

DOI: 10.21767/2171-6625.1000271

# Oscillatory Brain Activity in Morphological Parsing of Complex Words: Information Gain from Stems and Suffixes

Sainz LV<sup>1</sup>, Javier S Sainz<sup>2</sup> and Miguel Lázaro<sup>2</sup><sup>1</sup>Department of Neurosurgery, Fundación Jiménez Díaz Hospital, Spain<sup>2</sup>Department of Experimental Psychology: Cognitive Processes, Campus de Somosaguas, Universidad Complutense de Madrid, Spain**\*Corresponding author:** Dr. Laura V. Sainz, Department of Neurosurgery. Fundación Jiménez Díaz Hospital, Avda. Reyes Católicos, 2 28040, Madrid, Spain, Tel: 34676952241, 34916322827; E-mail: laura.sainz@fjd.es**Rec Date:** July 27, 2018; **Acc Date:** October 19, 2018; **Pub Date:** October 22, 2018**Citation:** Sainz LV, Sainz JS, Lazaro M (2018) Oscillatory Brain Activity in Morphological Parsing of Complex Words: Information Gain from Stems and Suffixes. J Neurol Neurosci Vol.9 No.5:271.

## Abstract

In this study, behavioral and EEG measurements were taken while participants performed two priming lexical decision experiments on complex words. In Experiment I stems of high and low family size were used as primes. Behavioral results show an inhibitory effect for stem family size whereas time-frequency responses (TFR) show significant oscillatory brain activity in the range of beta-band and theta-band on right and left temporal sites respectively, both related to lexical status of word patterns. In Experiment II, in which suffixes of high and low family size were used as primes, the effect of family size is facilitatory. Concerning ERP analysis on waveform amplitudes, an early significant lexical status effect emerges although it disappears over time. No significant oscillatory brain activity emerges concerning time frequency responses (TFR). According to an information-gain probabilistic model, the participants modulate their responses in terms of the information provided by the different morphemes used as primes.

**Keywords:** Morphological processing; Stem type frequency; Affix type frequency; Oscillatory brain activity; Theta band; Beta band

**Abbreviations:** EEG: Electroencephalogram; ERP: Event Related Potentials; TFR: Time-Frequency Response

## Introduction

Morphological decomposition of complex words, i.e., words composed of two or more morphemes, remains a controversial issue in Psycholinguistics. Derivation is the more conspicuous expression of word formation in Romance and other European languages and therefore has been the object of much attention. Derivational morphology, unlike inflectional morphology, alters one or several of the properties associated with an item listed in the lexicon (the grammatical category, the conceptual semantics, the number of arguments of a base

– sub-categorization – and the selectional restrictions), and entails the combination of a bound morpheme (prefix or suffix) to a base or stem to create new words. The role of stems and affixes in word recognition is a matter of much discussion and has been studied by psycholinguists during decades through different experimental variables. One of these variables is the so-called family size. Schreuder and Baayen [1] were the first to report an effect by which words with more morphological relatives were responded to faster in a lexical decision task than words with few morphological relatives. The family size effect was initially assessed for stems (how many different words contain a certain stem, including inflectional and derivative words), but it can also be assessed for suffixes (how many different words contain a certain affix). Both are measures of lexical productivity. Family size effects have been observed in many languages such as English, Finnish, French, Dutch, German, Hebrew or Spanish and in both adults and children [2-7].

## Research Methodology

Researchers have studied the family size effect not only through behavioral data, but also through ERP recordings, although with inconsistent results. For example, Kwon, Nam and Lee [8] studied the stem family size in a lexical decision task and failed to find a N400 difference between items of high and low family size. This window is of particular interest given that the family size effect has been hypothesized to be semantic in nature [1,9-11] and this late window is assumed to be sensitive to semantic process. In contrast to the study by Kwon et al. [8], Mulder, Schreuder and Dijkstra [12] in Dutch, reported significant family size effects in an unprimed lexical decision task, with a reduced N400 for words with a high family size.

Comparison of the studies by Kwon et al. [8], and Mulder, et al. [12] is complex given the different writing and morphological systems involved. In the specific case of European languages, the writing system is, in most cases, the same but, importantly, there are very relevant morphological differences that can be of relevance when explaining and understanding different results. Let us compare the case of Spanish (Romance) and some Germanic languages. In Spanish,

Lázaro and Sainz [4] compared words of low and high family size, the number of words associated with a given word, with means of 8.8 and 2.3 respectively while Verhoeven L and Carlisle [13] explained that the rich compounding system of Dutch generates family sizes of 500 members for the word *werk* (work) and almost 300 for the word *hoofd* (head) while in English the word “man” has a family of 200. This enormous contrast can be interpreted as that stems differ in their informational content across languages. This enormous difference in the distribution of stems and affixes across languages should play a role in both, the ability of a stem to prime an affix, or the ability of an affix to prime a stem.

The different distribution of stems and affixes across languages is highly relevant as it allows us to explain different results (facilitative stem family size effects for Dutch, English or Finnish but inhibitory for Spanish) and make some new predictions. Specifically, for Spanish, Lázaro and Sainz [4] carried out a study in which family size for stems (exp. 1) and suffixes (exp. 2) was studied with post-masked priming. Complex words and complex pseudo-words were grouped in high and low family size groups and primed by their stems (*trabajo* → *trabajador* [work → worker]) or by their suffixes (*dor* → *trabajador*; *dor* → \**mesador* [er → \**tabler*])- it is worth noting that in Spanish it has been demonstrated that suffixes can prime complex words [14,15]. In their first experiment, high family size stems elicited slower reaction times than low family size stems (an inhibitory effect of the family size) while in the second experiment, the results showed facilitative effects for affix family size i.e., the stem family size effect was found to be inhibitory while the affix family size effect was found to be facilitative. These results are interesting since finally the stem- and suffix-family-size counts for Spanish replicate the previously mentioned difference between the stem family size of Dutch or English vs. Spanish.

In Spanish, there is enormous asymmetry between stem and suffix family size. While the condition of high stem family size reached a mean value of 8.5, the high suffix-family-size value reached 1,775. From this asymmetry, and in accordance with the previous statement, the opposite direction of the stem and suffix family size effect reflects that stems are better predictors (more informative) of lexical identification than affixes. Because there are few candidates compatible with the presentation of stems, we assume by hypothesis that the prime presentation of stems provides readers with a large amount of information about the stimuli to be read. In sharp contrast, there are thousands of candidates compatible with the presentation of suffixes (stems in languages such as English or Dutch). Therefore, according to a probabilistic approach to word identification, the affix prime does not greatly reduce uncertainty in the reader. Stems activate fewer candidates and therefore the probabilities for any possible lexical recognition are much larger than the probability of a certain candidate when a suffix is presented as a prime. Being sensitive to pattern regularity, the human brain draws information from regular patterns, a view entirely compatible with connectionist models [16]. The results of Lázaro and Sainz [4] paved the way towards the rationale of this experimental series: the differences emerging from the manipulation of

stem productivity and affix productivity do not arise from its very linguistic nature but from the fact that they contain different information, that is, information about how each component contributes to reduce uncertainty about word identity. Both stems and affixes are jointly meaningful sub-lexical patterns of words [17]. The current study aims to assess the contribution of stems and affixes used as primes to lexicality judgments of words -congruent or existing stem + affix combinations- and pseudo-words -incongruent or not existing stem + affix combinations- both composed of real stems and affixes. The main purpose is to examine the compatibility of stem + affix concatenations and the contribution of stems and affixes to word recognition depending on whether the target constitutes a legal lexical entry or not. This experimental manipulation has been employed by Longtin and Meunier [18,19]. van Jaarsveld, Coolen and Schreuder [20] observed that novel compounds (i.e., unfamiliar combinations of nouns) took more time to be rejected as words when constituent nouns had a larger family size. The concatenation of real stems and affixes in new derived forms may produce a differential dynamic of morphemic retrieval [21], a phenomenon that can be exploited to study stem + affix compatibility.

According to our hypothesis in this previously outlined experimental series, stems and affixes provide language users with different information about lexical status of a word candidate. Although the actual mechanism remains uncharacterized, a probabilistic approach to morphological processing is in the core of parallel distributed processing models: the informational content of stems and affixes depend on linguistic distribution of their use. Indeed, informational content is not an intrinsic and static property of word constituents but a dynamic one that can change across time. Baayen, Milin, Đurđević, Hendrix and Marelli [22] implicitly take this view when they propose a model of discriminative learning to account for type frequency effects in morphological processing without appealing to the existence of morphemes as independent lexical units. The semantic approach is adopted by Marelli and Baroni [17] both stems and affixes to evaluate semantic compatibility.

## A probabilistic approach to morphological processing

The idea that stems and affixes provide different information to language users fits easily with the theory of information pioneered by Shannon [23]. Shannon's theory provides all of the formal prerequisites for a modern formulation of the classic model of lexical decomposition. By using the theory of information, we do not mean that words are grouped in paradigms or that word constituents are organized separately as independent lexical units. It would be begging the question to state that stems and affixes constitute lexical units with the same lexical status or state that they differ in semantic content before learning about their contribution to the lexical status of a word candidate. Stems and affixes might have a different impact in word recognition due to the different information residual ( $I_R$ ), encompassed by

stems and affixes in the recognition of a complex word. Their actual role could depend on configurational properties of the language at hand, the different distribution of stems and affixes in the lexicon.

Moscoso-del-Prado-Martín, Kostic and Baayen [24] formalized the information residual ( $I_R$ ), of a given word according to its lexical distribution, and stated that response latency in a lexical decision task is inversely proportional to  $I_R$ . When a whole word is presented, lexical decisions depend basically on word frequency of this word ( $I_W$ ) and that of all derived words with the same lexical base ( $I_{WP}$ ). The more frequent a word is, the more easily it is recognized. This mechanism does not necessarily operate in a sequential way. The notion of information residual ( $I_R$ ) can be generalized to morphological decomposition of complex words and pseudo-words.

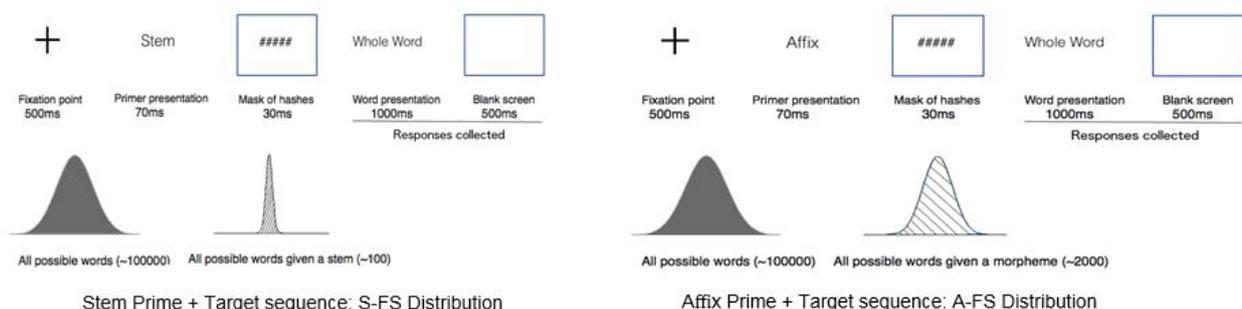
The presentation of a fragment might contribute greatly to the parsing of the target letter string, and more so if the fragment constitutes a lexical subunit of the target. A stem prime (S-FS) reduces uncertainty more than an affix prime (A-FS) by reducing variability of the distribution of possible words. There is an indefinite number of stems almost as many as the number of words in existence. In sharp contrast, affixes constitute a less populous set, the number of its members is finite and fixed, although they are used relatively more frequently in language. This stringent difference between stems and affixes parallels a difference between open-class words and closed-class words, for which Bastiaansen, et al. [25] have found a different oscillatory neuronal dynamic. In general terms, it can be stated that the informational content of a prime  $I_p$  is given by

$$I_p = - \log_2(F(\text{prime})/N) - \log_2(\text{prob}(\text{prime})) \quad (1)$$

Since, according to equation 1,  $\text{prob}(\text{stem})$ , that is, the frequency of stems in language use is relatively smaller than that of affixes,  $I_p$  is greater for stems than for affixes. The presentation of a prime activates a distribution of possible words correctly including that prime. Since, the distribution of primed stems (S-FS) contains fewer items than the distribution of primed affixes (A-FS), the entropy of distribution of possible words containing that prime is smaller for stems than for affixes. If, given a prime, the number of possible words is more recognizable, a pseudo-word can be discarded more easily in a lexical decision task. A stem is a better predictor of lexicality of a target than an affix. The uncertainty  $H_p$  is given by equation 2,

$$H_p = - \sum_{i=1}^n p_i * \log_2[p_i] \quad (2)$$

Where  $p_i = \frac{F_{\text{word}}}{F_{\text{prime}}}$  Fword=frequency of a word in the lexicon, and Fprime=frequency of the prime, obtained by summing across the frequency of all the words containing that prime in the lexicon. Figure 1 illustrates changes in the distribution of stems and affixes when they are used according to the sequences of the experiments presented below. According to Miller [26], our capacity for processing information is at most 2.8 bits, which limits the possible categorization of a stimulus to, on average, 7 equally probable chunks. It is actually being discussed whether working memory span can be even restricted to 3 or 4 chunks [27]. The rationale in both cases is the same: the ability of working memory to keep active a set of informational chunks is limited. This hypothesis fits well with the model of conscious processing of Dehaene and Changeux [28].



**Figure 1** Trial sequences for priming experiments I and II.

We would need a formal theory about how working memory operates according to stimulus complexity through pointers to long-term memory rather than by storing actual items. Having a limited capacity does not mean that working memory is the only cognitive resource in word recognition. The entropy of the distribution of possible words for a stem might be close to Miller's or Cowan's estimates, but although small it is usually greater than working memory capacity. Given a stem, a hundred of possible words can be considered a good

estimate. Therefore, a subject can maintain active the set of possible words weighted according to their frequency in language use, which represents a large part of the entire distribution. And the process could take place in parallel.

Given an affix, a subject cannot maintain more than few possible words, an amount too small and under-representative proportion of the total number of possible words. Extending formalization of Moscoso-del-Prado-Martín et al. [24], the impact of a prime in word recognition of complex words and

pseudo-words can be expressed by equation 3. In this equation the residual information provided by the presentation of the primer is added to the terms of the equation proposed by Moscoso-del-Prado-Martín et al. [24]. The sequential presentation first, of a primer and, then, the whole word containing that primer may result in combined effects as a product of probabilities. Therefore, the log of this expression will involve the sum of residual information of the primer and the entire word. The resulting equation in 3 includes information provided by the primer IP and the entropy of the primer HP jointly equates to the entropy of the distribution of all possible words containing that primer in the lexicon.

$$I_T = \left( I_{RW} - HW \right) + \left( IP - HP \right) \quad (3)$$

$I_{RWord}$                        $I_{RPrime}$

### Looking for neurocognitive properties of stems and affixes

Most of the studies examining word recognition of complex words have collected behavioral measures. As far as we know, there is still no evidence of neural activation to support the contrasting effect of stems and suffixes on word recognition. The formalization of the stem/affix asymmetry according to Shannon's theory [23], and its expression in the preceding equations, allow us to evaluate brain activity in the course of word recognition of complex words. The immediate consequence of this formalization is that, drawing on a post-masked priming design, stems should be better predictors of lexicality of complex words and pseudo-words than affixes. In fact, when a stem is provided as a prime, it would be possible to keep in mind most of the alternative completions of a limited set of legal lexical entries weighted according to their frequency, whereas when an affix is provided as a prime, it simply would not be possible to keep in mind as many possible words as required.

Moreover, if, in the context of these experiments, the system searches for a possible word, the completion of a lexical pattern should be easier after a stem prime (High and Low S-FS) than after an affix prime (High and Low S-FS). In a lexical decision task, in which words and pseudo-words are both composed of real stems and affixes differing in type frequency [4,18], the only differences between existing and non-existing lexical entries would be frequency or novelty and the system's capacity to retain completions of possible words according to the available information, in our terms,  $I_{RPrime}$ .

Although there is no obvious correspondence between brain activity and neurocognitive functions, there is growing body of evidence showing that alpha and beta frequencies of brain potentials mainly reflect quantitative differences between word-patterns according to their lexical status, while theta frequencies reflect qualitative differences mainly concerning brain areas involved in language processing [25,29]. Activity in the beta frequency range has been linked to active inhibition of ongoing processes [30]. Theta oscillations and to some extent gamma oscillations have mostly been associated with storage and retrieval of information from long-

term memory [31,32]. In addition, working memory processes have been related to theta oscillations [33,34]. Some ERP studies have shown consistent differences in word processing between familiar (complex words) and novel words (complex pseudo-words) when time-frequency analyses are conducted: novel words elicit lower power in the theta band (4-8 Hz) than existing words, reflecting lexical access [35] in the left inferior prefrontal and temporal cortex, known to be involved in lexical storage [36].

Although standard event-related potential (ERP) analysis provides information about the ongoing processes in real time, it provides a limited view of oscillatory phenomena because, broadly speaking, the latter are not phase-locked to the event in question and are thus largely cancelled out when single trials are averaged. Despite their complexity, time-frequency analyses can reveal a complex of lexicality and working memory span effects which can be used to examine lexical processing. Such lexical status and working memory span effects can be observed in the theta (4-7 Hz), alpha (8-12 Hz), beta (13-30 Hz), and gamma (30-100 Hz) frequency bands. Bastiaansen et al. [25] have found a theta power increase over left temporal areas for open-class words, but not for closed-class words. The left temporal theta increase may index, according to Bastiaansen et al. [25], the activation of a network involved in retrieving the lexical-semantic properties of the open-class items.

In the present study, we replicate the behavioral experiments conducted by Lázaro and Sainz [4] but, in addition, Event Related Potentials (ERPs) are also taken. ERP analyses in the time domain have shown a N400 component reflecting morphological processing [37]. In Koester and Schiller [38], using morphological priming in overt speech, ERPs were used to test more directly whether morphological priming originates at word form level or whether it has an independent status. These authors reported an onset of the N400 effect that is similar to the estimated onset of morphological encoding proposed by Indefrey and Levelt [39]. Koester and Schiller [38] and, Lensink, Verdonschot and Schiller [40] obtained an increased amplitude value for morphologically primed target words in the latency range 400-550 ms post-stimulus. The use of post-masked priming paradigm in the framework of a lexical decision task allows us to examine the time course of brain activity, examining different ERPs components according to the information provided by a prime. The rationale for post-masking a stimulus pattern is to make sure that stimulus processing does not remain any longer than the time required for the given duration.

Based on our current understanding of oscillatory phenomena, we would expect the pattern of brain activity for complex words to be different from that for complex pseudowords according to stem+affix lexical compatibility, these differences would involve beta and theta frequencies. In addition to these lexicality effects, we expect increases in beta- and theta-band brain activity to accompany the processing of stem-primed lexical candidates. For the impoverished affix-primed lexical candidates, this should be much less obvious, in

accordance with the different informational content of stems and affixes. Specifically, according to the model presented above, subjects might keep trace of a word after a stem prime (S-FS) but fail to do so after an affix prime (A-FS).

This difference should emerge across time in the frequency analysis, as expressed by an oscillatory theta and beta brain activity for stem-primed targets but not for affix-primed targets. A different oscillatory brain activity should be obtained between legal and illegal lexical entries only in the case that stems are used as primers in a lexical decision task. Theta band frequencies should show significant inhibitory effects for pseudo-words than for words in left temporal sites, whereas beta band frequencies should show significant inhibitory effects for pseudo-words than for words in right temporal sites. When used as primes, stems, unlike suffixes, activate brain areas committed to lexical processing and working memory to keep lexical candidates active so as to compare them with targets.

## General Method

### Participants

All participants were Spanish speakers enrolled in undergraduate courses and took part in the experiment in exchange for credits. They were all students from the Complutense University of Madrid, had normal or corrected-to-normal vision, and no documented history of reading difficulties or neural/psychiatric disorders. No student had participated in any other experiment of similar features nor had experience with this experimental design. All were right-handed according to the Edinburgh Handedness Inventory [41]. The study was carried out in accordance with the Declaration of Helsinki and approved by the ethics committee of the Psycholinguistics Research Lab. All participants gave informed consent before starting the experiments. Participants were informed of the aims of this study after the experiment. All the students were assigned either to the first or the second experiment randomly in order to minimize sample effects.

### Stimulus material and procedures

The actual stimuli are described as the relevant experiment is being presented. Stimulus patterns were presented white-on-black centered on an LCD screen, controlled by SuperLab 4.0 software [42]. In a quiet room, laptop screen was located 65 cm from participants' eyes. In the center of the screen, with visual angles around  $0.8^{\circ}$ - $4^{\circ}$  in width according to subject-screen distance, subjects were serially presented, in a sequence, a fixation point "+" for 500ms to focus attention on the point where the stimulus was to appear, and then a post-masked word-string prime for 70ms. According to evidence of word visual processing, this prime duration guarantees pre-conscious sub-lexical processing [14,43].

Although the debate about optimal measures of conscious perception continues, it is important to acknowledge that objective assessments, wagering indices and subjective reports are generally in excellent agreement. In visual masking, the

conscious perception thresholds derived from objective and subjective data are essentially identical across subjects ( $r^2=0.96$ , slope  $\approx 1$ ) [44,45]. Empirical evidence indicates that selection can occur without conscious processing [46]. Conscious processing occurs only if it achieves not only differentiation (i.e., the isolation of one specific content out of a vast repertoire of potential internal representations) but also integration (i.e., the formation of a single, coherent, and unified representation, where the whole carries more information than each part alone) [47].

The minimum duration required for preconscious stimulus processing actually depends on the task being performed, stimulus modality and attentional blink [28]. A basic distinction is whether the non-conscious stimulus is subliminal or preconscious. A subliminal stimulus is one in which the bottom-up, stimulus driven information is so reduced as to make it undetectable, even with focused attention. A preconscious stimulus, by contrast, is one that is potentially visible, but which, on a given trial, is not consciously perceived due to temporary distraction or inattention. Pockett [48,49] suggests that while as much as 500 ms may be required if complicated judgements are being made concerning the data, in other cases stimuli can produce basic sensations in as little as 50-80 ms. This is broadly in line with Efron [50], who estimates that a minimum of 60-70 ms of neural processing time is required for simple auditory and visual stimuli reaching the brain to result in experience.

As in Marslen-Wilson, et al. [19], the prime was either the stem or the affix of the target for both words and pseudo-words. Following the prime, a mask of hashes (#) was presented for 30 ms, then the target remained on the screen for 1 second or until subjects responded. Once the target disappeared, a blank screen was shown for 500 ms. Responses were also collected in this window. Readers were asked to make a lexical decision while their brain activity was recorded using a 64-electrode acti-CAP and a BrainVision Recording System. Response times (RT's), Error Rates (ER) and EEG measurements were taken as dependent measures. RT's were registered from the word onset until subjects made a response. Participants were instructed to judge as quickly as possible whether the letter strings presented were existing words or not, avoiding errors. The session started with a training period, where a sample of 10 representative trials was presented [51].

### EEG recordings and TFR analyses

Recordings were performed in an electrically shielded chamber with voluntary healthy subjects (18-26 years old). Participants were seated comfortably in a chair and were asked to relax and remain still while they performed a visual lexical decision task. The EEG data were recorded continuously using 64 Ag/AgCl electrodes placed in an acti-CAP -for which each electrode includes an integrated circuit for noise filtering- with 2 Brainamp amplifiers, according to the extended 10/20 International System (51) plus de right mastoid, all of them referenced to the left mastoid. EEG recordings were taken for all the active electrodes, although, for the purpose of these

analyses, all, excluding the central ones, are grouped according to brain topography in regions of interest (ROIs: Left Frontal [LF]; Right Frontal [RF]; Left Central [LC], Right Central [RC], Left Temporal [LT]; Right Temporal [RT]; Left Parietal [LP]; Right Parietal [RP]; Left Occipital [LO]; Right Occipital [RO]). Bipolar horizontal and vertical electro-oculograms (EOG) were monitored bipolarly for artifact monitoring via sub- and supraorbital electrodes and left and right external canthal montages, respectively. The ERP results (and analyses thereof on EEG signals) for the experiments being presented have been recorded in two different groups of participants, but by using the same equipment and following the same procedures.

All electrodes were referenced online to the left mastoid and re-calculated offline to averaged-mastoids reference [52], which has been demonstrated as the best way of obtaining word recognition [53]. The Ag/AgCl electrodes were attached to the scalp with adhesive cream in order to keep the electrode resistance below 3 k. An electrode placed on the subject's forehead acted as earth. The bandpass of the amplifiers was set from 0.05 Hz to 35 Hz. Eye movements were recorded through electrooculogram (EOG) and recordings with EOG higher than 70 V were rejected. All signals were digitized on-line with a sample frequency of 250 Hz so that for signals in the frequency range 0.05-35 Hz, the Nyquist-Shannon theorem is over satisfied. Since noise (signals that are not EEG/ERP) is considered to be a random process with zero mean value, the Signal Noise Ratio (SNR) of EEG/ERP was enhanced by averaging across all the valid trials of the experiment.

EEG preprocessing and analyses were conducted using software Analyzer 2.0. The continuous EEG record was divided into 1500 ms epochs, beginning 500 ms before the onset of the word target pattern, i.e., the epochs were time locked to the onset of word target, establishing a baseline starting at -500 ms. The time epoch extended from 500 ms prior to the probe onset, until the end of the response interval. The method described by Gratton, Coles and Donchin [54] was used to correct vertical blinks and horizontal eye movements. The artifacts (eye blinks, horizontal and vertical eye movements, muscle artifacts, etc.) were semi-automatically rejected offline, by eliminating epochs exceeding 75  $\mu$ V in any of the channels. After a visual inspection, remaining epochs that contained artifacts were eliminated throughout. Electrodes for which the signal became erratic were recomputed by triangulation of the nearby electrodes. No recomputed electrode yielded relevant positions in our main results.

Cross-trial average ERPs elicited by the stimuli were computed for each participant, and then averaged separately across subjects for every combination of Lexical Status (LS), and Prime Type Frequency -Family Size- (S-FS and A-FS) condition, according to the design, for every major brain area, by averaging amplitudes of all electrodes in the relevant area. In order to have an overall estimation of main modulations and their time courses, the statistical analyses were calculated for six consecutive time windows (70, 100, 170, 300, 400, 600) by averaging amplitudes over latency interval peaks (-100, +100 when possible), starting at stimulus onset and lasting

until the end of the epoch once DC and artifacts were removed. These windows appropriately fit the main findings, according to visual inspection of the data, while permitting us exploring the whole epoch statistically.

Only RT's and ERPs of correct responses less than 2 standard deviations from the average were used in analyses. Where appropriate, the Greenhouse- Geiser correction for violations of the sphericity assumption was applied and post hoc tests were Bonferroni-corrected. Following Bastiaansen, Oostenveld, Jensen and Hagoort [55], we computed the TFR of the ERP between 0.05 and 35 Hz and applied the same statistical analyses as done on the ERPs to investigate whether any time-frequency effects were driven by event-related brain activity. Decisions on the variables -baseline and latency interval peaks- are very conservative for the purposes of the analyses in the time domain. Overall, the mean rejection rate was 21.52% of the epochs and at least 34 trials were able to be analyzed for every condition.

To isolate the induced-type oscillations from the ERP components, we computed, for every subject, the average ERP, and extract oscillatory brain activity in the range of theta and beta frequency bands. Theta and beta frequency bands, which show distinct response patterns and reflect different cognitive processes [32,56] were computed not by averaging the entire frequency ranges but rather by selecting narrower bands based on the average Time-Frequency Representations (TFRs) of power, by computing the result of the convolution of complex Morlet wavelets with the EEG data, as described by Tallon-Baudry, Bertrand, Delpuech and Permier [57]. Time-Frequency Representations were computed for every time-frequency window and region of interest (ROI).

## Wavelet transforms

Wavelet transform was firstly introduced by Grossmann and Morlet [58], as an extension of the Fourier and Gabor transform. In a similar way to the Gabor transform, the wavelet transform uses a time window of varying width, wide for low frequencies and narrow for high frequencies. As a result, the time-frequency resolution is high and accurate for all frequencies revealing the time evolution of frequencies in the analyzed signal [59]. EEG responses were extracted separately for legal and illegal stem- + -affix combinations. A wavelet-based time-frequency analysis [57] was used to quantify event-related amplitude changes of EEG oscillations. The latency of the effects cannot be meaningfully related to the timing of the underlying neural processes. Although different types of measurements are difficult to relate, the appearance of quantitative and qualitative differences in beta band and theta band oscillatory neural activity may allow us to identify significant differences in the way morphological analysis is conducted for complex word patterns according to the lexical compatibility of their formatives, stems and affixes.

Several studies have been published highlighting adequacy of wavelets on ERP processing and explication. In this study, a continuous wavelet transform with a Morlet real mother wavelet was used. For wavelet normalization, an instantaneous amplitude normalization (Gabor normalization)

was used, taking logarithmic steps for computing real values - spectral amplitudes. Each ERP signal was decomposed in 7 levels and the desirable bands were obtained, 16-32 Hz (beta), 8-16 Hz (alpha), 4-8 Hz (theta) and 0-4 Hz (delta). Note that because of the filter settings (0.05 to 35 Hz) used in the EEG recordings, we were not able to analyze activity in the gamma frequency range. The data padding was set to periodic to cope with the boundary effects. The actual ranges of frequency bands, however, were automatically determined online by the software Analyzer 2.0. For subsequent analyses, the actual theta band brain activity was centered at 6.022 Hz within the range of a Gaussian filter (3.619-8.445 Hz); the actual beta band brain activity was centered at 18.769 within the range of a Gaussian filter (11.261- 26.276 Hz).

This representation of the signal offers the advantage of multi-resolution decomposition. The functions are generated by dilation and translation of unique admissible scaling and mother wavelet functions and constitute an orthonormal basis [58]. In signal analysis, scaling functions are considered as low pass filters, and mother functions as high pass filters. According to these decisions, ERP can be decomposed by levels which correspond to the traditional bands of physiological EEG. Each band frequency is analyzed through peak extraction of its component amplitude at different time windows. Therefore, each band component amplitude represents an average of the component amplitudes for each frequency within the range of the band. For example, for the theta band frequency the average moves from 3.619 to 8.445 Hz as a filter centered at 6.022 Hz. Finally, we compute a unique component for that particular band. At desired time windows (N70, N170, P100, P200, P300, P400, P600), the value of the component amplitude is analyzed and compared between experimental conditions to detect and examine relevant differences concerning underlying cognitive processes released by stimulus presentation.

## Experiment I: Stem Priming

In this experiment, complex words and pseudo-words are equally composed of real stems and affixes, with both patterns only differing in lexical compatibility. Lexical candidates are primed by their corresponding stems. Different early lexical compatibility effects are expected for words and pseudo-words. Since, according to our hypothesis, stem primes are better predictors than affix primes, beta- and theta-band brain activity is expected to reflect the processing of stem-primed lexical candidates.

### Method

**Participants:** Eighteen right-handed students, with mean handedness scores of 82.4%, ranging from 50% to 100% (14 women, and 4 men, mean age 23.2 years) took part in this experiment. EEG data of four participants were excluded due to excessive movement artifacts or equipment failures.

**Stimuli:** A set of 72 stimuli was used in this experiment, 18 words and 18 non-words per SFS condition. Every complex word was selected from ESPAL [59]. Illegal concatenations of

real stems and suffixes were used as complex non-words. All the visual stimuli, primes and letter-strings were presented in the center of the screen.

**Design:** A 2 (Lexical Status: Word vs. Pseudo-word) x 2 (Stem Family Size (S-FS): High vs. Low) factorial design was used. Affix Family Size (A-FS), Frequency (F), Neighborhood Density (ND), and Letter-string Length (LL) were controlled. Syllable Length (SL) was also controlled although it correlates with Letter-string Length (LL). **Table 1** shows the descriptive statistics.

**Table 1** Means and Deviations for Stem (S) and Affix (A) Family Size (FS) patterns in Experiment I. (Freq: Frequency; ND: Neighbor Density; LL: Letter-Length).

Stem Prime	S-FS	A-FS	Freq	ND	LL
Mean High-FS	8.85	1308	2.3	1.5	7.8
SD High-FS	1.8	1091	1.86	0.88	1.1
Mean Low-FS	2.35	1572	2.3	1.05	8
SD Low-FS	0.9	1026	2.4	0.6	1.07

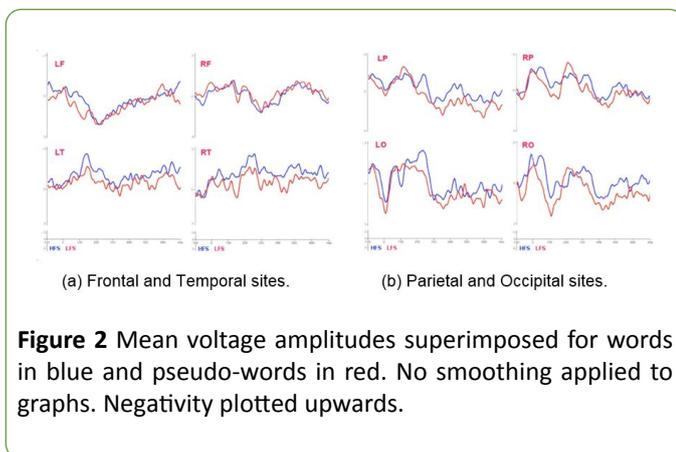
## Results

**Behavioral results:** Overall accuracy in the lexical decision task was high (97.53%) and did not vary across the main experimental conditions. However, a significant interaction between Lexical Status (LS) and Stem Family Size (FS) effects on error rates was found in the analysis on subjects ( $F(1,17)=30.80$ ,  $MSE=2.60$ ,  $p<.001$ , exact observed Power=0.999 (WHFS: 3.22%, W-LFS: 1.56%, PW-HFS: 1.28%, PW-LFS: 3.83%). As stated in data analyses section above, all p-values were Bonferroni-corrected. P-values and degrees of freedom are presented in the text as required.

Regarding the RT analysis, misses and RTs more than 2 standard deviations from the mean were removed (7.8%). Main effects of Lexical Status emerge in the ANOVAs conducted on subjects ( $F(1,17)=42.47$ ,  $MSE=3989.90$ ,  $p<0.001$ , exact observed Power=1.000 (Word [W]: 656 ms vs PseudoWord [PW]: 753 ms)), and on items ( $F(2,140)=87.77$ ,  $MSE=2943.13$ ,  $p<.001$ , exact observed Power=1.000 (W: 656 ms vs PW: 741 ms)). A main Stem Family Size (FS) effect was found to be significant in the ANOVAs on subjects ( $F(1,17)=13.45$ ,  $MSE=528.09$ ,  $p<0.002$ , exact observed Power=0.932, High FS: 715 ms vs. Low FS: 695 ms), and in the ANOVAs on items ( $F(2,140)=4.52$ ,  $MSE=2943.13$ ,  $p<0.05$  High FS: 708 ms vs. Low FS: 689 ms, exact observed Power=0.561). The Lexical Status (LS) x Stem Family Size (FS) interaction is significant, ( $F(1,17)=8.03$ ,  $MSE=803.75$ ,  $p<.05$ ; W-HFS: 676, W-LFS: 637, PW-HFS: 754, PW-LFS: 753, exact observed Power=0.762), ( $F(2,140)=9.41$ ,  $MSE=2943.13$ ,  $p<0.005$ ; W-HFS: 679, W-LFS: 632, PW-HFS: 736, PW-LFS: 745, exact observed Power=0.861). These behavioral data show, high family size stems have an inhibitory effect on lexical decisions

for word letter chains. Lexical decisions are typically taken after the entire word pattern is being presented.

**EEG analyses:** Data at electrodes belonging to a particular ROI were averaged at the individual subject level and used as dependent variables in the statistical analyses. An ANOVA for repeated measures was performed for each time/frequency range, with the factors word string lexical status, stem family size and ROI. ANOVAs conducted on raw mean voltage amplitudes across different time intervals show significant differences according to Lexical Status after stimulus onset, 70 ms: ( $F(1,13)=14.454$ ,  $MSE=3.691$ ,  $p<0.005$ ); 100 ms: ( $F(1,13)=11.191$ ,  $MSE=4.327$ ,  $p<0.01$ ); 170 ms: ( $F(1,13)=5.064$ ,  $MSE=8.647$ ,  $p<0.05$ ); 300 ms: ( $F(1,13)=18.079$ ,  $MSE=5.204$ ,  $p<0.005$ ); and 400 ms: ( $F(1,13)=19.763$ ,  $MSE=8.978$ ,  $p<0.001$ ). **Figure 2** shows ERP waveforms across time and regions of interest. Analyses on ROI averages show consistent lexicality effects across regions in selected time windows.



**Figure 2** Mean voltage amplitudes superimposed for words in blue and pseudo-words in red. No smoothing applied to graphs. Negativity plotted upwards.

**Table 2** Statistic F for averaged amplitudes and mean square error by regions of interest (ROI): ANOVAs for Lexical Status. (Left/Right Temporal, [LT/RT] Right Parietal, Right Occipital [RP/RO]).

ROI	N70	P100	N170	P300	P400	P600
LT <sup>F</sup> (1,13)	9.92	9.9	20.24	9.73	5.78	5.44
LT MSE	1.28	1.1	1.18	1.37	2.03	3.26
RT <sup>F</sup> (1,13)	19.3	17.9	18.5	15	6.3	5.41
RT MSE	0.96	1.01	1.09	1.27	1.9	2.61
RP <sup>F</sup> (1,13)	11.7	12.8	13.1	14	3.74	-
RP MSE	2.18	2.09	2.12	2.16	3.74	-
RO <sup>F</sup> (1,13)	4.84	6.24	6.84	12.7	4.32	-
RO MSE	3.36	2.39	2.33	1.91	4.48	-

Differences between words and pseudo-words are due to lexical compatibility of stem+affix concatenations. **Table 2**

presents the significant results for the relevant windows and **Table 3** the mean voltages in  $\mu\text{V}$  for every region and time window. On average, words involve more inhibitory brain activity than pseudo-words across ROIs and time intervals.

These results replicate the main behavioral effects of lexicality obtained by Lázaro and Sainz [4] and show how word and pseudowords are processed differently at particular regions. As depicted in **Table 3**, stimulus patterns induce more inhibitory brain activity in right hemispheres areas -occipital, parietal and temporal- for word patterns, than for pseudo-word patterns. The same pattern emerges in the only significant left-brain area: the left temporal region.

**Table 3** Means of ERP-Amplitudes for Lexical Status and Regions of interest. (W: Words; PW: Pseudoword. (Left/Right Temporal, [LT/RT] Right Parietal, Right Occipital [RP/RO]).

ROI	N70	P100	N170	P300	P400	P600
LT <sup>F</sup> (1,13)	9.92	9.9	20.24	9.73	5.78	5.44
LT MSE	1.28	1.1	1.18	1.37	2.03	3.26
RT <sup>F</sup> (1,13)	19.3	17.9	18.5	15	6.3	5.41
RT MSE	0.96	1.01	1.09	1.27	1.9	2.61
RP <sup>F</sup> (1,13)	11.7	12.8	13.1	14	3.74	-
RP MSE	2.18	2.09	2.12	2.16	3.74	-
RO <sup>F</sup> (1,13)	4.84	6.24	6.84	12.7	4.32	-
RO MSE	3.36	2.39	2.33	1.91	4.48	-

No apparent effect related to levels of stem family size is obtained. The **Figure 2** displays mean voltage amplitudes for words and pseudowords according to regions of interest. While no relevant differences are obtained involving frontal sites, qualitative differences emerge between words and pseudo-words involving occipital, parietal and temporal regions of the right hemisphere. Interestingly, differences involving the lexical status of a word candidate emerge in both left and right temporal sites.

**Table 4** Statistic F and mean square error for Beta-band-brain-activity: Lexical Status ANOVA. (RT: Right Temporal; W: Word, PW: Pseudoword).

ROI	N70	P100	N170	P300	P400	P600
RT <sup>F</sup> (1,13)	6.48	6.36	5.65	6.35	6.58	6.07
RT MSE	0.26	0.25	0.25	0.28	0.44	0.58

**(a) Beta-band brain activity (13-30 Hz):** A different beta-band brain activity for words and pseudo-words is visible in

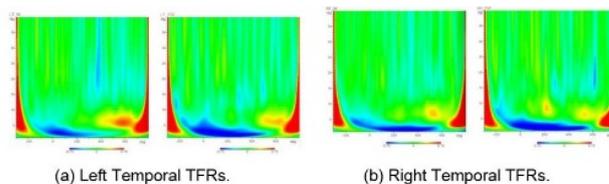
the average TFR peaking across 70-600 time-windows, after stimulus onset. Separate repeated-measures ANOVAs conducted on every time-frequency band shows a lexical status effect on beta-band time frequencies: pseudo-words elicited slightly more inhibition than existing words in the presence of stem-primers, but in specific brain areas. Beta-band TFR is significant in every time-window at right temporal sites (**Table 4** for statistical tests, and **Table 5** for mean TFRs).

**Table 5** Beta-band-brain-activity: Lexical Status Means. (RT: Right Temporal; W: Word, PW: Pseudoword).

ROI	N70	P100	N170	P300	P400	P600
RT W	-0.3	-0.3	-0.32	-0.31	-0.32	-0.28
RT PW	-0.63	-0.63	-0.62	-0.64	-0.75	-0.78

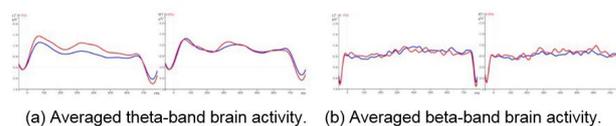
The **Figure 3** displays jointly TFRs of the power changes elicited by words and pseudo-words, for both left and right temporal regions of interest (ROIs).

Beta-band brain activity at left frontal sites fails marginally to reach significance at N70 ( $F(1,13)=3.91$ ,  $MSE=0.203$ ,  $p=0.053$ ), and at N170 ( $F(1,13)=3.09$ ,  $MSE=0.223$ ,  $p=0.084$ ), while it does reach significance at P100 ( $F(1,13)=4.42$ ,  $MSE=0.199$ ,  $p<0.05$ ). Additionally, a beta-band brain activity fails to reach significance at N400 ( $F(1,13)=3.01$ ,  $MSE=0.127$ ,  $p=0.088$ ) for Words in Left Central ROI, the only trace involving a stem family size effect in this analysis, with Stem-FS (HFS: -0.757; LFS: -0.440).

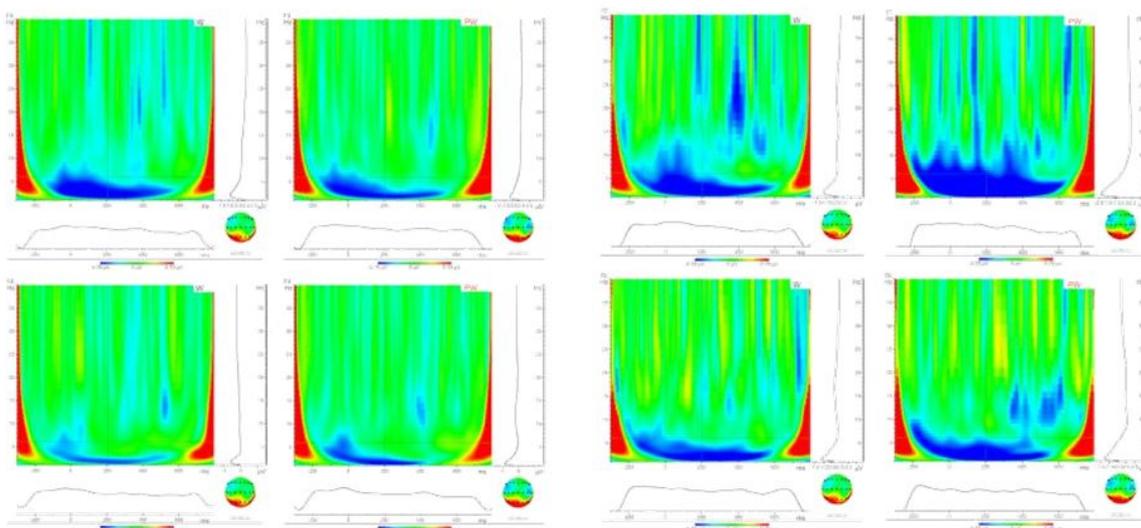


**Figure 3** Time-Frequency Responses (TFRs) of the spectral amplitude changes elicited by words (left) and pseudowords (right) in each figure. TFRs are given for left and right temporal regions of interest (ROIs).

In accordance with **Tables 4 and 5**, we might observe from the **Figure 4** that beta-band brain activity differ for both stimulus patterns according to lexical status, typically different for both hemispheres.



**Figure 4** TFRs of the spectral amplitude's changes elicited by stimulus patterns superimposed for words in blue and pseudo-words in red. Notice the differences between words and pseudowords in theta band in left temporal sites as well the differences in right temporal sites in the beta band that reach significance.



**Figure 5** Time-Frequency Responses and scalp topographies at 200 ms. post stimulus onset of the spectral amplitude changes elicited by words (left) and pseudowords (right). TFRs are given for representative channels (Frontal: F3, F4, and Temporal: T7 and T8), with blue representing spectral amplitude decreases and red depicting spectral amplitude increases for words and pseudowords. Theta band mean (6.022 Hz) brain activity on the horizontal line. Again, left temporal sites present differences between words and pseudowords in the theta band. So as right temporal sites in the beta band.

A closer view can be obtained by looking at representative channels from temporal sites in the **Figure 5**. The **Figure 5** presents TFRs and head mappings for 200 ms in the theta-band brain activity range. Notwithstanding, it is easy to see that the same figure may be applied to capture beta-band brain activity.

In accordance with **Tables 4 and 5** the **Figure 5** shows that while beta-band TFRs are different for words and pseudo-words at temporal right sites, the comparison between words and pseudo-words does not result in significant differences at temporal left sites. The TFRs figures look noisy than data from the tables given limitations of wavelet analyses with BrainVision analyzer software.

**(b) Theta-band brain activity (4-7 Hz):** A different theta-band brain activity for words and pseudo-words is visible in the average TFR peaking across 70-600 time-windows, after stimulus onset. Separate repeated-measures ANOVAs conducted on every time-frequency band shows a lexical status effect on theta-band time frequencies: pseudo-words elicited slightly more inhibition than existing words in the presence of stem-primers, but in specific brain areas. Theta-band TFR is significant in every time-window at left temporal sites (**Table 6** for statistical tests, and **Table 7** for average TFR).

**Table 6** Statistic F and mean square error for Theta-band-brain-activity: Lexical Status ANOVA. (Regions of interest (ROI): LT: Left Temporal; W: Word, PW: Pseudoword).

ROI	N70	P100	N170	P300	P400	P600
LT <sup>F</sup> (1,13)	10.15	10.02	9.31	7.04	4.57	5.25
LT MSE	0.4	0.41	0.42	0.42	0.47	0.56

**Table 7** Averaged theta-band-brain-activity: Lexical status means. (Regions of interest (ROI): LT: Left Temporal; W: Word, PW: Pseudoword).

ROI	N70	P100	N170	P300	P400	P600
LT W	-0.56	-0.55	-0.54	-0.51	-0.54	-0.41
LT PW	-1.07	-1.06	-1.04	-0.95	-0.92	-0.84

Marginally non-significant lexical status effects involve left frontal sites at N70 ( $F(1,13)=3.77$ ,  $MSE=0.360$ ,  $p=0.057$ ,  $P100$  ( $F(1,13)=3.87$ ,  $MSE=0.370$ ,  $p=0.054$ ) and N170 ( $F(1,13)=3.70$ ,  $MSE=0.378$ ,  $p=0.059$ ). No other effects are significant. The **Figure 5** TFRs and head mappings for 200 ms in the theta-band brain activity range. Let us notice, by looking at head mappings in this figure, that while pseudowords seem to activate both left and right temporal sites in the theta-band range, words seem to involve mainly left temporal sites. Theta power has been found to react on a variety of linguistic manipulations. Theta power may at least partially be explained by the large overlap in the neural systems underlying mnemonic processes that relate memory and linguistic processing, the kind of processes that take place in lexical access and word recognition.

## Discussion

The behavioral results of this first experiment fully replicate those obtained by Lázaro, and Sainz [4]. The data obtained show an inhibitory effect of stem Family Size. As in the original paper, these results can be interpreted as a reflection of a lexical competition among all candidates co-activated by the stems presented as primes. Therefore, once different candidates compete for lexical recognition, readers need more time to decide about the lexicality of the items. In this sense, it is relevant to consider that pseudo-words were also composed of real stems and affixes. In the context of words and pseudo-words composed by real morphemes, readers can not benefit from the mere existence of the stem and have to process both stems and suffixes. This process enables the lexical competition we observe in the behavioral results.

Regarding ERP analysis the results show a weak late but non-significant effect for stem-family size while a significant early lexical status effect emerges involving right occipital, right parietal and left- and right temporal sites in every time-window. The lexical status effect of an incoming word target is a too much strong effect for not to emerge as subjects have to decide whether a stem + affix word candidate is a legal lexical entry or not. As a consequence, amplitudes differ for words and pseudo-words composed of real stems and affixes. As already stated, after being presented with a stem prime, participants await a derivational morpheme to complete their lexical decision. Meanwhile, most of the word models compatible with the stem become active. Lexical decisions are taken on the basis of whole-word access but the game is almost closed when a stem is provided as a prime. The extended life of stem-primers is reflected in oscillatory brain activity.

During encoding, beta power was commonly considered responsible for ongoing processing. A consistent and significant beta-band TFR appears to involve Right Temporal sites, while subjects should combine competing cues coming from stems and affixes to decide whether the lexical candidate is a lexical entry or not. Right temporal sites become active when subjects have to assess the linear order of the formatives of a lexical candidate. It has been suggested that beta response is related to a functional role in language processing of visual input [56]. Since a priming paradigm is being used, keeping a lexical pattern in mind might help to assess lexicality to the extent that the prime sufficiently reduces uncertainty about a possible target. A priming paradigm may impose sequential processing of different informational cues, one provided by the prime and the other provided by the presentation of the entire lexical candidate. Hanslmayr and Staudigl [60] and Meeuwissen, Takashima, Fernández and Jensen [61] localized subsequent memory effects in the same range as ours in the lower beta band brain activity to the left inferior frontal gyrus, which is known to be involved in semantic processing.

In contrast to more episodic effects of beta synchronization, theta synchronization has been shown to be associated with memory encoding and retrieval of consolidated patterns, that is, existing lexical entries. The current data show a consistent

and significant theta-band brain activity involving left temporal sites, the brain areas involved in lexical access, lexical retrieval and word identification. Let us notice that beta-band brain activity involves right temporal sites, while theta-band brain activity involves left temporal sites. Theta-band brain activity during encoding is usually higher in response to subsequently recalled items and to correctly recognized old items than to new items during retrieval [35,60,61]. Theta-band TFR has been related to long-term memory traces and lexical access. Pseudo-words have been found to elicit lower theta power than real words [35]. This difference is usually largest in the left inferior prefrontal and temporal cortex, known to be involved in lexical storage [36]. A larger power increase in left temporal theta was also observed when participants read semantically rich open-class words like ours (nouns, verbs, and adjectives) versus words with less semantic content such as determiners and prepositions [25]. In our case, brain activity patterns in the beta-band and theta-band ranges are related to the process of filtering those word candidates that can be always broken down into real stems and affixes. Under this condition, word recognition requires a careful examination of a word candidate. The process requires to examine stem-affix compatibility and the consultation of word models in the lexicon. The task becomes much easier when stems are provided as primes.

Theta synchronization may play a crucial role in the activation of a lexical candidate [25]. Our results indicate that existing lexical entries activate the left temporal site, assessing the lexicality of the incoming word pattern. As a subject is being presented with a stem-prime, lexical candidates are kept in the memory to some extent and these representations, in turn, activate the target those candidates are compared to. As a matter of fact, stems taking an affix are quite few as compared with the number of affixes taking a stem. Our results show significant differences between words and pseudo-words in both theta and beta oscillatory brain activity. These results happen to emerge because stem primes sufficiently reduce uncertainty about which lexical entries are likely. Our hypothesis will be tested again in the following experiment under the same conditions but by using affix prime in a lexical decision task.

## Experiment II: Suffix Priming

In this experiment, which complements the previous one, the same rationale is applied: stems and affixes might serve equally well, in principle, as primes in a masked priming lexical decision task, except for the way both formatives, stems and affixes, concatenate in a complex target, while everything else controlled. The only difference between stem and affix primes concerns linear order and its relative contribution to lexical access according to the gain-information probabilistic model presented above.

In the first experiment, we reported that stems presented for 70 ms activate lexical candidates with those roots. Since a highly similar procedure is followed in this second experiment, we might postulate that suffix primes should prime lexical candidates as stem primes do. However, while between 2 and

13 lexical candidates became active after a stem prime in the first experiment (the lowest and highest Stem-FS of our stimuli), between 50 and 3,100 candidates (lowest and highest Affix-FS) should become active in this second experiment to get parallel results. This enormous number of lexical candidates would certainly overburden the limited capacity of working memory, producing very different results. Thus, we expect a different pattern of results for high-frequency and low-frequency affix family size, no lexicality effects and no trace or reduced beta and/or theta brain activity involving priming of possible lexical candidates.

The rationale for expecting a different pattern of results for stem and affix primes is not related to the semantic content of these word formatives. Suffix primes forming part of lexical candidates are poor predictors of lexicality. Because suffixes generate poor lexical representations of real lexical entries -or too many lexical entries to manage them properly, we do not expect lexical compatibility effects. We do expect, however, negligible beta- and theta-band brain activity to accompany the processing of affix primed lexical candidates. Likewise, latency measurements (RT's) to existing words will be faster than to non-existing lexical candidates. The impoverished nature of suffixes is not an intrinsic property of affixes, but a consequence of the number of appearances in the lexicon, that is, a consequence of how much information is gained from its apprehension. Suffixes are poor predictors of lexical status of a word candidate since they are recognized only to the extent that are not fused with the stems which they are concatenated to.

In sharp contrast with this probabilistic view, a semantic approach would interpret actual suffix effects as a result of semantic content of stems and affixes. In terms of lexicality, in this second experiment, primes are bound morphemes, while in the first experiment they were, theoretically, free morphemes, an unlikely case in Spanish for the stems used since they usually are bound morphemes. Baayen, Feldman and Schreuder [62,63], and Ford et al. [2] have suggested that stems and affixes differ in semantic status, a statement also shared by Pastizzo and Feldman [64]. The reduced semantic load of an affix permits the concatenation of an affix with a diversity of lexical stems. From this perspective, lexical compatibility of both formatives, stems and affixes should play a role while target lexicality is being computed. Then, lexical compatibility should emerge as a difference between words and pseudo-words. Our predictions go beyond this approach.

In our view, differences between stem and affix primes are not specifically due to their semantic content but to their different contributions in reducing uncertainty about possible targets. Therefore, instead of differences in terms of pattern lexicality we expect differences in predictive value of both formatives. Thus, we expect differences between suffixes according to their relative productivity. These differences should emerge in the analysis of brain responses in the time domain, but not in measurements in the time frequency domain, that is, in beta and theta brain activity. It does not make sense to keep traces of multiple lexical candidates across time when an affix is used as a prime. Whether an affix can be

distinguished from a word base depends on its distribution in the lexicon; if an affix becomes fused with the word base, it cannot be distinguished as a predictor of lexicality. In both cases, having an affix does not represent enough information gain for recognizing a word candidate. As a consequence, no difference is expected concerning oscillatory brain activity.

## Methods

**Participants:** Eighteen right-handed students, with mean handedness scores of 87%, ranging from 52 to 100% (15 women, and 3 men, mean age 20.7 years) took part in this experiment.

**Stimuli:** A set of 100 stimuli was used in this experiment: 50 words and 50 non-words, with words selected from the same corpus as before. Half of the stimuli were high Affix-FS and half were low Affix-FS. Illegal concatenations of the same real stems and suffixes were used as complex non-words.

**Design:** A 2 (Lexical Status: Word vs. Pseudo-word) OE 2 (Affix Family Size (S-FS): High vs. Low) factorial design was used. Stem Family Size (Stem-FS), Frequency (F), Neighborhood Density (ND), and Letter-string Length (LL) were controlled. **Table 8** shows the descriptive statistics.

**Table 8** Means and deviations for Stem (S) and Affix (A) Family Size (FS) patterns in Experiment II.

Suffix-Prime	S-FS	A-FS	Freq	ND	LL
Mean High-FS	9.32	1775	6.20	1.84	7.88
SD High-FS	1.62	811.9	4.40	0.90	1.00
Mean Low-FS	9.36	104	5.80	1.68	7.72
SD Low-FS	1.22	74.9	4.90	0.70	1.20

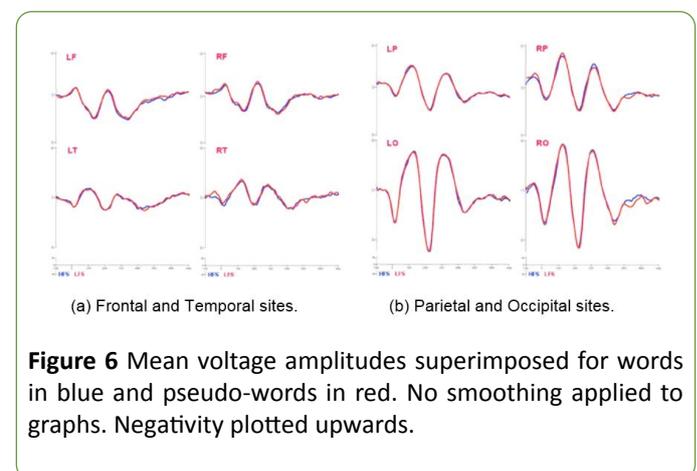
(Freq: Frequency; ND: Neighbor Density; LL: Letter-Length).

## Results

**Behavioral results:** As stated in general method section above, all p-values for both behavioral and EEG results were Bonferroni-corrected. P-values and degrees of freedom are presented in the text as required. Overall accuracy in the lexical decision task was extremely high (98.9%). Error rates did not vary across experimental conditions. Hence, only latency measures are presented. As for the RT analysis, misses and RTs more than 2 standard deviations from the mean were removed (7.8%). The main lexical status effect reached significance in the analyses by participants ( $F(1,17)=123.31$ ,  $MSE=1635.125$ ,  $p<0.001$ , exact observed Power=1.000; Word: 689 ms, Pseudo-word: 796 ms) and by items ( $F(1,97)=2848.25$ ,  $MSE=124.65$ ,  $p<0.001$ ; exact observed Power=1.000, Word: 691ms, Pseudo-word 810 ms). The main effect of Affix-FS reached significance both in the analyses by subjects ( $F(1,17)=22.565$ ,  $MSE=241.978$ ,  $p<0.001$ , exact observed Power=0.994, HFS: 734, LFS: 751) and in the analyses by items  $F(1,97)=97.947$ ,  $MSE=124.65$ ,  $p<0.001$ , exact observed Power=1.000, HFS: 740, LFS: 762. High Affix-FS

stimuli required 22 ms less than Low Affix-FS stimuli (740 ms vs. 762 ms, respectively). No other effect reaches significance.

**EEG results:** ANOVAs conducted on raw mean voltage amplitudes across ROIs show significant differences according to lexical status at 70 ms. ( $F(1,17) = 6.64$ ,  $MSE=0.95$ ,  $p<0.05$ ) and at 300 ms. ( $F(1,17) = 3.405$ ,  $MSE=3.416$ ,  $p<0.05$ ) after stimulus onset. After breaking down mean voltage amplitudes by ROIs, these lexical status effects seem to disappear: despite averaged voltage amplitudes being larger for pseudo-words than for words, no significant main effects of lexical status emerge at any specific site. A significant effect is found involving affix family size at left temporal sites, but the actual effects are only significant at early stages at 70 ms after stimulus onset ( $N70: F(1,17)=4.05$ ,  $MSE=0.690$ ,  $p<0.05$ ,  $0.383 \mu\text{V}$  for high-family-size suffixes and  $0.049 \mu\text{V}$  for low-family-size Suffixes) and at 170 ms after stimulus onset ( $N170: F(1,17)=4.01$ ,  $MSE=0.881$ ,  $p<0.05$ ,  $0.385 \mu\text{V}$  for high-family-size suffixes and  $-0.015 \mu\text{V}$  for affix low-family-size suffixes). Stimulus patterns induce a greater brain activity for high-family-size than for low-family-size suffixes, involving left temporal sites at time intervals N70, N170, but brain activity decreases over time. Separate repeated-measures ANOVAs were conducted on every time-frequency band to examine whether, as in the first experiment, oscillatory brain activity was involved in processing of affix-primed words and pseudo-words. No significant main or interactive effects are found for either beta or theta brain activity. **Figure 6** show ERP waveforms across time and regions of interest.



**Figure 6** Mean voltage amplitudes superimposed for words in blue and pseudo-words in red. No smoothing applied to graphs. Negativity plotted upwards.

## Discussion

The behavioral data replicates once again the results of Lázaro and Sainz [4]. RTs to existing words are faster than to non-existing lexical candidates composed of real stems and affixes. The data not only show a significant effect of Lexical status, but also a significant and facilitatory effect of suffix-family size. As expected, RTs to existing words are faster than to non-existing stimuli, and words composed by suffixes with large morphological relatives are faster responded than words with suffixes of smaller morphological relatives. This is a facilitative effect of Family size, in opposition to the inhibitory effect observed in the first experiment. We interpret that this opposition is explained by the different information that

primes in the first and second experiment provide. In this second experiment the primes are coherent with thousands of possible candidates and therefore, and contrary to what happened in the first experiment, there is no option for a lexical competition. Instead the system benefit from the amount of candidates as found by Bertram, Schreuder and Baayen [65], and Schreuder and Baayen [1] among others. With respect to the ERP analysis, the mean voltage amplitudes in the time domain show early significant effects for Affix-Family Size at Left Temporal sites, and Lexical Status. In this case they are not linked to specific regions and, both effects are in selected time-windows. The information gain obtained from affix processing rapidly declines in such a way that early effects induced by the presentation of a real affix have a short life. As a consequence, no significant oscillatory brain activity is found in the beta and theta bands. Since affixes occur relatively more often in the lexicon than stems, affix primes do not contribute to reduce uncertainty about possible lexical candidates. This finding coincides with that of Kaczer, et al. [21] who suggest that participants may focus more on the separate constituents specifically in the case of non-existing compounds but fail to take them into account when uninformative primes are provided.

## General Discussion

The behavioral results obtained in both experiments completely replicate those found by Lázaro and Sainz [4]. As in their study, in the first experiment the use of stems as primes of pseudo-words made up by existing morphemes created an experimental context in which lexical competition emerges among all possible candidates activated by stem priming. The informational value of the stem priming was great because there were only few possible suffixes that can be lexically concatenated to the stems. In opposition, in the second experiment the information provided by suffixes used as primes is scarce given the huge number of candidates that can be lexically concatenated to them. For this reason, instead of finding again an inhibitory effect of Family Size, we observe a facilitatory effect. This time the information provided by suffixes speeds up the recognition of stimuli, because suffixes can be ignored in order to anticipate an incoming word target. After being presented with a stem, participants await a suffix to complete their lexical decision. Information gain from an affix in this last step is lower than in the previous step when a stem is provided. The role played by stems is reflected in oscillatory brain activity. A number of studies have reported differences in oscillatory responses to coherent, recognizable, meaningful stimuli vs. unrecognizable, incoherent or non-meaningful stimuli. Complex words sound familiar to language users as opposed to pseudo-words composed of stems and affixes in a non-familiar way. Different spectral responses have been observed according to lexical familiarity and composition. Pseudoword patterns were built by randomly combining stems and affixes of well-learned word patterns. Lexical status of a word candidate drives the observed differences in brain responses [66] giving support to the implementation of a simulation model that distinguish

between words and pseudowords according to brain activity [67].

The right temporal beta component represents the ongoing process by which participants examine the possible concatenations of a stem with a limited set of possible affixes. The left temporal theta component represents a qualitative difference between existing and non-existing stem-affix complex patterns. As the theta frequency component is only found after processing a stem-prime, it might be related to retrieving a lexical entry. The stem prime is used to inhibit non-existing lexical candidates. Support for this interpretation arises from a number of studies demonstrating theta brain activity during retrieval of information from long-term memory [31,32]. Moreover, this interpretation fits the results of brain imaging studies of language comprehension that show activations of the left middle/superior temporal areas during the retrieval of a lexical entry [6]. Theta-band brain activity may represent aspects of language processing indicative of lexical access to lexicon. This interpretation is supported by the results obtained in the second experiment.

In the second experiment, the presentation of a suffix as a prime represents a less significant reduction of uncertainty than the presentation of a stem as a prime in the first experiment. Behavioral results show early significant effects of lexical status and morpheme-family size, but in fact no trace of beta- and theta-band brain activity is found in the analyses of oscillatory responses. Given a morpheme, the set of all possible stems is certainly greater than the set of all possible morphemes which implies, in turn, a larger gain in information in the presentation of the whole word when affixes are used as primes. Lexicality cannot be anticipated from the presentation of an affix in the context of words and pseudo-words. Since it is not possible to anticipate any legal stem + affix compounds, no lexical candidates become active, and, therefore, no oscillatory responses are observed. Moreover, an early significant effect of affix family size effect, observed in the ERP amplitude analyses, emerges from the presentation of an affix-prime, but the effect disappears in the course of processing. Behavioral results show facilitatory effects of suffix family size, which are higher for the condition of high family size than for the condition of low family size. The absence of oscillatory brain activity relating left temporal sites across time reveals that affix family size effects are late effects emerging from the presentation of the entire lexical candidate.

The focus of conducting these two experiments is determined the informational gain of being presented with stems or affixes as reflected by oscillatory brain activity, once behavioral measurements have shown a different pattern of responses for stems and affixes in a lexical decision task that takes place after the entire word is being processed, a result congruent with the semantic nature of family size effects.

The experimental results of both experiments are congruent with the equation 3 in a number of aspects. Behavioral results of the previous experiments demonstrate that stem priming produces shorter time responses (W: 656 ms, PW: 753 ms) than affix priming (W: 689 ms, PW, 796 ms). Since time responses are inversely related to residual information

according to Moscoso-del-Prado-Martín et al. [24], this difference is well predicted by equation 3 since the resultant total residual information for stem primers is greater than for affix primers. Under the same rationale, the different results in ERPs analysis for stem and suffix priming are supported by the constraints of equation 3. For stem priming there is no stem-family-size effect. On the contrary, we do encounter affix-family-size effect for affix priming. The difference in entropy values of the distribution of possible words elicited by the stem primers, between high family and low family stems is informationally meaningless from this point of view. Both groups have an entropy value HP less or comparable to the cognitive limit of information processing i.e., Miller's limit. On the other hand, the differences between the entropies of these distributions for affix primers of high and low family size can be substantially greater. In general, entropies HP for affix primers are greater than the Miller's limit but some low family affix primers can have values smaller than this limit. The broader spectrum of entropy values in regard to the cognitive limit in affix priming may be the reason of the early family size effect that is encountered in the second experiment. This line of reasoning is supported by equation 3 because of the different primers being used. Taking stem and affix primers in the edges of their family distribution, we have for example low family size affix "-eja" (as from "moral"-"moraleja", meaning cautionary tale) with  $H_{peja}=1.89577$  and high family size affix "-ero" with  $H_{pero}=4.03267$ . And inversely, the same phenomenon occurs for stem primers that sit at the stem family size distribution edges: low family size stem like "barc-" (stem for "ship")  $HP_{barc}=0.815615$ , and high family size stem "dent-" (stem for "dental"),  $HP_{dent}=2.87067$ . If primer residual information is the variable that decides word processing in priming, this different stem and affix distributions would be a unique continuous spectrum with no qualitative cognitive differences between stems and affixes besides its informational content. Differences in the results of this experimental setting can be attributed to the fact that most of stems behave like low entropy primers in sharp contrast with most of affixes. An apparent discrepancy of significant effects arises between behavioral and ERP measurements. However, it should be taken into account that both kinds of measurements are collected at different temporal paces. Whereas behavioral measurements are offline measurements, once subjects have processed the entire stimulus pattern, ERP measurements are online ones. Family Size effects arise at the time of a lexical decision, not before the stimulus pattern has been processed, as a result of integration processes.

Although it is usually difficult to integrate different measurements under a common framework, the simple information-gain probabilistic model that we present may contribute to solve some of the problems that a complete explanation of morphological parsing should confront. In one hand, the analysis of latency effects come up against many obstacles as they are to be related to oscillatory brain activity. The temporal resolution of time-frequency analysis is inherently poor compared to ERP measures [68,69], with lower frequencies more than with higher frequencies. The use of

sliding windowing wavelet transforms leads to some additional temporal smearing of effects. The latency of the effects cannot be meaningfully related to the timing of the underlying neural processes. On the other hand, as ERP components are concerned, the scalp topographies of frequency components use to be markedly different from the scalp topography of the different ERP components reported in the ERP analysis of the same data [70]. Most ERP studies show different effects for particular manipulations of linguistic materials, and it is difficult to see that a time-frequency analysis produce any effects unique to specific linguistic properties. However, this usual case has no parallel in our research: our finding of qualitative differences in low-frequency -theta band-oscillatory neural activity between stem-primed words and pseudo-words support the notion that lexical compatibility of stem + affix strings are assessed in qualitatively different ways. At this stage alternative explanations cannot be excluded. Notwithstanding, our data are compatible with the notion of a functional role for oscillatory synchrony in binding together different aspects of the linguistic input to obtain a unified and coherent representation of the input entirely compatible with the provisions of our information-gain probabilistic model.

Consistent and significant differences emerge between words and pseudo-words in the first experiment when a stem prime is being provided; no similar effect emerges in the second experiment. Instead a clear suffix family size effect is found, but the effect decreases across time as shown by ERP analyses. Jointly, both experiments show that stems contribute to determine most of lexical candidates in the lexical decision task, but the same is not true for affixes. Oscillatory brain activity analyses provide information about the underlying mechanisms. Affixes fail to elicit a useful set of possible words that represent a significant proportion of the total distribution of possible lexical entries in informational terms. Since, according to our hypothesis, the differential oscillatory brain activity expresses memory retrieval processes, no differences between words and pseudo-words emerge in the time-frequency analysis for affix priming. This correlates with the fact that affixes -by applying equation 3- do not represent any advantage to predict the word lastly presented; thus the lexical decision must be finally made as the entire word becomes available. Considering the limited capacity of working memory for information processing suggested by Miller [52], the values for the expression of HP are consistent with the measured differences between stem and affix priming. For stem primer "obr-" in the word obrero, the HP has a value of 0.605996 bits that is considerably less than the limit 2.8. On the contrary for the affix primer "-ero" a higher value is encountered (4.03267) that surpasses Miller's limit. Again, this endorses our view about the differential informational content of stems and affixes as they are provided as primers, the difference between stems and affixes resting on its respective distribution in the lexicon. In our view, there is no reason for stems and affixes playing a role other than the way these particular units are distributed across the lexicon. Assuming that stems and affixes differ in meaningfulness or lexical status without providing an explanation of the underlying mechanisms involved would be dodging the question, a common fallacy. The proposition that

stems and affixes have different lexical status cannot be taken for granted. We did not assume that stems and affixes are from a different cognitive nature. Stems and affixes can be different because of their distribution in language use. Our results provide some rationale to this statement.

The results obtained in these two experiments may fit well within a lexical model that conceives suffixes as dependent lexical units, since while the stems activate left temporal sites, affixes become active just when whole-word representations become available. In fact, suffixes are not learnt independently from the lexical entries which they form part of such as word and paradigm morphology defends [71], eschewing the morpheme as theoretical construct. It is difficult to see how the recognition of complex words might be achieved by the simultaneous decomposition of lexical candidates into their morphemic units [72]. Someone might suggest that morphological decomposition of lexical candidates could be in our case a priming artifact. However, since demands for words and pseudo-words are the same in both experiments, our results argue against an interpretation that attributing the results obtained to task demands.

## Conclusion

To gain understanding of the nature of language processing in morphological parsing further experiments should be conducted. Deepening in the concept of information processing, effects of using different length letter strings in priming should be elucidated. By using letter chunks not serving as morphological units of complex compounds we could shed light on the role played by morphological parsing – if any – in word recognition, that is whether stems and affixes correlate with or represent actually linguistic and/or cognitive units. An interesting approach in this respect is represented by prefix priming as compared with stem priming. As it has been used in this research, suffix priming could have been an epiphenomenon of sequential reading from left to right if no evidence had been obtained in terms of differential brain activity for stems and affixes. However, our experimental design can be extended to prefix priming. By using prefix priming we can neutralize left-to-right reading effects without changing our equations. Moreover, to unravel whether the adopted probabilistic approach can be applied in other linguistic contexts beyond this experimental design, different primers can be used after sorting these primers according to its residual information, this way allowing us to predict a word target according to the informational content of different word fragments. It would be quite interesting to show discontinuities or phase changes depending on whether this word fragments constitute lexical units or not. However, our findings show that stems and affixes are assessed according to their informational content and this result on the activation of different brain regions and specific time frequency responses. Our results show that stems and affixes play a different role in lexical activation, a role than can be modeled by an information-gain distribution-based model without appealing to the semantic content of the formatives of a complex word pattern.

## Acknowledgements

The authors thank Manuel Martin-Loeches and two anonymous reviewers for their comments on an earlier version of this article.

## Disclosure Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Funding

This research has been conducted exclusively with resources available of the Laboratory of Psycholinguistics in the Psychology Department and with equipment owned by the company Braintools lent freely for that purpose.

## References

1. Schreuder R, Baayen H (1997) How complex simplex words can be. *Journal of Memory and Language* 37: 118-139.
2. Ford MA, Davis MH, Marslen-Wilson WD (2010) Derivational morphology and base morpheme frequency. *Journal of Memory and Language* 63: 117-130.
3. Keller DB, Schultz J (2014) Word formation is aware of morpheme family size. *PLoS One* 9: e93978.
4. Lázaro M, Sainz JS (2012) The effect of family size on Spanish simple and complex words. *Journal of Psycholinguistic Research* 41: 181-193.
5. Moscoso-del-Prado-Martín F, Deutsch A, Frost R, Schreuder R, De Jong NH, et al. (2005) Changing places: A cross-language perspective on frequency and family size in Dutch and Hebrew. *Journal of Memory and Language* 53: 496-512.
6. Nicoladis E, Krott A (2007) Word family size and French-speaking children's segmentation of existing compounds. *Language Learning* 57: 201-228.
7. Pykkänen L, Feintuch S, Hopkins E, Marantz A (2004) Neural correlates of the effects of morphological family frequency and family size: A MEG study. *Cognition* 91: B35-B45.
8. Kwon Y, Nam K, Lee Y (2012) ERP index of the morphological family size effect during word recognition. *Neuropsychologia* 50: 3385-3391.
9. Diependaele K, Duñabeitia JA, Morris J, Keuleers E (2011) Fast morphological effects in first and second language word recognition. *Journal of Memory and Language* 64: 344-358.
10. Feldman LB, O'Connor PA, Moscoso-del-Prado-Martín F (2009) Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form then meaning accounts of word recognition. *Psychonomic Bulletin and Review* 16: 684-691.
11. Feldman LB, Milin P, Cho KW, Moscoso-del-Prado-Martín F, O'Connor PA (2015) Must analysis of meaning follow analysis of form? A time course analysis. *Frontiers in Human Neuroscience* 9: 111.

12. Mulder K, Schreuder R, Dijkstra T (2013) Morphological family size effects in L1 and L2 processing: An electrophysiological study. *Language and Cognitive Processes* 28: 1004-1035.
13. Verhoeven L, Carlsle JF (2006) Introduction to the special issue: Morphology in word identification and word spelling. *Reading and Writing* 19: 643-650.
14. Duñabeitia JA, Perea M, Carreiras M (2007) Do transposed-letter similarity effects occur at a morpheme level? Evidence for morpho-orthographic decomposition. *Cognition* 105: 691-703.
15. Lázaro M, Illera V, Sainz J (2016) The suffix priming effect: Further evidence for an early morpho-orthographic segmentation process independent of its semantic content. *Quarterly Journal of Experimental Psychology* 69: 197-208.
16. Seidenberg MS, McClelland JL (1989) A distributed, developmental model of word recognition and naming. *Psychological Review* 96: 523-568.
17. Marelli M, Baroni M (2015) Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics. *Psychological Review* 122: 485-515.
18. Longtin CM, Meunier F (2005) Morphological decomposition in early visual word processing. *Journal of Memory and Language* 53: 26-41.
19. Marslen-Wilson WD, Bozic M, Randall B (2008) Early decomposition in visual word recognition: Dissociating morphology, form, and meaning. *Language and Cognitive Processes* 23: 394-421.
20. van Jaarsveld HJ, Coolen R, Schreuder R (1994) The role of analogy in the interpretation of novel compounds. *Journal of Psycholinguistic Research* 23: 111-137.
21. Kaczer L, Timmer K, Bavassi L, Schiller NO (2015) Distinct morphological processing of recently learned compound words: An ERP study. *Brain Research* 1629: 309-317.
22. Baayen RH, Milin P, Đurđević DF, Hendrix P, Marelli M (2011) An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review* 118: 438-481.
23. Shannon CE (1948) A mathematical theory of communication. *The Bell System Technical Journal* 379-423.
24. Moscoso-del-Prado-Martín F, Kostić A, Baayen RH (2004) Putting the bits together: An information theoretical perspective on morphological processing. *Cognition* 94: 1-18.
25. Bastiaansen M, Van der Linden M, Ter Keurs M, Dijkstra T, Hagoort P (2005) Theta responses are involved in lexical-semantic retrieval during language processing. *Journal of Cognitive Neuroscience* 17: 530-541.
26. Miller GA (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63: 81.
27. Cowan N (2015) George Miller's magical number of immediate memory in retrospect: Observations on the faltering progression of science. *Psychological Review* 122: 536-541.
28. Dehaene S, Changeux JP (2011) Experimental and theoretical approaches to conscious processing. *Neuron* 70: 200-227.
29. Bakker I, Takashima A, van Hell JG, Janzen G, McQueen JM (2015) Changes in theta and beta oscillations as signatures of novel word consolidation. *Journal of Cognitive Neuroscience* 27: 1286-1297.
30. Pfurtscheller G, Lopes da Silva FH (1999) Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology* 110: 1842-1857.
31. Burgess AP, Ali L (2002) Functional connectivity of gamma EEG activity is modulated at low frequency during conscious recollection. *International Journal of Psychophysiology* 46: 91-100.
32. Klimesch W, Doppelmayr M, Stadler W, Pöhlhuber D, Sauseng P, et al. (2001) Episodic retrieval is reflected by a process specific increase in human electroencephalographic theta activity. *Neuroscience Letters* 302: 49-52.
33. Jensen O, Tesche CD (2002) Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience* 15: 1395-1399.
34. Tesche CD, Karhu J (2000) Theta oscillations index human hippocampal activation during a working memory task. *Proceedings of the National Academy of Sciences USA* 97: 919-924.
35. Krause CM, Grönholm P, Leinonen A, Laine M, Säkkinen AL, et al. (2006) Modality matters: The effects of stimulus modality on the 4- to 30-Hz brain electric oscillations during a lexical decision task. *Brain Research* 1110: 182-192.
36. Marinković K, Rosen BQ, Cox B, Kovacević S (2012) Event-related theta power during lexical-semantic retrieval and decision conflict is modulated by alcohol intoxication: Anatomically constrained MEG. *Frontiers in Psychology* 3: 121.
37. McKinnon R, Allen M, Osterhout L (2003) Morphological decomposition involving non-productive morphemes: ERP evidence. *Neuroreport* 14: 883-886.
38. Koester D, Schiller NO (2008) Morphological priming in overt language production: Electrophysiological evidence from Dutch. *Neuroimage* 42: 1622-1630.
39. Indefrey P, Levelt WJM (2004) The spatial and temporal signatures of word production components. *Cognition* 92: 101-144.
40. Lensink SE, Verdonschot RG, Schiller NO (2014) Morphological priming during language switching: An ERP study. *Frontiers in Human Neuroscience* 8: 995.
41. Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9: 97-113.
42. Stimulus presentation software Superlab 4.0. (2006) Pedro, CA: Cedrus Corporation.
43. Carreiras M, Perea M, Vergara M, Pollatsek A (2009) The time course of orthography and phonology: ERP correlates of masked priming effects in Spanish. *Psychophysiology* 46: 1113-1122.
44. Del Cul A, Dehaene S, Leboyer M (2006) Preserved subliminal processing and impaired conscious access in schizophrenia. *Arch Gen Psychiatry* 63: 1313-1323.
45. Del Cul A, Dehaene S, Reyes P, Bravo E, Slachevsky A (2009) Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain* 132: 2531-2540.
46. Koch C, Tsuchiya N (2007) Attention and consciousness: Two distinct brain processes. *Trends Cogn Sci* 11: 16-22.
47. Tononi G (2008) Consciousness as integrated information: A provisional manifesto. *Biol Bull* 215: 216-242.

48. Pockett S (2002) On subjective back-referral and how long it takes to become conscious of a stimulus: A reinterpretation of Libets data, *Consciousness and Cognition* 8: 45-61.
49. Pockett S (2003) How long is "now", *Phenomenology and the specious present, Phenomenology and the Cognitive Sciences* 2: 55-68.
50. Efron R (1967) The duration of the present. *Annals of the New York Academy of Sciences* 138: 713-729.
51. American Clinical Neurophysiology Society (2006) Guideline 5: Guidelines for standard electrode position nomenclature. *American Journal of Electroneurodiagnostic Technology* 46: 222-225.
52. Lehmann D (1987) Principles of spatial analysis. *Handbook of Electroencephalography and Clinical Neurophysiology* 1: 309-354.
53. Rudell AP, Hua J (1997) The recognition potential, word difficulty, and individual reading ability: On using event-related potentials to study perception. *Journal of Experimental Psychology: Human Perception and Performance* 23: 1170-1195.
54. Gratton G, Coles MG, Donchin E (1983) A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology* 55: 468-484.
55. Bastiaansen M, Oostenveld R, Jensen O, Hagoort P (2008) I see what you mean: Theta power increases are involved in the retrieval of lexical semantic information. *Brain and Language* 106: 15-28.
56. Weiss S, Mueller HM (2012) Too many betas do not spoil the broth? The role of beta brain oscillations in language processing. *Frontiers in Psychology* 3: 201.
57. Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J (1997) Oscillatory gamma-band (30-70 Hz) activity induced by a visual search task in humans. *The Journal of Neuroscience* 17: 722-734.
58. Grossmann A, Morlet J (1984) Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM Journal on Mathematical Analysis* 15: 723-736.
59. Louis PMAK, Rieder A (1997) *Wavelets theory and applications*. John Wiley and Sons, NY, USA.
60. Hanslmayr S, Staudigl T (2014) How brain oscillations form memories? A processing based perspective on oscillatory subsequent memory effects. *Neuroimage* 85: 648-655.
61. Meeuwissen EB, Takashima A, Fernández G, Jensen O (2011) Evidence for human fronto-central gamma activity during long-term memory encoding of word sequences. *PLoS One* 6: e21356.
62. Düzel E, Penny WD, Burgess N (2010) Brain oscillations and memory. *Current Opinion in Neurobiology* 20: 143-149.
63. Baayen RH, Feldman LB, Schreuder R (2006) Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language* 55: 290-313.
64. Pastizzo MJ, Feldman LB (2004) Morphological processing: A comparison between free and bound stem facilitation. *Brain and Language* 90: 31-39.
65. Bertram R, Schreuder R, Baayen H (2000) Effects of family size for complex words. *Journal of Memory and Language* 42: 390-405.
66. Moseley RL, Pulvermüller F (2014) Nouns, verbs, objects, actions, and abstractions: local fMRI activity indexes semantics, not lexical categories. *Brain Lang* 132: 28-42.
67. Garagnani M, Lucchese G, Tomasello R, Wennekers T, Pulvermüller F (2017) A Spiking Neurocomputational Model of High Frequency Oscillatory Brain Responses to Words and Pseudowords. *Frontiers in Computational Neuroscience* 18.
68. Scott SK, Blank CC, Rosen S, Wise RJ (2000) Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123: 2400-2406.
69. Knosche TR, Bastiaansen MC (2002) On the time resolution of event-related desynchronization: A simulation study. *Clinical Neurophysiology* 113: 754-763.
70. Brown CM, Hagoort P, Ter Keurs M (1999) Electrophysiological signatures of visual lexical processing: Open and closed-class words. *Journal of Cognitive Neuroscience* 11: 26-281.
71. Blevins JP (2003) Stems and paradigms. *Language* 79: 737-767.
72. Meunier F, Longtin CM (2007) Morphological decomposition and semantic integration in word processing. *Journal of Memory and Language* 56: 457-471.