureur	écurosse	écu
42	magister	magiment	magierache	magie
43	lentille	lentier	lentache	lent
44	aveugle	aveument	aveustre	aveu
45	cachalot	cachesse	cachille	cache
46	goulot	goulon	gouluche	goule
47	escalope	escalon	escalune	escale
48	boisson	boisseur	boissule	bois
49	pharaon	pharier	pharille	phare
50	potasse	potette	potile	pot
51	baraque	barrier	barope	bar

Table 1

Table 1: Mean item duration (in ms) for the whole stimulus and the embedded word within each condition.

	TS condition	PS condition	NS condition
Mean duration of whole stimulus			
words	590	578	630
pseudo-suffixed nonwords	618	603	608
non-suffixed nonwords	681	649	691
Mean duration of embedded word			
words	303	280	397
pseudo-suffixed nonwords	296	280	293
non-suffixed nonwords	299	277	298

Table 2

Table 2: Summary of the total number of trials per condition and the average proportion of the trials retained over all 22 participants (n=22) for words, pseudo-suffixed nonwords and non-suffixed nonwords within the truly suffixed (TS), pseudo-suffixed (PS) and non-suffixed (NS) conditions. Standard deviations are presented in parentheses.

	TS condition	PS condition	NS condition
Total Epoch Number (n=22)			
words	817	846	800
pseudo-suffixed nonwords	884	854	894
non-suffixed nonwords	934	926	930
% Epochs Retained (n=22)			
words	70 (11.6)	72 (11.9)	64 (10.2)
pseudo-suffixed nonwords	75 (15.3)	73 (13.1)	76 (12.4)
non-suffixed nonwords	80 (10.7)	79 (11.9)	81 (11.8)
Stimuli examples			
words	<i>pochette</i>	<i>mouette</i>	<i>fortune</i>
pseudo-suffixed nonwords	<i>pochure</i>	<i>mouesse</i>	<i>forteur</i>
non-suffixed nonwords	<i>pochique</i>	<i>mouipe</i>	<i>fortaque</i>

Table 3

Table 3: Mean lexical decision times and error rates for word targets averaged across subjects. Standard deviations are shown in parentheses.

Item type	Reaction times (ms)	Error rates (%)
Truly suffixed	959 (74)	14.1 (6.4)
Pseudo-suffixed	943 (70)	12.4 (8.7)
Non-suffixed	1012 (84)	14.8 (8.9)

Table 4

Table 4: Mean lexical decision times and error rates for nonword targets averaged across subjects. Standard deviations are shown in parentheses.

Item type	Reaction times (ms)	Error rates (%)
Pseudo-suffixed nonwords	1042 (114)	7.0 (7.8)
Non-suffixed nonwords	1034 (95)	2.8 (5.3)

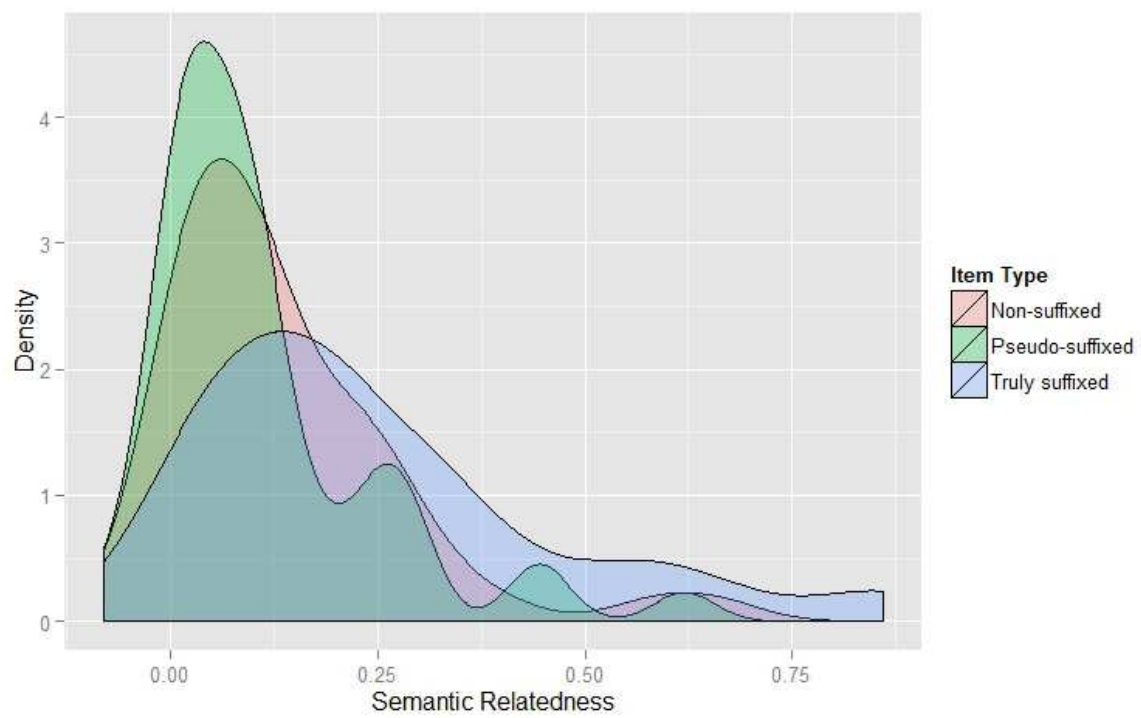
Figure 1

Figure 1: Distribution of semantic relatedness values across Item Types, based on the Latent Semantic Analysis Web facility (Landauer & Dumais, 1997).

Figure 10

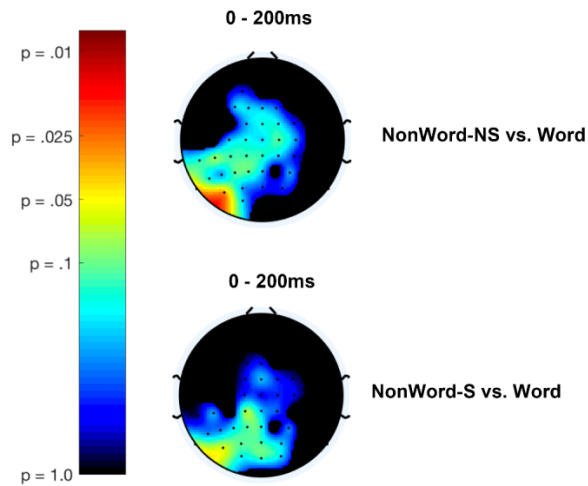


Figure 10. Topographies of the log normalized $(-\log_{10}(p))$ p-values over the 0-200ms time window resulting from the (top) non-suffixed nonword vs. word and (bottom) suffixed nonword vs. word comparisons carried out by permutation test with FDR correction for individual electrodes.

Figure 2

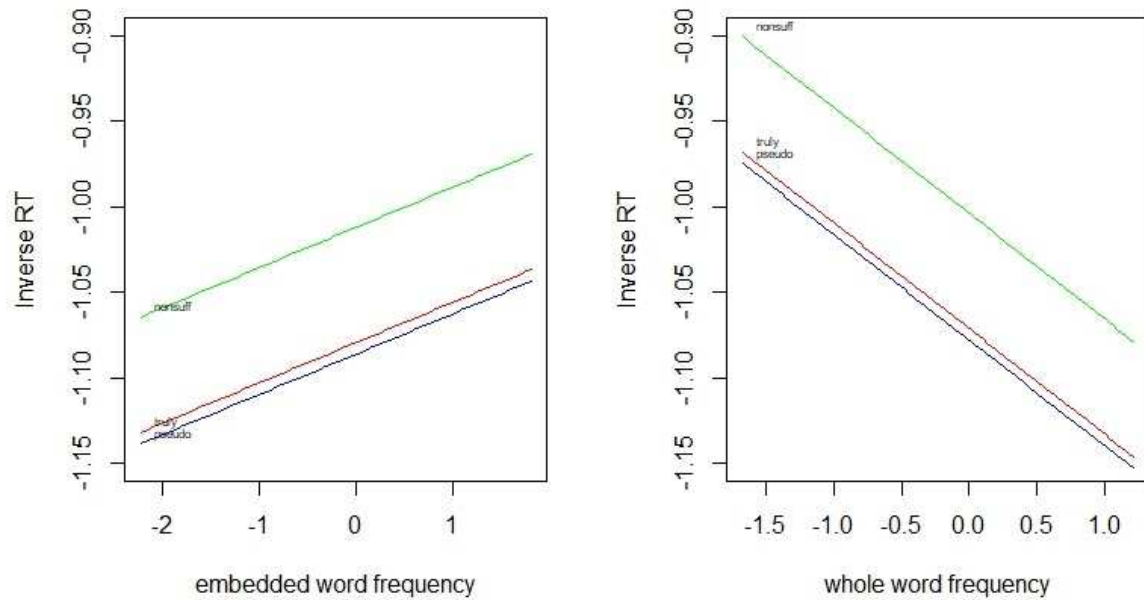


Figure 2: Inverse reaction times (RT) as a function of embedded word frequency (left panel) and whole word frequency (right panel). The non-suffixed condition is displayed in green, the truly suffixed condition in red, and the pseudo-suffixed condition in blue. Frequency measures were subtitle word frequencies extracted from the Lexique database (New, et al., 2004; 2007), transformed into Zipf frequencies (Van Heuven, et al., 2014) and centered to avoid spurious correlations.

Figure 3

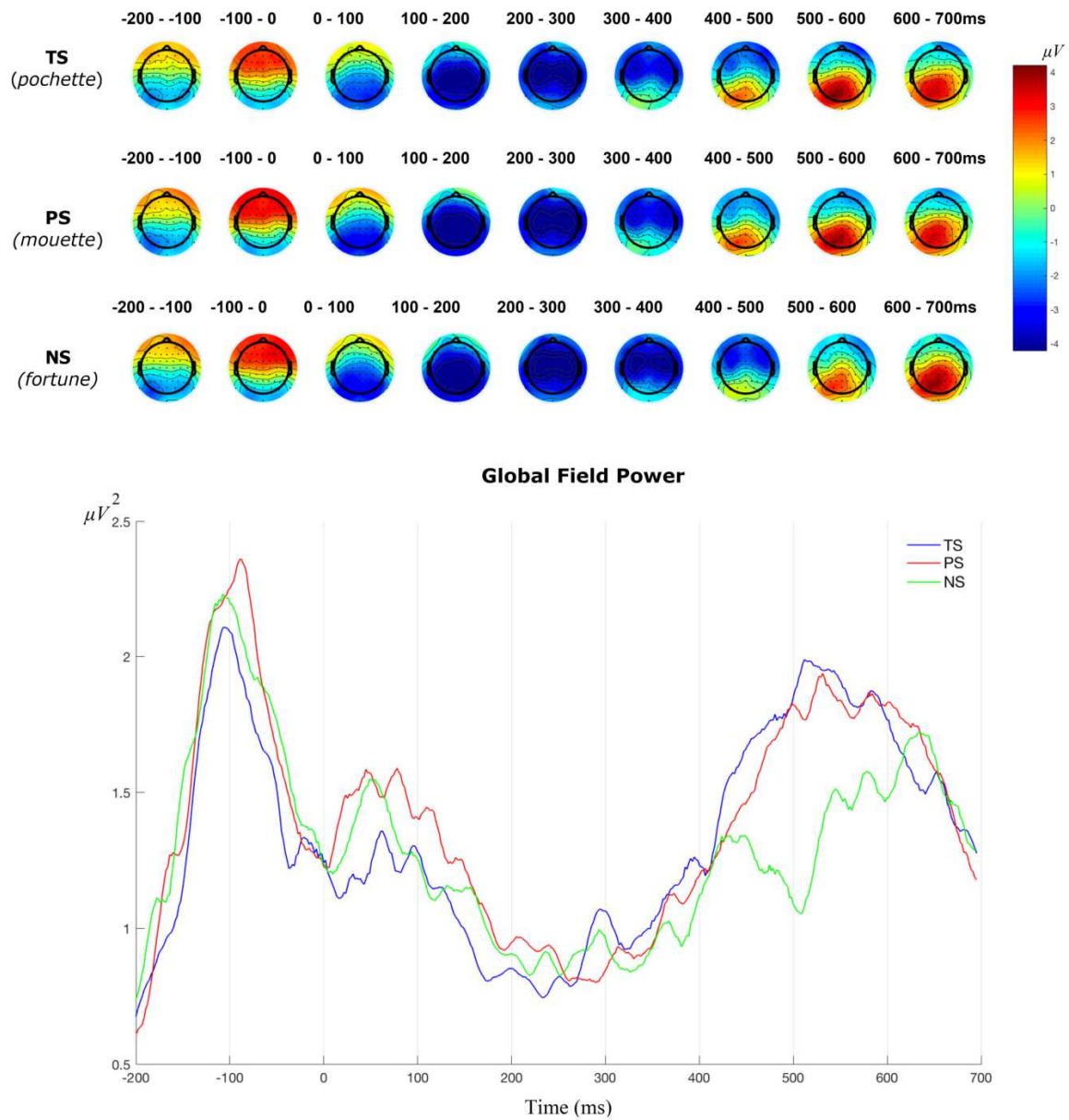


Figure 3: (Top) Scalp maps of the mean activity over 100ms time windows from -200 to 700ms for the truly-suffixed (TS), pseudo-suffixed (PS) and non-suffixed (NS) condition. (Bottom) A plot of the Global Field Power (GFP) of the three word conditions. We can see clearly that, according to the GFP, the NS activity diverges from that of the TS and PS conditions over a time window spanning 400 to 600ms.

Figure 4

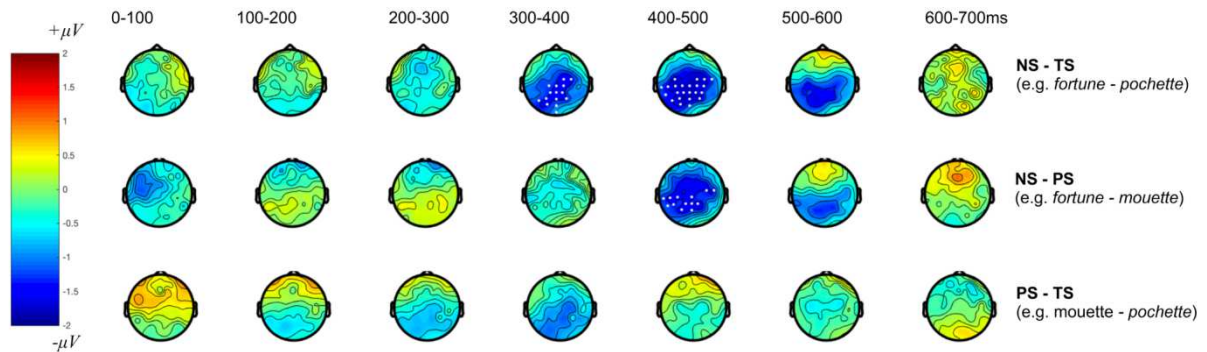


Figure 4: Results of the cluster-corrected permutation test for non-suffixed vs. truly-suffixed words (NS vs. TS), non-suffixed vs. pseudo-suffixed words (NS vs. PS) and pseudo-suffixed vs. truly-suffixed words (PS vs. TS). For all three comparisons, the topographies of the raw effect (NS – TS, NS – TS, PS-TS) are presented over time in 100ms time steps. Those spatio-temporal points presenting statistically significant ($p \leq .025$, two-tailed) differences according to the cluster-corrected permutation test indicated by white dots. The results reveal statistically significant differences over central and parietal electrodes bilaterally in the 300ms to 500ms time window for NS vs. TS and the 400ms to 500ms time window for NS vs. PS. The cluster-corrected permutation test did not reveal any statistical difference for PS vs. TS.

Figure 6

Word - nonword comparisons in the truly-suffixed (TS) condition
(e.g. *pochette* - *pochure* - *pochique*)

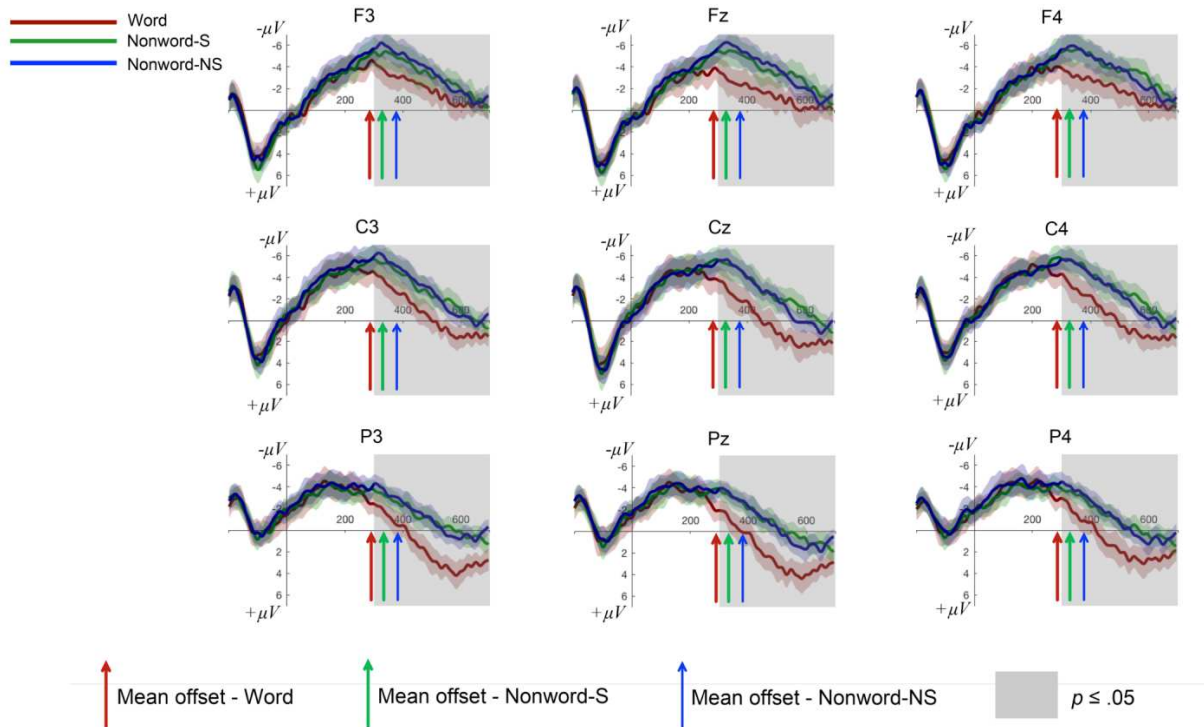


Figure 6: Grand-average ERPs of words (Word), non-suffixed nonwords (Nonword-NS) and suffixed nonwords (Nonword-S) for the truly-suffixed condition. Nine individual electrodes from frontal (F3, Fz, F4), central (C3, Cz, C4) and parietal (P3, Pz, P4) regions are presented and 95% confidence intervals (CI) are shown. For each electrode, time windows presenting a significant difference ($p \leq .05$) between word and both suffixed and non-suffixed nonwords according to a permutation test with FDR correction are highlighted. The mean offset times for words, suffixed nonwords and non-suffixed nonwords are indicated by a red, green and blue arrow, respectively.

Figure 7

Word - nonword comparisons in the pseudo-suffixed (PS) condition
(e.g. *mouette* - *mouesse* - *mouipe*)

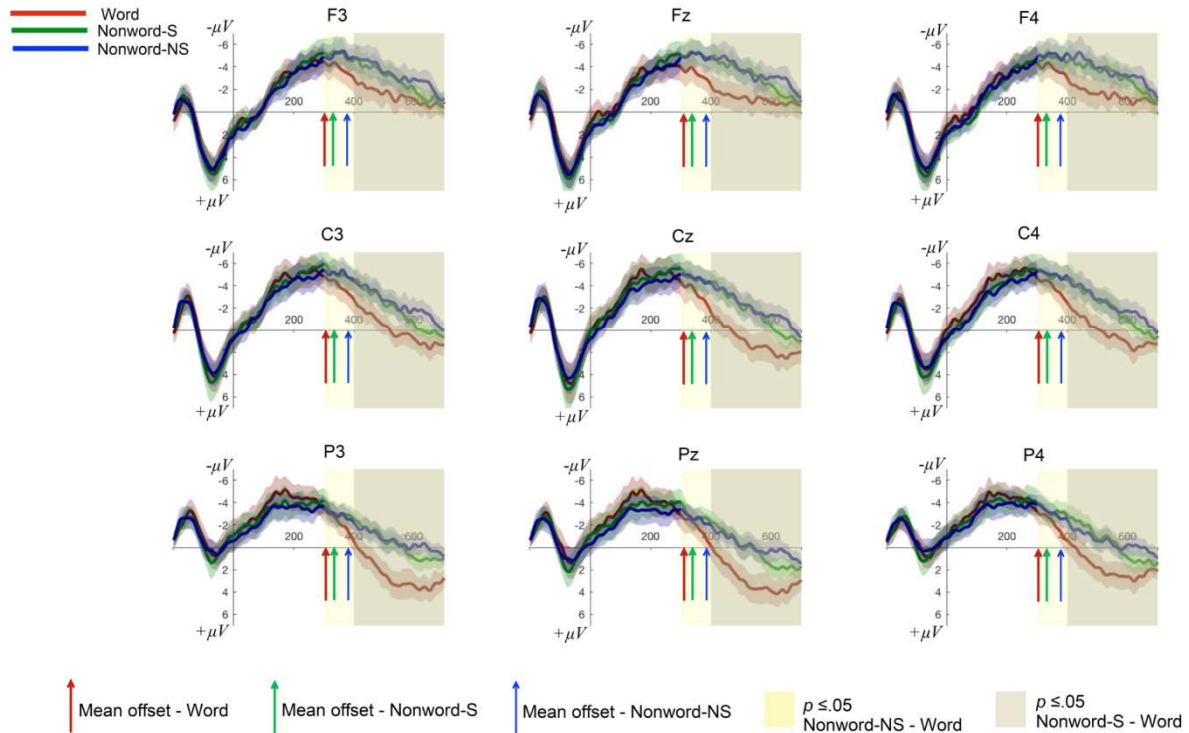


Figure 7: Grand-average ERPs of words (Word), non-suffixed nonwords (Nonword-NS) and suffixed nonwords (Nonword-S) for the pseudo-suffixed condition. Nine individual electrodes from frontal (F3, Fz, F4), central (C3, Cz, C4) and parietal (P3, Pz, P4) regions are presented and 95% confidence intervals (CI) are shown. For each electrode, time windows presenting a significant difference ($p \leq .05$) between word and both suffixed and non-suffixed nonwords according to a permutation test with FDR correction are highlighted. For the non-suffixed nonword vs. word comparison, a significant difference emerges 300ms after the T0 point and continues until 700ms, this is highlighted in yellow. However, for the suffixed-nonword vs. word comparison a significant difference emerges later at 400ms and continues until 700ms; this time interval is highlighted in gray. The mean offset times for words, suffixed nonwords and non-suffixed nonwords are indicated by a red, green and blue arrow, respectively.

Figure 8

Word - nonword comparisons in the non-suffixed (NS) condition
(e.g. *fortune* - *forteur* - *fortaque*)

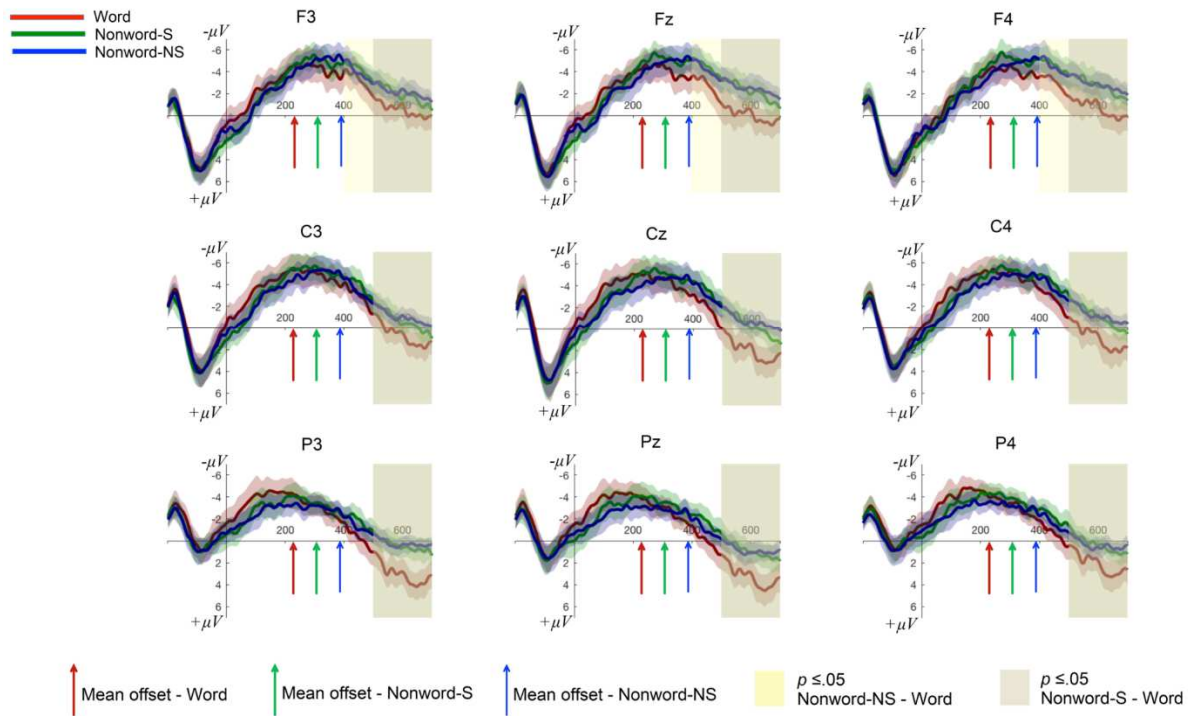


Figure 8: Grand-average ERPs of non-suffixed words (Word), non-suffixed nonwords (Nonword-NS) and suffixed nonwords (Nonword-S). Nine individual electrodes from frontal (F3, Fz, F4), central (C3, Cz, C4) and parietal (P3, Pz, P4) regions are presented and 95% confidence intervals (CI) are shown. For each electrode, time windows presenting a significant difference ($p \leq .05$) between word and both suffixed and non-suffixed nonwords according to a permutation test with FDR correction are highlighted. For the non-suffixed nonword vs. word comparison, a significant difference emerges 400ms after the T0 point over frontal electrodes and continues until 700ms, this is highlighted in yellow. However, for the suffixed-nonword vs. word comparison a significant difference emerges later at 500ms and continues until 700ms; this time interval is highlighted in gray. The mean offset times for words, suffixed nonwords and non-suffixed nonwords are indicated by a red, green and blue arrow, respectively.

Figure 5a

Word: Truly-Suffixed (TS) vs. Non-Suffixed (NS)
(e.g. *pochette* - *fortune*)

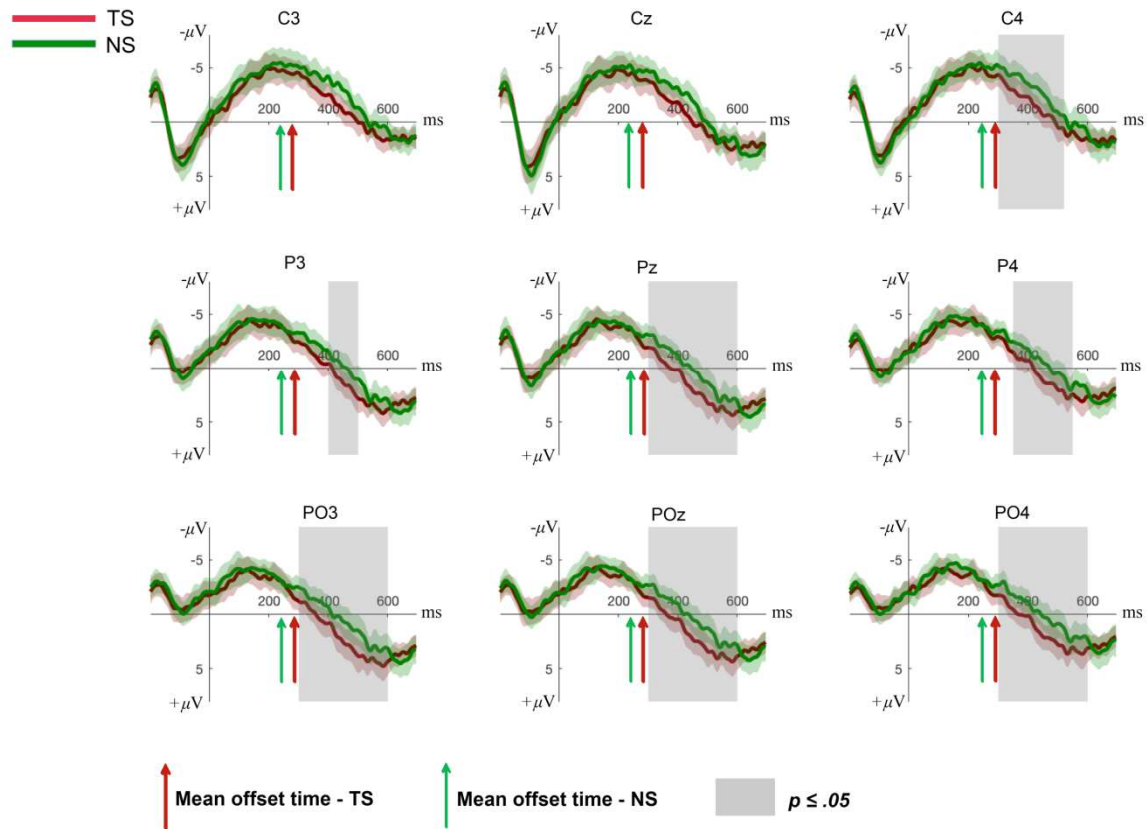


Figure 5a: A comparison of the grand-average ERPs of truly-suffixed and non-suffixed words. Nine individual electrodes from frontal (C3, Cz, C4), central (P3, Pz, P4) and parietal (PO3, POz, PO4) regions are presented and 95% confidence intervals (CI) are shown. For each electrode, time windows (with a minimum duration of 10ms) presenting a significant difference ($p \leq .05$) according to a permutation test with *fdr* correction are highlighted. The mean stimulus offset time of TS and NS words are indicated by a red and green arrows, respectively.

Figure 5b

Word: Pseudo-Suffixed (PS) vs. Non-Suffixed (NS)
(e.g. *mouette* - *fortune*)

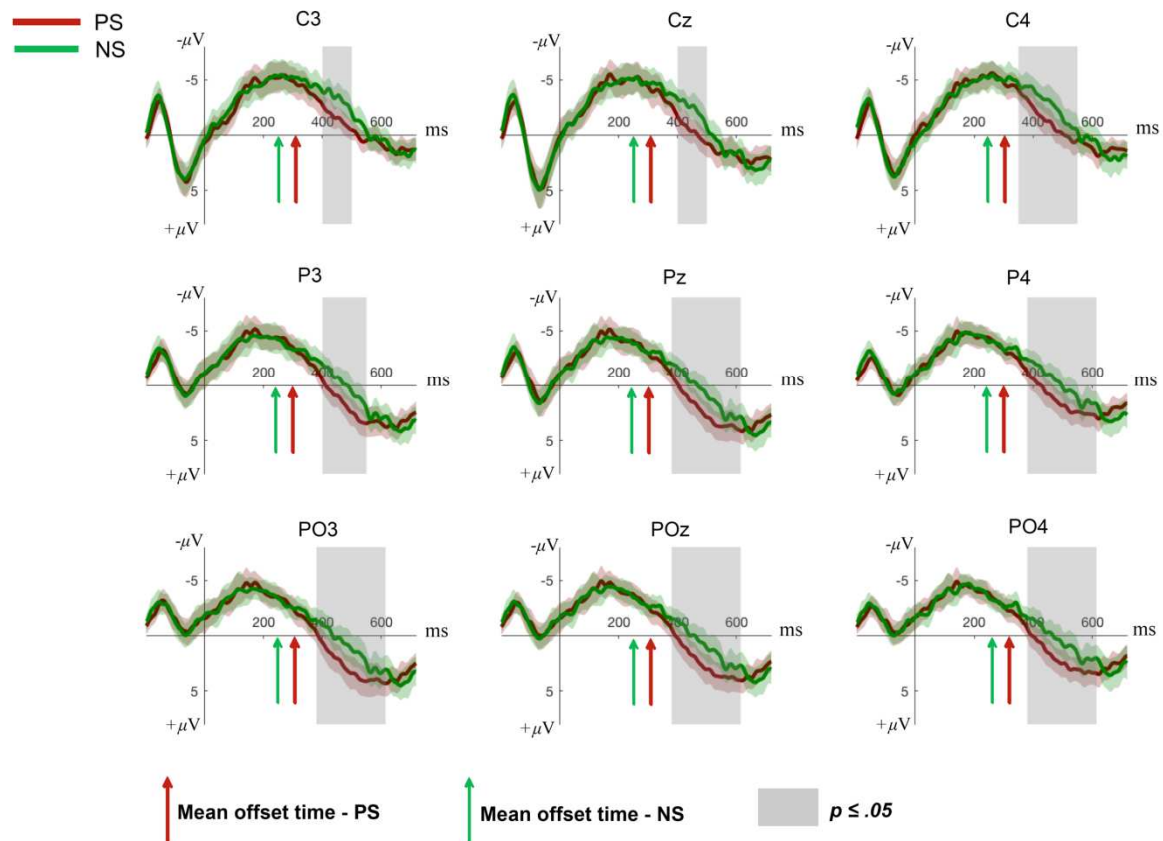


Figure 5b: A comparison of the grand-average ERPs of pseudo-suffixed and non-suffixed words. Nine individual electrodes from frontal (C3, Cz, C4), central (P3, Pz, P4) and parietal (PO3, POz, PO4) regions are presented and 95% confidence intervals (CI) are shown. For each electrode, time windows (with a minimum duration of 10 ms) presenting a significant difference ($p \leq .05$) according to a permutation test with *fdr* correction are highlighted. The mean stimulus offset time of PS and NS words are indicated by a red and green arrows, respectively.

Figure 5c

Word: Truly-Suffixed (TS) vs. Pseudo-Suffixed (PS)
(e.g. *pochette* - *mouette*)

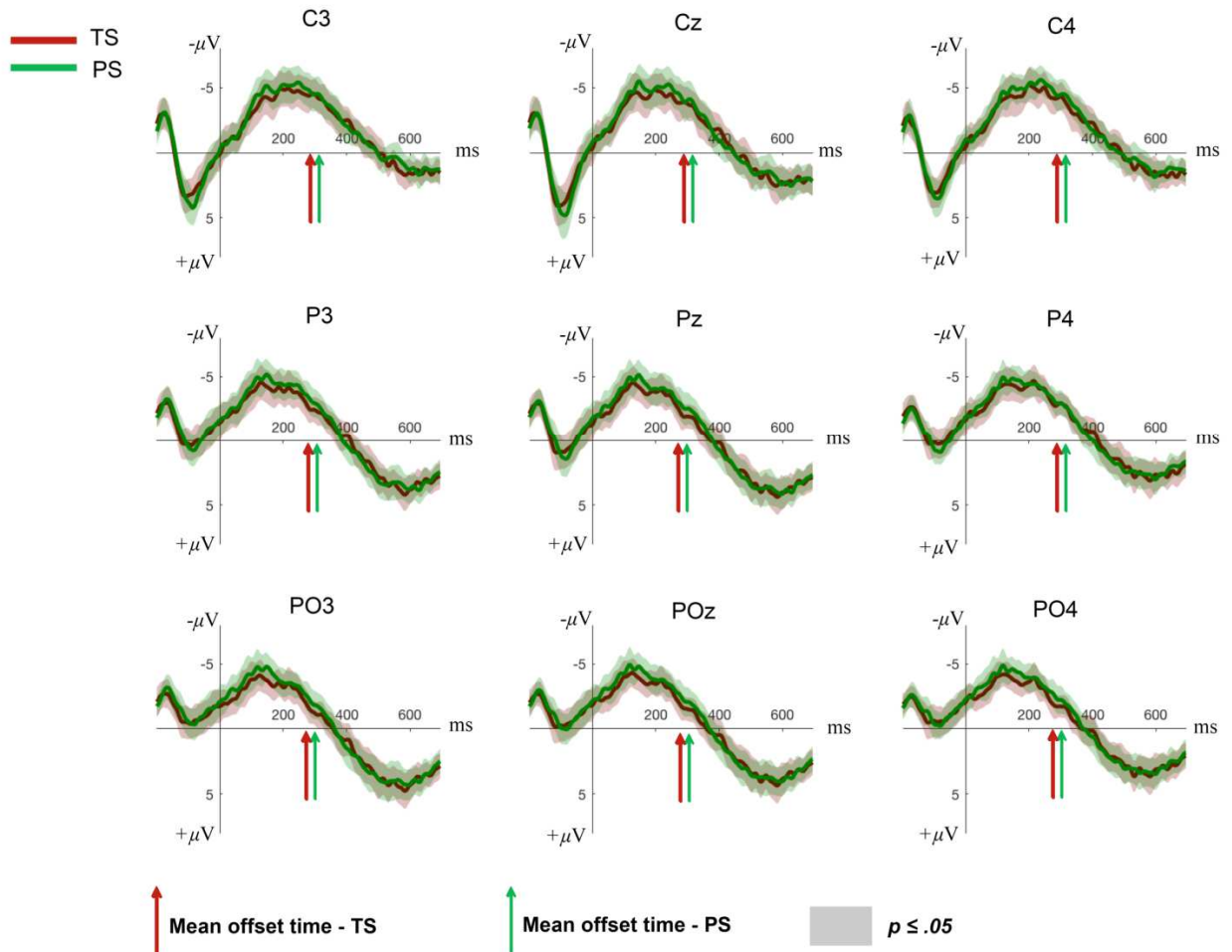


Figure 5c: A comparison of the grand-average ERPs of truly-suffixed and pseudo-suffixed words. Nine individual electrodes from frontal (C3, Cz, C4), central (P3, Pz, P4) and parietal (PO3, POz, PO4) regions are presented and 95% confidence intervals (CI) are shown. For all electrodes, no time-window presents significant ($p \leq .05$) differences according to a permutation test with *fdr* correction. The mean stimulus offset time of TS and PS words are indicated by a red and green arrows, respectively.

Figure 9

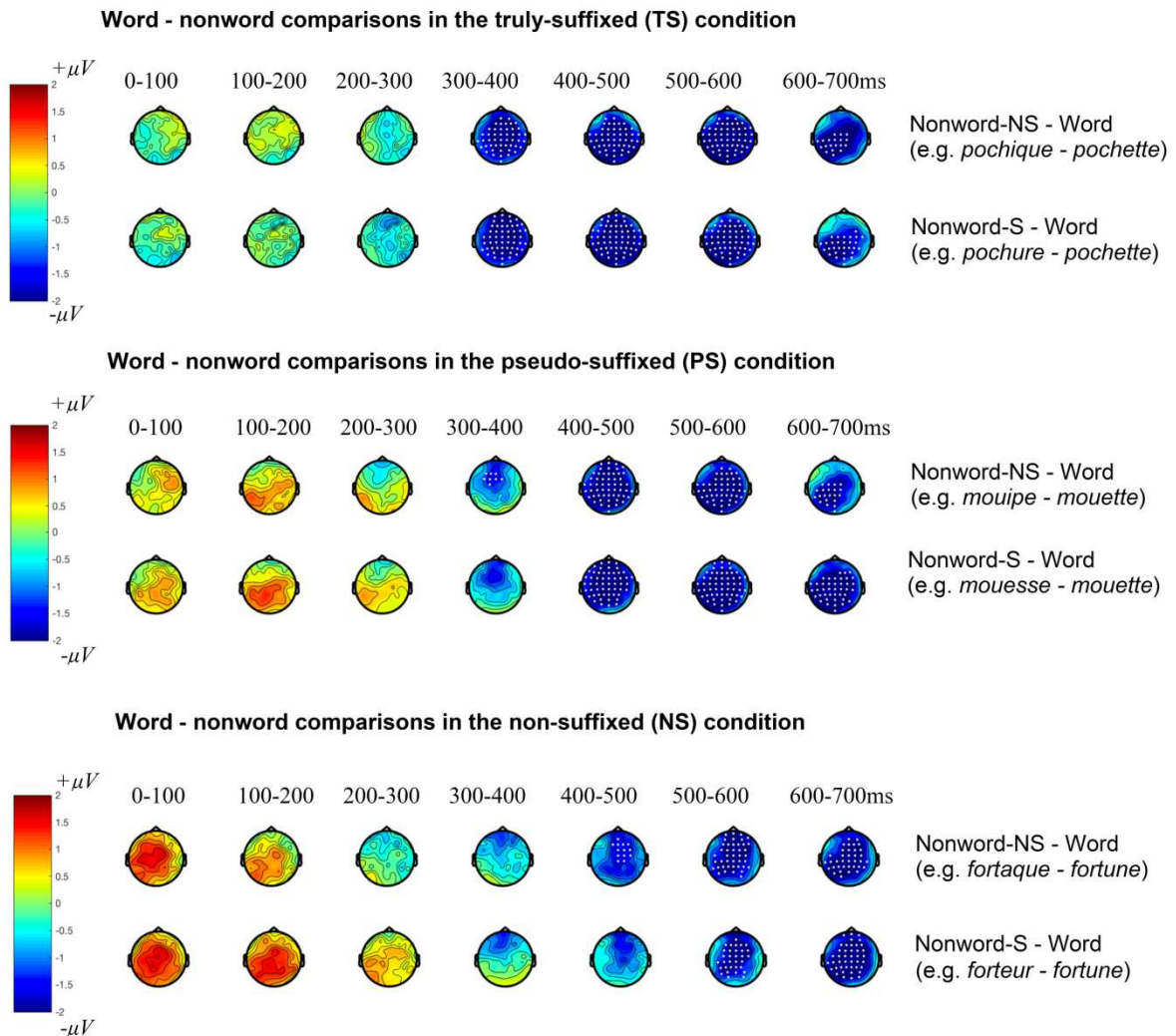


Figure 9: Comparisons of suffixed nonwords, non-suffixed nonwords, and words (TS, PS, NS). For each word condition, a non-suffixed nonword vs. word and suffixed nonword vs. word comparison was carried out by applying a cluster-corrected permutation test over the post-stimulus interval (0-700ms) and over all 64 electrodes. For each comparison, topographies of the raw effect (e.g. non-suffixed nonword – truly suffixed word) are presented as a function of time in 100ms time steps. Those spatio-temporal points presenting statistically significant differences ($p \leq 0.025$, two-tailed) are indicated by white dots.