ERP Correlates of Inhibitory and Facilitative Effects of Constituent Frequency in Compound Word Reading

Marta Vergara-Martínez[1,3], Jon Andoni Duñabeitia[1] , Itziar Laka[2], & Manuel Carreiras[4,1]

[1] Instituto de Tecnologías Biomédicas, Universidad de La Laguna, Spain
[2] Euskal Herriko Unibertsitatea-Universidad del País Vasco, Spain
[3] University of California, Davis, USA
[4] Basque Center on Cognition Brain and Language, Donostia, Spain

Number of pages: 29
Number of Figures: 4
Number of Tables: 1

Address for correspondence:
Marta Vergara-Martínez
Departamento de Psicología Cognitiva
Universidad de La Laguna
Campus de Guajara
38205 Tenerife (Spain)
phone: +34 922 317515
fax: +34 922 317461
e-mail: mverga@ull.es

Abstract

In the process of reading compound words, those with high-frequency second constituents are recognized faster than the ones with low-frequency second constituents. However, the role of the first constituent still remains unclear. In the present study, the time course of the frequency effects for both constituents was assessed using Basque compound words embedded in sentences while electrophysiological measures (ERPs) were recorded (Basque is a language with a high frequency of compound words, both right and left-headed). Subjects responded to comprehension questions that were not focused on the compound words. The results revealed that high-frequency first constituents elicited larger negativities starting very early (100-300 ms time window), while low-frequency second constituents elicited larger N400 amplitudes than high-frequency second constituents. Following an activation-verification framework, we argue that the early negativity difference reflects candidate triggering, whereas the N400 difference for the second constituent reflects the cost of its selection and integration for the whole-word meaning to be accessed.

Section: Cognitive and Behavioral Neuroscience
Keywords: Compound word processing, constituent frequency, early negativity, N400, visual word recognition

1. Introduction

The process of visual word recognition involves an early processing stage in which sublexical units (graphemes, morphemes, syllables or words) match their respective mental representation in lexical memory. Most previous research has focused on the processing of simple words, while the recognition of complex words is still poorly understood. Nonetheless, we –either as speakers, writers or readers– very frequently use complex words. The present study aims to understand how we comprehend a morphologically complex type of words: compound words (words that are composed by two or more free-standing lexemes which prototypically belong to major lexical categories: nouns, verbs and adjectives).

A critical issue in the study of morphological processes is to ascertain whether multimorphemic words are represented in their full form (full-listing models; e.g., Butterworth, 1983; Bybee, 1995) or whether only morphemes are stored and combined when complex words are processed (decompositional or full-parsing models; e.g., Libben et al., 1999; McKinnon et al., 2003; Taft, 2004; Taft & Forster, 1976). An intermediate approach is provided by dual-route proposals, according to which only novel multimorphemic words are processed in a decomposed manner, and lexicalized or frequent complex words are processed in a whole-word manner (e.g., Caramazza et al., 1988; Baayen et al., 1997; Schreuder and Baayen, 1995: Isel, Gunter & Friederici, 2003). From the first point of view, decomposition may not be mandatory to access the meaning of a compound word. It predicts no role for morphological-level consituents. The meaning of a word like *milkman* might be accessed from the processing of its constituents, but this might take place at a supra-lexical level (Fowler, Napps & Feldman, 1985; Giraudo & grainger, 2001; Plaut & Gonnerman, 2000). However, from a decompositional perspective, a lexeme parsing mechanism would enforce the conjunction of *milk* and *man* in order to recognize the meaning of *milkman*. Constituent morphemes would be activated early and automatically during lexical access. The implication of this hypothesis is that the recognition of the whole-word would be modulated by the properties of each lexeme. The third approach assumes both a whole-word process and a sequential decomposition process (Bertram & Hyönä, 2003; Inhoff, Radach & Heller, 2000; Pollatsek, Hyönä & Bertram, 2000; Taft, 1994): the meaning of the whole compound would be achieved while at the same time segmentation mechanisms would operate on the progressive recognition of the compound constituents. This last type of models suggest that both decompositional and whole-word processing routes are available, and factors such as prosody and word frequency may determine which route will success when accomplishing the meaning of a compound word. For example, Isel et al. (2003) present a model where prosodic cues (such as the length of the first constituent in auditory compound processing) may determine the route (direct or decompositional) in which the meaning of a compound is accessed. A longer duration of the first morpheme indicates a "non-compound monomorphemic route configuration" while shorter duration of the first morpheme allows identifying it as the onset of a compound word, and so the system would then activate a decompositional route.

Most of the empirical evidence on the processing of compound words has pointed to a pre-lexical decomposition of compound words into their constituent lexemes, at least when the constituents are of higher frequency than the whole-compound (e.g., Lima & Pollatsek, 1983; Andrews, 1986; Sandra, 1990; Zwitserlood, 1994; Hyönä & Pollatsek, 1998; Pollatsek et al., 2000; Andrews et al., 2004; Juhasz et al., 2003; see also Fiorentino & Poeppel, 2007, for evidence of early morphological-structure based computation in a MEG experiment). Constituent priming studies have provided insight about the extent to which the constituents are activated when accessing the whole-word representation of a compound word (see Isel, Gunter & Friederici, 2003, for review). Evidence from unmasked and masked constituent priming experiments (Libben et al., 2003; Monsell, 1985; Zwitserlood, 1944; Forster & Davis, 1984; Shoolman & Andrews, 2003; Duñabeitia, Laka, Perea & Carreiras, in press; Duñabeitia, Perea & Carreiras, in press) indicates that both constituents facilitate the recognition of the compound word. In these studies, similar facilitative priming effects have been obtained for the initial and final constituents with respect to the unrelated priming condition (e.g., book-BOOKSHOP and shop-BOOKSHOP vs. glass-BOOKSHOP). This has been interpreted as a parallel activation of separate representations of the morphemic constituents.

In contrast to the pattern of evidence from constituent priming studies, results from studies which have manipulated each constituent's frequency have shown a different pattern of processing. Although there is evidence showing that high frequency constituents facilitate the recognition of the compound (Hyönä & Pollatsek, 1998; Juhasz, Starr, inhoff & Placke, 2003; Shooolman & Andrews, 2003; Taft, 1979), there is an open debate regarding the impact of each lexeme and its frequency on the processing of compound words. The two first studies that explored the relative influence of the frequency of each constituent lexeme recorded eye movements (Hyönä & Pollatsek, 1998; Pollatsek, Hyönä and Bertram, 2000). Their results showed that the frequency of both constituents influenced the time spent in reading a compound word. High-frequency first constituents had shorter fixations in early measures (Hyönä & Pollatsek, 1998), while high-frequency second constituents had shorter reading times in late measures (Pollatsek et al., 2000). The authors proposed a dual-route model to account for this pattern of results. A decomposition route would allow the constituents be identified and processed while the direct route would be based on the whole-word processing. These two routes would operate in parallel and comprehension would be accomplished as a result of a "race" between them. From this perspective, readers access the first constituent first, and subsequently access the second one, in a serial fashion. This parallel race model predicts a supremacy effect for the first lexeme only at initial stages of lexical access, whereas there would be supremacy for the second lexeme later in processing. At a very early stage of processing (undetected by lexical decision or naming tasks, but noticeable in first fixation duration in silent reading; see Hyönä & Pollatsek, 1998) the first constituent might be processed. Later in processing the second lexeme of the compound becomes relevant, not only in terms of frequency, but also in terms of syntactic properties: the inflectional morphological properties and the syntactic category of the whole compound word are specified by the second lexeme. This dissociation of the frequency effect for the initial and ending lexeme clearly supports a morphemic/lexical decomposition in lexical access.

However, in these studies frequency manipulations were never manipulated orthogonally, and only the frequency of a single constituent was manipulated each time, keeping the frequency of the other constituent invariant. To overcome this limitation, Juhasz et al. (2003) presented participants with compound words that included first and second constituent lexemes of high and low-frequency (orthogonal manipulation) in three different tasks (lexical decision, naming and on-line sentence reading). Their results confirmed a robust second constituent lexeme effect: compound words with a high-frequency second constituent were read faster than those with a low-frequency second constituent. However, they found a trend towards facilitative effects of the first lexeme frequency. These results were interpreted in terms of a meaning-based account, by proposing a privileged role for the second constituent, due to a semantic headedness effect: the meaning of the compound words was mainly determined by the ending lexemes. Under this framework, constituent frequency effects would only show up for the head constituent. With a similar manipulation, Andrews et al. (2004) failed to find a significant effect of the frequency of either constituent in an eye movement experiment, but they found a non-significant facilitative trend for both constituents.

Andrews et al. also linked the second constituent effect to the headedness of the English compound words. Taken together, these studies on constituents' frequency point to a decompositional hierarchical model, since it is the head constituent the one which guides the segmentation process.

Interestingly, in a recent study, Duñabeitia, Perea & Carreiras (2007a) presented evidence from two lexical decision experiments in Basque and Spanish that supported the second constituent frequency effect (in line with Juhasz et al., 2003, Andrews et al., 2004, and Pollatsek et al., 2000). Two parallel experiments were run in two languages that differ largely in two main points regarding compound word creation: productivity and headedness. Basque is highly productive in compound creation. In contrast, Spanish is a language with a low number of compound words. Spanish compound words tend to be right-headed (i.e., the whole-word meaning is driven by the meaning of the second constituent, like in English; *pasatiempo* for *pastime*, *pasa*[*pass*] + *tiempo*[*time*]). Basque compound words have more distributed headedness groups, including a large group of initial headed compound words (e.g., *pastime* in Basque is translated as *denborapasa*, composed by *denbora*[*time*] + *pasa*[*pass*]). The two lexical decision experiments (Duñabeitia et al., 2007) included an orthogonal manipulation of the frequency of the constituents, keeping the whole-word frequency invariant. Results confirmed previous data, showing a significant second lexeme effect in both languages, and no cues for any first constituent significant effect. Moreover, the absence of a general headedness effect pointed to a blind-to-semantics morphological decomposition (e.g., Taft, 1994; Rastle et al., 2004). Duñabeitia and colleagues suggested that an activation-verification framework (e.g., Paap, et al., 1982; Reichle et al., 1998) could be efficiently adapted to account for their findings with transparent compounds in the following way: first, the initial constituent would be treated as a separated unit (i.e., tea in teacup), not only activating the corresponding lexeme (tea) but also, activating possible candidates to a higher or lower degree depending on its frequency, including the whole compound (teacup). The verification process for the recognition of the whole compound cannot be carried out until the second unit shows up (the second constituent, cup). The second lexeme fires a new process, activating its corresponding lexeme and compound word (cup and teacup). Then, the verification process can be satisfactorily carried out, as the ending lexeme closes the orthographical, morphological, lexical, syntactic and semantic retrieval. Therefore, one could expect the frequency of the ending lexeme to play a large role at later stages, given that the second lexeme is the one that accomplishes the activation-verification procedure.

The main goal of the present investigation is to trace the time course of each constituent's frequency by measuring the ERPs elicited by compound words in which first and second constituents' frequency was manipulated. One of the most important ERP components related to (although not exclusive to) language processing is the N400, a negative peak with maximum amplitude around 400 ms after stimulus onset elicited in response to the processing of a word or any other meaningful stimulus. This component typically shows a centro-parietal scalp distribution in the visual modality (Kutas & Hillyard, 1980; Kutas & Van Petten, 1994; Kutas & Federmeier, 2000). Among other things, the N400 has been found to correlate with lexico-semantic aspects of single word processing (Kutas & Federmeier, 2000). On the one hand, it has been shown that the amplitude of this negativity is an inverse function of lexical frequency: the amplitude is larger for low-frequency than high-frequency words, which is thought to reflect greater processing demands on a lexico-semantic candidate selection level of processing (Neville et al., 1992; Bentin et al., 1985; Van Petten & Kutas, 1990; Van Petten, 1995). ERP research regarding the processing of compound words is very scarce (e.g., El Yagoubi et al., in press; Koester et al., 2004; Koester et al., 2007; Pratarelli, 1995), and has not yet investigated the effect of frequency of constituents. However, the N400 has also been related to the lexical-semantic integration of compound constituents (Koester et al., 2007).

The behavioral pattern for the second constituent frequency manipulation in our previous work (Duñabeitia, Perea & Carreiras, 2007a) has shown facilitative effects of frequency for the second constituent. Thus, we would expect larger N400 for second constituents with lower compared to those with higher lexical frequency. However, predictions for the ERP results on the first constituent's frequency manipulation are not so straightforward. Most of the research has not found a clear pattern of results for the first constituent (Juhasz et al., 2003). In fact, the effects from the first constituent seem to be noticeable only during silent reading, and they are undetected in lexical decision experiments or naming tasks (see Hyönä & Pollatsek, 1998). One of our main hypotheses shall be based on the timing for each of these effects, since previous results have shown that first constituent frequency has an impact on early eye measures while second constituent has an impact on later measures (Pollatsek et al., 2000). Based also in the proposal of an activation-verification framework for compound processing, we could expect that the processing of the first constituent would activate not only its own lexical representation but would also activate possible candidates to a higher or lower degree depending on its frequency. In a similar direction, Holcomb, Grainger and O´Rourke (2002) showed a significant influence of orthographic neighborhood in the standard window of the N400 (350-550 ms) and even earlier effects were obtained for the same stimuli while presented in a semantic categorization task (150-350 ms): larger negativities were observed for words with large compared to those from small orthographic neighborhoods. Their interpretation was that stimuli with larger number of neighbors lead to increased levels of activation and that this increase in activation could be associated with some measures of global lexical activity, reflected as an increase of N400 amplitude. In the same direction, Kounios & Holcomb (1992) has shown that the N400 was larger when words had more semantic associations.

Considering preceding behavioral evidence, we predicted different ERP effects for the first and second constituents which would reflect the sequential time-course of processing that has been already proposed (Hyönä & Pollatsek, 1998). If, as stated before, first and second constituents are accessed in a serial way, with the initial constituent being processed earlier than the final, this pattern should have a reflection in our ERP data, reinforcing the view of a temporal processing distinction. This is a highly interesting issue that could help to determine whether the pattern of results that has been obtained in previous eye-movement experiments, and that have constituted the basis of the compound word reading models of visual word recognition, correspond to the pattern of results that emerge at an electrophysiological level. Regarding the manipulation on constituents' frequency, and according to the "race" model, one would expect larger N400 amplitudes for low compared to high frequency second constituent. This and other accounts based on behavioral data (Andrews et al., 2004; Juhasz et al., 2003) do not provide any specific prediction for the first constituents' frequency manipulations. However, and if our hypothesis of activation-verification holds true, we would expect larger negativities for high compared to low frequency first constituents (according to the proposal of Holcomb et al., 2002, where the N400 and even early starting negativities could also reflect an increase of the global lexical activity). This is the opposite pattern that should be expected according to the eye movements data found for the first constituent.

ERP effects of compounds' constituent frequency were measured while participants were presented with compound words inserted in sentences, in which the frequency of the first and the second constituent was manipulated orthogonally (high-high, high-low, low-high and low-low). Up to now, the vast majority of the data supporting a decomposition of compound words during visual word recognition has been obtained in sentence reading experiments.

However, the observed effects in single word presentation experiments are inconclusive at distinguishing between pre-lexical and post-lexical effects, despite the fact that frequency manipulations have yielded different effects for the first and the second constituent. In addition, it is difficult to investigate the time-course of the processing in these experiments with lexical decision or naming latency and accuracy. For these reasons, compound words were included in sentences. Moreover, since the race model proposed by Hyönä and colleagues (e.g., Hyönä and Pollatsek, 1998; Pollatsek, Hyönä and Bertram, 2000) clearly predicts differences in the time-course of the access to the constituents of the compound word, and they have collected evidence with compound words inserted into sentences, it was desirable to follow a similar procedure. The experiment presented here was done in Basque, an agglutinative language. We believe that it is always desirable to explore a sample of natural languages as wide as possible, and in general, to make sure one is on solid grounds regarding language-processing results. From this perspective, Basque is as good a language to study morphological processing as any other. Besides, Basque is a particularly well-suited language for our scientific purposes: this language is very productive in compound formation, far more so than its neighboring Spanish, for instance. Thus, we were able to use a large number of right- and left-headed compounds, what was required to run the study properly.

Subjects were asked to read the sentences (presented word by word) for comprehension. ERPs were time locked to the onset of the compound words and recorded for 800 ms.

2. Results

The ERP grand averages, time-locked to the onset of the four different word groups are represented in Figures 2 and 3, separately for the first and the second constituent frequency manipulation. As shown in the two figures, compound words elicited the N1-P2 complexes followed by a wide negativity starting at about 300 ms and lasting until 700 ms post-stimulus approximately. The next negativity that is shown at around 750 ms corresponds to the N1 component elicited by the second word of the sentence. Visual inspection reveals an early sustained negativity for high- compared to low-frequency first constituents over most areas. No differences seem to be occurring for the second constituent frequency manipulation in early windows. In addition, the following negative component seems to show differences for both the first and the second constituent frequency. Three windows of analysis were selected: an early window starting from 100 until 300 ms, a second window from 300 ms until 450 ms, and a late window between 350 and 650 milliseconds. In addition, voltage maps (see Figure 4) show a different topographical distribution for the frequency of the first and second constituents. The statistical analyses confirmed the differences described after visual inspection. Mean amplitudes were obtained for different time windows. For each window, a repeated-measures ANOVA was performed, including *electrode regions* (anterior, medial and posterior), *hemisphere* (left and right), *frequency* of the *first constituent* (High and Low) and *frequency* of the *second constituent* (High and Low). Where appropriate, critical values were adjusted using the Greenhouse-Geisser (1959) correction for violation of the assumption of sphericity. Effects for the electrode region factor or for the hemisphere factor are only reported when they interact with the experimental manipulations.

<Insert Figures 2, 3 and 4 about here>

2.1 100-300 epoch

The ANOVA with the average values of the 100-300 ms time epoch only showed a main effect of first constituent frequency [*F*(1,22)=8.52, *p*<.01; MSE= 91.02]. Larger negativities were observed for high-frequency first constituents compared to low-frequency ones (.812 μV). The difference between the low and high second constituent frequency (.409 μV) was not statistically significant in the same time window [*F*(1,22)=1.40, *p*=.24; MSE= 23.11].

2.2. 300-450 epoch

The ANOVA with the average values of the 300-450 ms time epoch showed no effect of the first constituent frequency [*F*(1,22)=1.37, *p*=.25, MSE: 24.75; difference: .424 μV]. Regarding the second constituent, no main effect of frequency was obtained [*F*(1,22)=2.28, *p*=.14, MSE: 41.51; difference: .548 μV]. However, a significant interaction between second constituent frequency and hemisphere was observed [*F*(1,22)=6.51, *p*<.05; MSE: 3.57]. Simple comparisons showed significant differences between high and low second constituent frequency over the right hemisphere [F(1,22)=4.56, p<.05; MSE: 34.7; diff: .709 μV : left hemisphere: F<1], with larger negativities for low-frequency compared to high-frequency ones.

2.2. 450-650 epoch

The ANOVA with the average values of the 450-650 ms time epoch showed a main effect of the first constituent frequency [*F*(1,22)=4.30, *p*<.05, MSE: 99.88; difference: .851 μV], with larger negativities for the high compared to the low frequency condition. Regarding the second constituent, no main effect of frequency was obtained [*F*(1,22)=2.64, *p*=.118, MSE: 62.16; difference: .671 μV]. However, and again, a significant interaction between second constituent frequency and hemisphere was observed [*F*(1,22)=5.40, *p*<.05; MSE: 2.97]. Simple comparisons showed significant differences between high and low second constituent frequency over the right hemisphere [F(1,22)=4.80, p<.05; MSE: 46.17; diff: .818 μV: left hemisphere: F(1,22)=1.31, p=.26; MSE: 18.96.17; diff: .524 μV], with larger negative amplitudes for low-frequency compared to high-frequency ones. Nevertheless, the first constituent frequency effect was in the opposite direction: larger negative values were obtained for the high constituent compared to the low constituent frequency.

3. Summary and general discussion

The ERP results were straightforward regarding both constituent frequency effects. Low-frequency second constituents elicited larger N400 amplitudes compared to high-frequency second constituents These differences started around 300 ms and lasted at about 650 ms. In contrast, larger negativities were obtained for high-frequency than low-frequency first constituents, starting much earlier than second constituent effects: they were observed at the 100-300ms time window.

Electrophysiological measures of passive word reading have allowed for the capture of subtle differences in cognitive processing that may not affect typical behavioral measures such as response times in a lexical decision task, or naming latencies in reading aloud experiments. In this study we aimed at exploring the validity of a widely spread model of compound word recognition which was grounded on behavioral data from eye-movements and supported by data from lexical decision and naming experiments (Hyönä & Pollatsek, 1998; see also Andrews et al., 2004; Duñabeitia et al., 2007a; Juhasz et al., 2003). According to this proposal, both constituent frequency manipulations would have been expected to induce similar ERP effects, since the model assumes that both constituents engage similar processes that simply occur in a different timing (access to the initial constituent is achieved earlier than access to the second constituent). In other words, behavioral facilitation from both constituents might have shown up on the ERP waveforms by affecting the same component, or if different components, at least in the same direction or with the same scalp distribution. As shown, this prediction was not observed in our data. Thus, it seems that each constituent might be engaged in different sub-processes occurring along meaning access of the whole compound word. Moreover, the difference in time-course and scalp distribution between the first and second constituent frequency effects supports this two stage framework of lexical access in compound word processing. As proposed in the Introduction, under an activation-verification account, the first constituent would trigger the activation of different candidates according to its frequency, while the second constituent would facilitate the recognition of the whole word. Second constituents might play a main role in the selection of the final lexical candidate among those that are triggered by the processing of the first constituent. The higher the frequency of the second constituent, the smaller the amount of effort displayed as amplitude voltage (ERP measures) or shorter RTs (behavioral measures) in the process of accessing the word meaning. This later stage would be similar to the verification stage in classical activation-verification frameworks (Paap et al., 1982; Reichle et al., 1998). Hence, at this point, it seems that an activation-verification model of compound word processing can be satisfactorily proposed, capturing the pattern of data that has been observed.

To our knowledge, this study constitutes the first ERP investigation of constituent frequency in compound word reading. We think that the differences observed regarding to each constituent frequency manipulation reveal that each constituent is operating differently during the lexical access of the whole compound. Apart from the differences regarding the time course of both effects (the first constituent frequency effects starting earlier than the second), not only the direction of the effects is reversed but the scalp distribution is different too. The dissociation of the frequency effects for each constituent on ERP measures is in line with previous behavioral evidence pointing to a pre-lexical decomposition of compound words into their constituent lexemes. Regarding the second constituent, behavioral studies have consistently shown facilitative effects of second constituent frequency (Juhasz et al., 2003; Andrews et al., 2004; Pollatsek et al., 2000). More recent results in Spanish and Basque have also shown facilitative second lexeme frequency effects (Duñabeitia et al., 2007a). However, behavioral effects of first constituent frequency are more subtle. Juhasz and colleagues reported a marginal main inhibitory effect in accuracy in their lexical decision experiment (Experiment 1). In the Spanish sub-experiment from Duñabeitia and colleagues the same effect arose in the error rate analysis, pointing to a sort of inhibitory effect of initial constituents' frequency. The present experiment confirms these intuitions from previous behavioral experiments, in the sense that a different temporal distribution of the constituents' frequency effects has been observed, as well as a different impact of each constituent's frequency in whole-word processing.

On the other hand and according to our hypothesis, the early starting negativities associated with high frequency first constituents could reflect the initial triggering of lexical candidates that in principle match with the first part of the word. The same type of pattern has been observed for words with large compared to low orthographic neighborhoods (Holcomb et al., 2002): not only N400 differences were observed regarding different orthographic neighborhood size, but this difference started earlier (150 ms) in the semantic categorization task. As in our present results, these differences started early but were present until 650 ms (although the difference was not statistically significant from 300 to 450 ms). According to our proposal and also to the previous results from neighborhood studies, the effect regarding the first constituent may reflect differences in the overall semantic activation generated by different stimuli (first constituents with different frequency). Interestingly, constituents' frequency also correlated with the size of the morphological family of the first constituent. High-frequency first constituents had much larger morphological families than low-frequency first constituents (High-High: 28; High-Low: 35; Low-High: 16; Low-Low: 19), which may be helping to increase the level of semantic activation for high compared to low frequency first constituents. In line with this proposal, Pylkkänen et al. (2004) reported an inhibitory effect of morphological family size on the peak latency of the M350 in a MEG experiment. This delay was interpreted to reflect competition among all the forms derived from the same root, all of which would have been activated initially (Pylkkänen et al., 2004). As shown in our ERP data, this process of semantic activation starts very early in time and is related to the processing of the first constituents of compound words, but the effect remains until late in time. The fact that no interaction had been obtained between each of the two constituents frequencies, shows that this long lasting negativity is reflecting processes which start earlier but proceed in parallel to those underlying the second constituent frequency effects. In this sense, the results are in accordance to the basic assumption of models of parallel processing of constituents (e.g., Hyönä & Pollatsek, 1998), which state that access to both constituents takes place during compound word recognition, together with processes of whole-word recognition. However, this model fails at predicting the different effects observed for first and second constituents.

On the other hand, the frequency of the second constituent facilitates the lexical selection of the compound, as shown by previous RT results and our present ERP results. This effect is similar in its peak latency, polarity and scalp distribution to the N400 (Kutas & Federmeier, 2000), and resembles the type of modulation previously observed by Koester et al. (2006). These last authors presented German opaque and transparent two-constituent compounds in the auditory modality while manipulating the gender agreement between a determiner and the initial (non-head) or last (head) constituent (Koester et al., 2004). Gender violations of initial constituents (and head constituents too: Koester et al., 2006) resulted in a left-anterior negativity (LAN) for both opaque and transparent compounds, which was interpreted as an index of morphosyntactic decomposition. Interestingly, they further compared the processing of opaque and transparent compounds in order to investigate processes related to semantic composition. Transparent compounds require semantic constituent integration (at least for low-frequent ones, as they may not be lexicalized) which would be absent in the case of opaque compounds (which must have a lexical entry independent of the relation between their constituents). This difference turned out to be reflected in larger negativities for the transparent compared to the opaque compound words, with the same centro-parietal distribution (and right-sided dominance) as the N400 effect (Kutas & Federmeier, 2000). This pattern of results was interpreted in terms of semantic constituent integration "cost" for low-frequent transparent compounds compared to low-frequent opaque compounds. This interpretation of the N400 is in accordance with the hypothesis of the *knowledge integration effort* (Holcomb, 1993), where the N400 amplitude is assumed to be directly proportional to the effort required by this integration process to fit each item in the representation. When two meanings have to be integrated in the process of accessing the semantic representation corresponding to the compound, there is a certain amount of knowledge associated to the first constituent (which in fact is the first to be encountered) which might need to be inhibited, while, at the same time, the second constituent is also being processed. On the one hand, the cost of this inhibition may reflect possible inhibitory processes associated with the N400 (Debruille, 2007), but it may well also be the case that it is not reflected in the N400 component as has been shown by the scalp distribution of the early starting negativity effect associated to high frequent first constituents. This is an interesting question that should be addressed in future studies. On the other hand, in the particular case of compound words, the integration process suggested by Holcomb might be guided by the second constituent (given the previous context which in this situation corresponds to the first constituent). The first constituent would not only trigger the activation of possible candidates, but would also determine the context into which the second constituent representation has to be integrated. It is then possible that the N400 associated to the second constituent frequency effect in this type of "context" reflects integration processes.

In sum, the present study is concerned with the processing of compound words that are presented visually. We have reviewed the results from different paradigms regarding the study of compound word processing (priming studies and constituent frequency studies) and have proposed a working hypothesis in order to integrate the pattern of results obtained in different languages. We have investigated the effect of each constituent's frequency on the processing of Basque compound words by measuring ERP responses to words embedded in sentences. This experiment showed an early starting negativity modulated by the frequency of the first constituent as well as an N400 effect modulated by the frequency of the second constituent. ERP effects are consistent with an interpretation of the results in terms of an activation-verification framework (see Duñabeitia et al., 2007a). The first constituent would trigger the activation (and competition) of different representations to which it is associated. Retrieval of the whole representation of the compound word is carried out once the second constituent is recognized and integrated. The frequency of this constituent speeds up the integration of the final meaning and so is shown in ERP measures (smaller N400 amplitudes).

4. Experimental Procedure

Participants Twenty-three (16 female, mean age = 20, SD = 2.67) undergraduate students at the University of the Basque Country took part in this data collection, which took place at ELEBILAB, the Psycholinguistics Laboratory in Vitoria-Gasteiz. They received 3 € or course credit in exchange for their participation. All were Basque native speakers with no history of neurological or psychiatric impairment, and with normal or corrected-to-normal vision. All participants were right-handed, as assessed with an abridged Spanish version of the Edinburgh Handedness Inventory (Oldfield, 1971).

Materials We selected 120 compound words from the Basque database (Perea et al., 2006). All of these compound words were decomposable into their two constituent lexemes without adding or deleting any letter (e.g., the Basque word *ikuspuntu*, translated as *point of view*, can be decomposed into *ikus* [*view*] and *puntu* [*point*]). The mean length of the compound words was 9.25 letters (range: 6-12), and the mean frequency was of 8.53 appearances per million words (range: 0.28-77.3). The mean frequency of the first constituent was 313.17 and the mean length was 4.64. The mean frequency of the second constituent was 331.06 and the mean length was 4.65. These 120 compound words were divided into four groups of 30 words each, following an orthogonal manipulation based on the mean frequency of the constituents, while keeping whole-word frequency and length constant across conditions (see Table 1 for examples of each condition). Constituents in the different frequency conditions were controlled for length, number of orthographic neighbors and concreteness. Concreteness values were obtained from a questionnaire in which 20 native Basque speakers rated each of the constituents in a 1-to-7 Likert scale (1=more abstract, 7=more concrete). A different group of 30 Basque speakers completed another questionnaire in which they rated each compound word for semantic transparency, headedness and concreteness. For the transparency rating, participants had to rate each compound in a 1-to-7 scale (1=more opaque, 7=more transparent). A transparent compound was defined as that compound whose meaning could be directly obtained from the meaning of its constituents. On average, all the compounds were rated as highly transparent (mean transparency score=5.63). Moreover, no differences were obtained in the transparency degree of the compounds in each frequency condition (High-High=5.63, High-Low=5.55, Low-High=5.66, Low-Low=5.67; all ps>.55). For the concreteness rating, a similar 7-point scale was used (1=more abstract, 7=more concrete), and no differences were obtained across conditions (all ps>.30; see Table 1). For the headedness ratings, participants had to rate the compound words with 1 if they were left-headed (the head is the initial morpheme), with 3 if they were right-headed (the head is the final morpheme), or with 2 when no clear distinction could be done. In order to rate the semantic headedness of each compound word, participants were asked to identify the constituent that contributed the most semantic information in the meaning. Results showed that, on average, 43.97% of the compounds were rated as left-headed, 21.94% were rated as right-headed and 44.09% were rated as indistinctive (with no statistical differences across conditions, ps>.13).

<Insert Table 1 about here>

These 120 compound words were inserted into 120 sentences. All of the sentences started with a compound word, followed by six more non-compound words. Thus, all the sentences were seven words long. The sentences respected the natural syntactic order in Basque and were all meaningful sentences. The compound words were the first words of the sentences because in this way any context-related predictability effect could be avoided. By way of example, we transcribe here one of the sentences that was used in this experiment:

Bizkarzain gutxi dauzkaten politikariak oso baikorrak dira.

Politicians that have few bodyguards are very optimistic.

[Bizkarzain][gutxi][dauzkaten][politikariak][oso][baikorrak][dira].

[Bodyguard] [few] [have that] [politicians] [very] [optimistic] [are].

In order to avoid any kind of expectancy from participants to the presentation of a compound word in initial position, an extra set of 270 seven-word long sentences was created. The percentage of target sentences starting with a compound was 30%. Considering that the aim of the study was to explore compound word processing in natural reading, we wanted to avoid participants' expectancy to read a compound word in initial position. A lower proportion of fillers would have led to the development of the expectancy of finding a compound word in sentence-initial position, consequently making the frequency manipulation more salient. None of these filler sentences included any compound word in any position. Participants were presented with the total set of 390 sentences in random order. They were asked to answer to a true-false question about the content of the sentences just read every 5 sentences on average. These questions could appear only after a filler sentence. The question always referred to a statement from the preceding sentence, and participants had to press one of two buttons to answer if the statement was TRUE or FALSE (on average participants correctly responded to more than 90% of the questions). By this procedure, subjects read sentences passively and no explicit task was executed on the compound words or on the experimental sentences. Comprehension check is commonly used in sentence-reading studies (e.g. Balota, Pollatsek, & Rayner, 1985; Rayner, Sereno, & Raney, 1996).

Procedure A fixation point (''+'') appeared at the centre of the screen and remained there for 1000 ms. This fixation point was followed by a blank screen interval of 300 ms, and then the sentence was displayed word by word. Each word appeared for 300 ms and was followed by a 300-msec blank interval. Comprehension questions were displayed for 2000 ms. Following the questions, the two possible answers (TRUE and FALSE) remained on the screen for a maximum of 3000 ms or until the participant's response. Subjects responded with 2 buttons accordingly to the location of the TRUE and FALSE words on the screen (left or right, respectively). The TRUE and FALSE word locations were counterbalanced. Before starting the experimental phase 8 warm up practice trials were presented to the participants.

EEG recording and analyses Scalp voltages were collected from 58 Ag/AgCl electrodes which were mounted in an elastic cap (ElectroCap International, Eaton, USA, 10-10 system). The right mastoid was used as reference (see Figure 3). Eye movements and blinks were monitored with two further electrodes providing bipolar recordings of the horizontal and vertical electro-oculogram (EOG). Inter-electrode impedances were kept below 5 $K\Omega$. EEG was filtered with an analogue bandpass filter of 0.01-50 Hz and a digital 15 Hz low-pass filter was applied before analysis. The signals were sampled continuously throughout the experiment with a sampling rate of 500 Hz, and digitally re-referenced to linked mastoids.

<Insert Figure 1 about here>

Epochs of the EEG corresponding to 800 ms after target word presentation were averaged and analyzed. Baseline correction was performed using the average EEG activity in the 200 ms preceding the onset of the target stimuli as a reference signal value. Following baseline correction, epochs with simultaneous artifacts in at least 10 channels were rejected. This resulted in the exclusion of approximately 12 % of the trials. Those epochs free of ocular artifacts (blinks and eye movements) and muscular artifacts were averaged and analyzed

Six regions of interest were computed out of the 58 electrodes, each containing the mean of a group of 6 different electrodes. The regions were (see electrode numbers in Figure1): left-anterior (F5, F3, F1, C5A, C3A, C1A), left-medial (C5, C3, C1, TCP1, C3P, C1P), left-posterior (P5, P3, P1, CB1, P3P, P1P), right-anterior (F2, F4, F6, C2A, C4A, C6A), right-medial (C2, C4, C6, C2P, C4P, TCP2) and right-posterior (P2, P4, P6, P2P, P4P, CB2).

Acknowledgements

The research reported in this article has been partially supported by Grants, SEJ2006-09238, SEJ2007-60751/PSIC and CONSOLIDER-INGENIO CSD2007-00012 and 2010 CSD2008-00048 from the Spanish Ministry of Science and Technology, and by Grants BFI05.310, GIU06/52 from the Basque Government. Marta Vergara-Martínez was supported by a postdoctoral contract under the Spanish Fulbright-MEC commission.

References

Andrews, S., 1986. Morphological influences on lexical access: lexical or nonlexical effects? J. Mem. Lang. 25, 726–740.

Andrews, S., Miller, B., Rayner, K., 2004. Eye fixation measures of morphological segmentation of compound words: There is a mouse in the mousetrap. Eur. J. Cogn. Psychol. 16(2), 285-311.

Baayen, R.H., Dijkstra, T., Schreuder, R., 1997. Singulars and plurals in Dutch: evidence for a parallel dual-route model. J. Mem. Lang. 37, 94–117.

Balota, D.A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. Cognitive Psychology, 17, 364-390.

Barber, H.C.A., Vergara, M., Carreiras, M., 2004. Syllable-frequency effects in visual word recognition: evidence from ERPs. Neuroreport 15(3), 545-548.

Bentin, S., McCarthy, G., Wood, C.C., 1985. Event-related potentials, lexical decision and semantic priming. Electroencephalogr. Clin. Neurophysiol. 60(4):343-355.

Butterworth, B., 1983. Lexical representation. In B. Butterworth (Ed.), Language production Vol. 2, London: Academic Press. pp. 257–294.

Bybee, J., 1995. Regular morphology and the lexicon. Lang. Cogn. Processes. 10, 425–455.

Caramazza, A., Laudanna, A., Romani, C., 1988. Lexical access and inflectional morphology. Cognition 28, 297–332.

Carreiras, M., Ferrand, L., Grainger, J., Perea, M., 2005. Sequential effect of phonological priming in visual word recognition. Psychol. Sci. 16 (8), 585- 589.

Carreiras, M., Vergara, M., Barber, H.C.A., 2005. Early Event-related Potential Effects of Syllabic Processing during Visual Word Recognition. J. Cogn. Neurosci. 17(11): 1803-1817.

Debruille, J.B., 2007. The N400 potential could index a semantic inhibition. Brain Res. Rev. 56(2), 472-477.

Duñabeitia, J.A., Perea, M., Carreiras, M., 2007a. The role of the frequency of constituents in compound words: Evidence from Basque and Spanish. Psychon. Bull. Rev. 14, 1171–1176.

Duñabeitia, J.A., Perea, M., Carreiras, M., 2007b. Processing compound words: evidence from a multi-task approach. The first constituent inhibition effect and the second constituent facilitation effect. Poster presented at the 5th International Workshop on Morphological Processing, Marseille, France.

Duñabeitia, J.A., Laka, I., Perea, M., Carreiras, M., in press. Is Milkman a superhero like Batman? Constituent morphological priming in compound words. Eur. J. Cogn. Psychol.

Duñabeitia, J.A., Marín, A., Avilés, A., Perea, M., Carreiras, M., in press. Constituent priming effects: Evidence for preserved morphological processing in healthy old readers. Eur. J. Cogn. Psychol.

Duñabeitia, J.A., Perea, M., Carreiras, M., in press. Does darkness lead to happiness? Masked suffix priming effects. Lang. Cogn. Processes.

Fiorentino, R., Poeppel, D., 2007. Compound words and structure in the lexicon. Lang. Cogn. Processes. 22 (7), 953 - 1000.

Forster, K.I., Davis, C., 1984. Repetition priming and frequency attenuation in lexical access. J. Exper. Psychol. Learn. Mem. Cogn. 10, 680-698.

Holcomb, P.J., 1993. Semantic priming and stimulus degradation: implications for the role of the N400 in language processing. Psychophysiology 30(1):47-61.

Hutzler F., Bergmann J., Conrad M., Kronbichler M., Stenneken P., Jacobs A.M., 2004. Inhibitory effects of first syllable-frequency in lexical decision: An event-related potential study. Neurosci. Lett. 372(3): 179–184.

Hyönä, J., Pollatsek, A., 1998. The role of component morphemes on eye fixations when reading Finnish compound words. J. Exp. Psychol. Hum. Percept. Perform. 24, 1612-1627.

Isel, F., Gunter, T.C., Friederici, A.D., 2003. Prosody-assisted head-driven access to spoken German compounds. J. Exper. Psychol. Learn. Mem. Cogn. 29(2), 277-288.

Jarema, C., Busson, R., Nikolova, K., Tsapkini, K., Libben, G., 1999. Processing compounds: A cross-linguistic study. Brain Lang. 68, 362–369.

Juhasz, B.J., Starr, M.S., Inhoff, A.W., Placke, L., 2003. The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. Br. J. Psychol. 94, 223-244.

Koester, D., Gunter, T.C., Wagner, S., Friederici, A.D., 2004. Morphosyntax, prosody, and linking elements: the auditory processing of German nominal compounds. J. Cogn. Neurosci. 16, 1647–1668.

Koester, D., Gunter, T.C., Wagner, S., 2007. The morphosyntactic decomposition and semantic composition of German compound words investigated by ERPs. Brain Lang. 102(1), 64-79.

Kutas, M., VanPetten, C., 1988. Event-related brain potential studies of language. In: P.K. Ackles, J.R. Jennings and M.G.H. Coles (Eds.), Advances in Psychophysiology, Volume 3, New York: JAI Press, pp.139–187.

Kutas, M., Van Petten, C., 1994. Psycholinguistics electrified. In: Gernsbacher, M.A. (Ed.), Handbook of Psycholinguistics. Academic Press, San Diego, pp. 83–143.

Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends Cogn. Sci. 4, 463–470.

Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. Science, 207, 203–205.

Libben, G., 1998. Semantic transparency in the processing of compounds: Consequences for representation, processing, and impairment. Brain Lang. 61 (1), 30–44.

Libben, G., Derwing, B.L., de Almeida, R.G., 1999. Ambiguous novel compounds and models of morphological parsing. Brain Lang. 68, 378–386.

Libben, G., Gibson, M., Yoon, Y.B., Sandra, D., 2003. Compound fracture: The role of semantic transparency and morphological headedness. Brain Lang. 84(1), 50-64.

Lima, S.D. Pollatsek, A., 1983. Lexical access via an orthographic code? The basic orthographic syllable (BOSS) reconsidered. J. Verbal Learn. Verbal Behav 22, 310–322.

McKinnon, R., Allen, M., Osterhout, L., 2003. Morphological decomposition involving non-productive morphemes: ERP evidence. NeuroReport, 14, 883–886.

Monsell, S., 1985. Repetition and the lexicon. In A. W. Ellis (Ed.), Progress in the psychology of Language, Vol. 2, Hillsdale, NJ: Erlbaum, pp. 147–195.

Neville, H., Mills, D.L., Lawson, D., 1992. Fractionating language: different neural subsystems with different sensitive periods. Cereb. Cortex 2, 244–258.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory, Neuropsychologia, 9, 97-113.

Paap, K.R., Newsome, S.L., McDonald, J.E., Schvaneveldt, R.W., 1982. An activation-verification model for letter and word recognition: The word superiority effect. Psychol. Rev. 89, 573-594.

Perea, M., Urkia, M., Davis, C. J., Agirre, A., Laseka, E., Carreiras, M., 2006. E-Hitz: A word-frequency list and a program for deriving psycholinguistic statistics in an agglutinative language (Basque). Behav. Res. Methods, 38, 610–615.

Pollatsek, A. Hyönä, J., Bertram, R., 2000. The role of morphological constituents in reading Finnish compound words. J. Exp. Psychol. Hum. Percept. Perform. 26, 820-833.

Pratarelli, M., 1995. Modulation of semantic processing using word length and complexity: an ERP study. Int. J. Psychophysiol. 19(3), 233-246.

Rastle, K., Davis, M.H., New, B., 2004. The Broth in my Brother's Brothel: Morpho-Orthographic Segmentation in Visual Word Recognition. Psychon. Bull. Rev. 11, 1090-1098.

Rayner, K., Sereno, S.C., & Raney, G.E. (1996). Eye movement control in reading: a comparison of two types of models. Journal of Experimental Psychology: Human Perception and Performance, 22(5), 1188-1200.

Reichle, E., Pollatsek, A., Fisher, D.L., Rayner, K., 1998. Toward a model of eye movement control in reading. Psychol. Rev. 105, 125-157.

Sandra, D., 1990. On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. Quarterly J. Exp. Psychol., 42A(3), 529-567.

Schreuder, R. Baayen, R.H., 1995. Modeling morphological processing. In: L. B. Feldman (Ed.), Morphological aspects of language processing. Erlbaum: Hillsdale, pp. 131–154.

Shoolman, N., Andrews, S., 2003. Racehorses, reindeers and sparrows: Using masked priming to investigate morphological influences on word identification. In S. Kinoshita S. Lupker (Eds.), Masked priming: The state of the art. Hove, UK: Psychology Press, pp. 241-278.

Taft, M., 1994. Interactive-activation as a framework for understanding morphological processing. Lang. Cogn. Processes 9, 271–294.

Taft, M., 2004. Morphological decomposition and the reverse base frequency. Quarterly J. Exp. Psychol. 57A, 745–765.

Taft, M., Forster, K.I., 1976. Lexical storage and retrieval of polymorphemic and polysyllabic words. J. Verbal Learn. Verbal Behav. 15, 607–620.

Van Petten, C., 1995. Words and sentences: Event-related brain potential measures. Psychophysiology, 32, 511-525.

Van Petten, C., Kutas, M., 1990. Interactions between sentence context and word frequency in event-related brain potentials. Mem. Cogn. 18, 380-393.

Yagoubi, R.E., Chiarelli, V., Mondini, S., Perrone, G., Danieli, M., Semenza, C., in press. Neural correlates of Italian nominal compounds and potential impact of headedness effect: An ERP study. Cognitive Neuropsychol.

Zwitserlood, P., 1994. The role of semantic transparency in the processing and representation of Dutch compounds, Lang. Cogn. Processes 9, 341–368.

Figure Captions

Figure 1. Schematic flat representation of the 58 electrode positions from which EEG activity was recorded (front of head is at top). The grouped electrodes are those included for the analyses in the six critical regions. The twelve double-rounded electrodes correspond to those displayed in Figures 2, 3 and 4.

Figure 2. Grand average ERPs corresponding to the compound words with high and low first constituent frequency (high and low) in twelve representative electrodes (two from each area of interest). Negative potentials are plotted upwards and each hash mask represents 200 ms.

Figure 3. Grand average ERPs corresponding to the compound words with high and low second constituent frequency (high and low) in twelve representative electrodes (two from each area of interest). Negative potentials are plotted upwards and each hash mask represents 200 ms.

Figure 4. (a) Topographical distribution of the first constituent frequency effect in terms of amplitude differences between two conditions: high-frequency first constituent minus low-frequency first constituent, averaged between 100 and 300 ms post-stimuli. Larger negative amplitudes for the high-frequent condition are reflected in negative difference voltage values over anterior areas. (b) Topographical distribution of the second constituent frequency effect in terms of amplitude differences between two conditions: high-frequency second constituent minus low-frequency second constituent, averaged between 450 and 650 ms post-stimuli. Larger negative amplitudes for the low-frequent condition are reflected in positive difference voltage values over the right hemisphere.

FIGURE 1

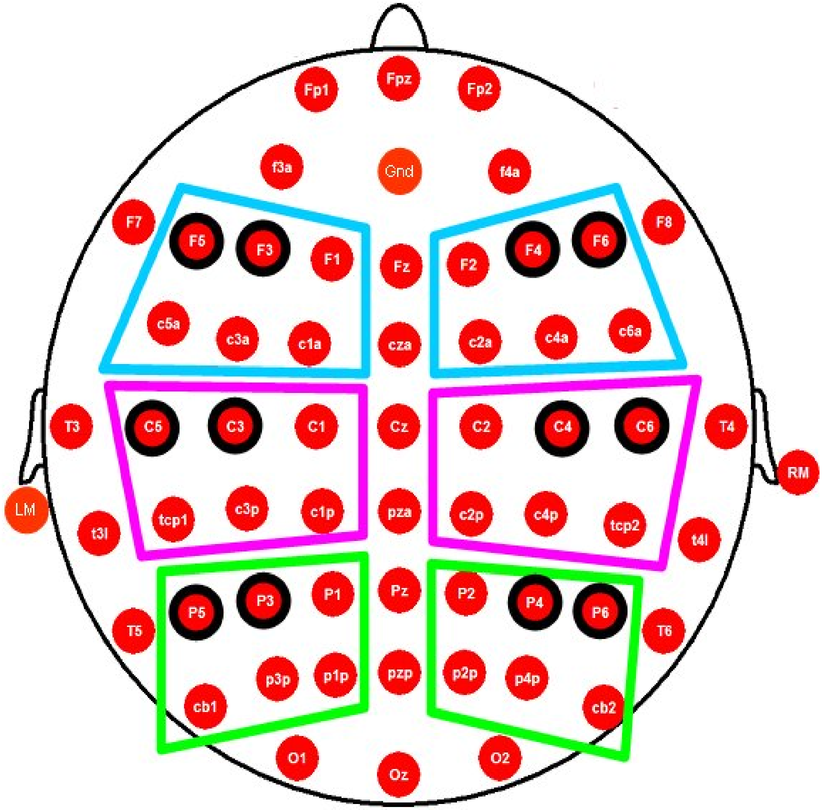
Fp1
Fpz
Fp2
f3a
Gnd
f4a
F7
F5
F3
F1
Fz
F2
F4
F6
F8
c5a
c3a
c1a
cza
c2a
c4a
c6a
T3
C5
C3
C1
Cz
C2
C4
C6
T4
RM
LM
t3l
tcp1
c3p
c1p
pza
c2p
c4p
tcp2
t4l
P5
P3
P1
Pz
P2
P4
P6
T5
T6
p3p
p1p
pzp
p2p
p4p
cb1
cb2
O1
Oz
O2

FIGURE 2

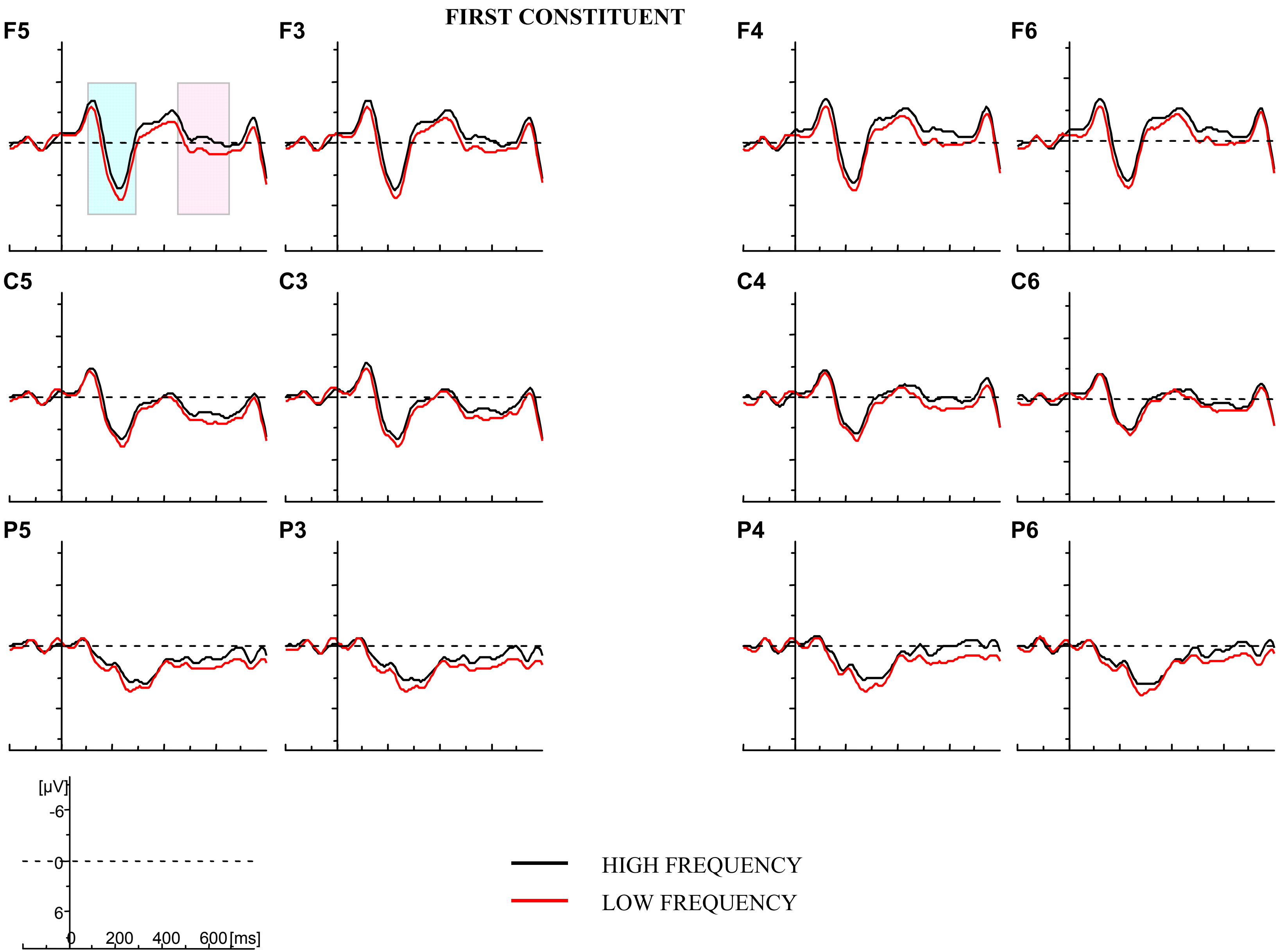
FIRST CONSTITUENT
F5
F3
F4
F6
C5
C3
C4
C6
P5
P3
P4
P6
[µV]
-6
0
6
0
200
400
600
[ms]
HIGH FREQUENCY
LOW FREQUENCY

FIGURE 3

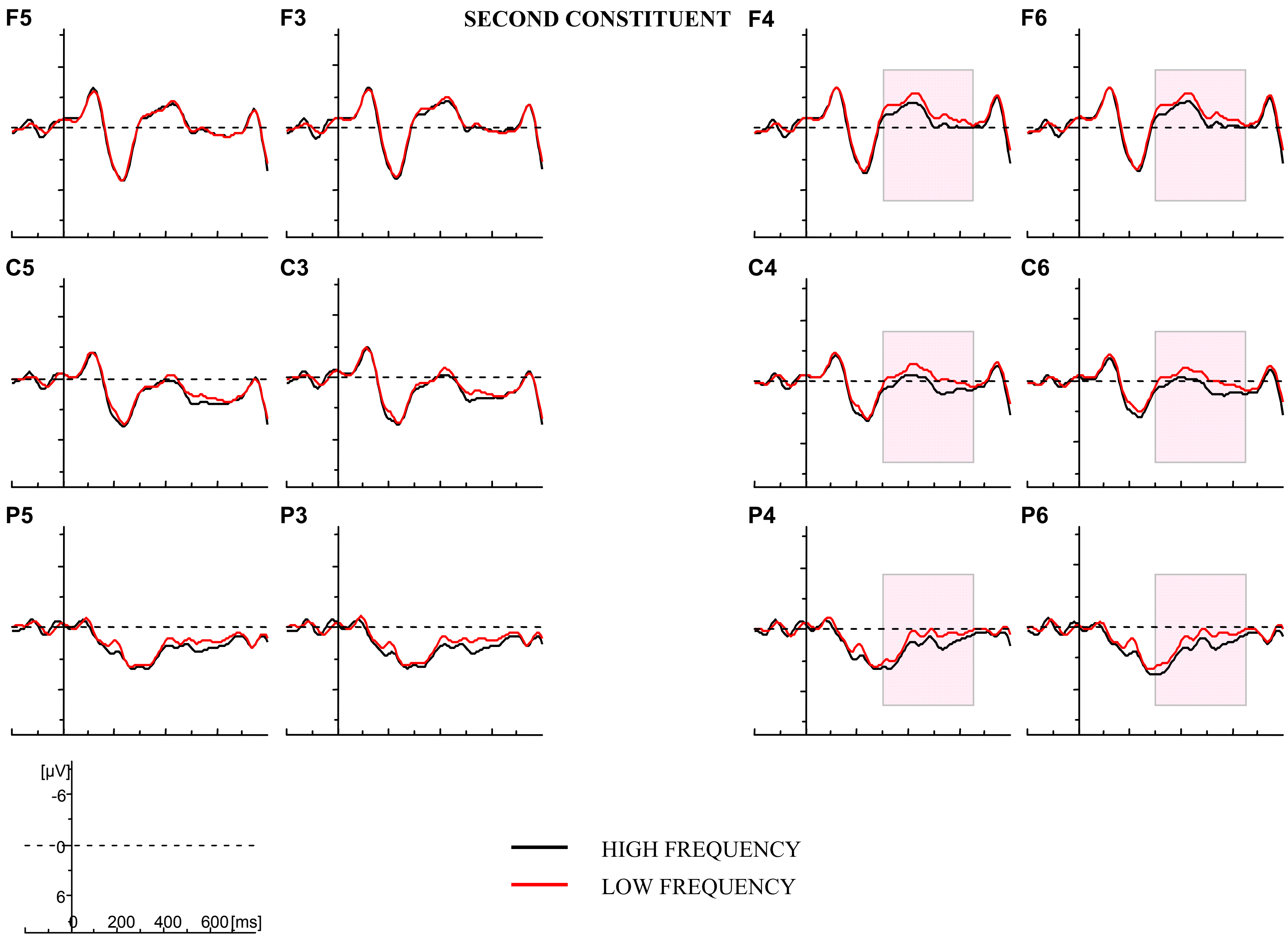
SECOND CONSTITUENT
F5
F3
F4
F6
C5
C3
C4
C6
P5
P3
P4
P6
[µV]
-6
0
6
0
200
400
600
[ms]
HIGH FREQUENCY
LOW FREQUENCY

FIGURE 4

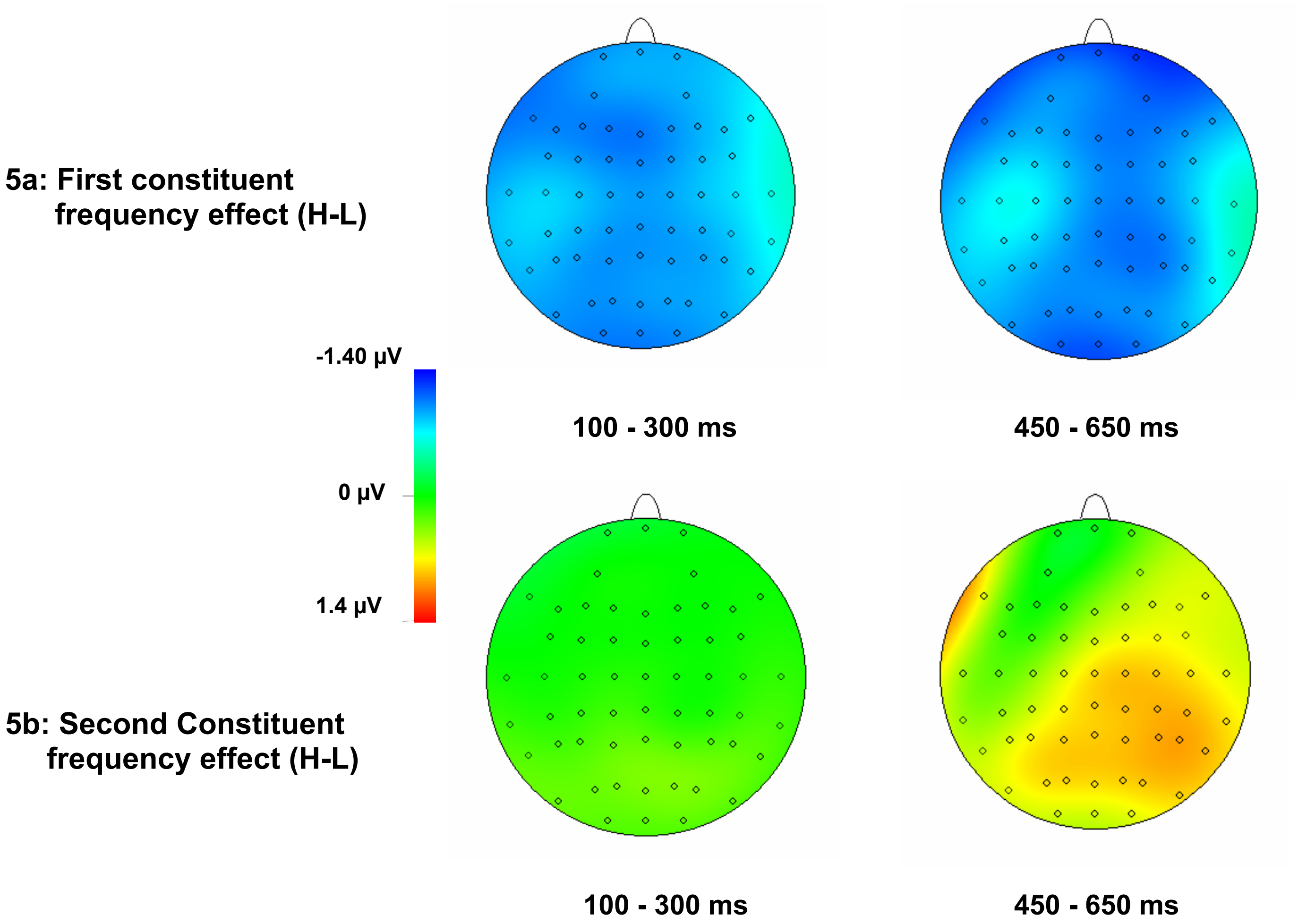
5a: First constituent frequency effect (H-L)
-1.40 µV
0 µV
1.4 µV
100 - 300 ms
450 - 650 ms
5b: Second Constituent frequency effect (H-L)
100 - 300 ms
450 - 650 ms

Table 1

Frequency (in appearances per million), length (in number of letters), number of orthographic neighbors and concreteness scores for the constituents and the compound words, in each experimental condition.

		Frequency			Length			Neighbors		Concreteness		
Condition	example	Word	C1	C2	Word	C1	C2	C1	C2	Word	C1	C2
High-High	Izenburu (title; izen=name, buru=head)	9.1	540.3	580.23	9.03	4.53	4.5	5.17	5.6	5.08	4.89	4.66
High-Low	Eskularru (glove; esku=hand, larru=skin)	9.22	549.18	91.86	9.1	4.53	4.57	4.73	5.47	5.11	5.36	4.61
Low-High	Elizgizon (priest; eliz=church, gizon=man)	8.34	84.73	577.34	9.2	4.7	4.5	4.2	4	5.26	5.28	4.88
Low-Low	Haitzulo (cave; haitz=stone, zulo=hole)	7.48	78.47	74.84	9.67	4.7	4.8	4.93	5.1	5.22	5.45	4.62

Figure

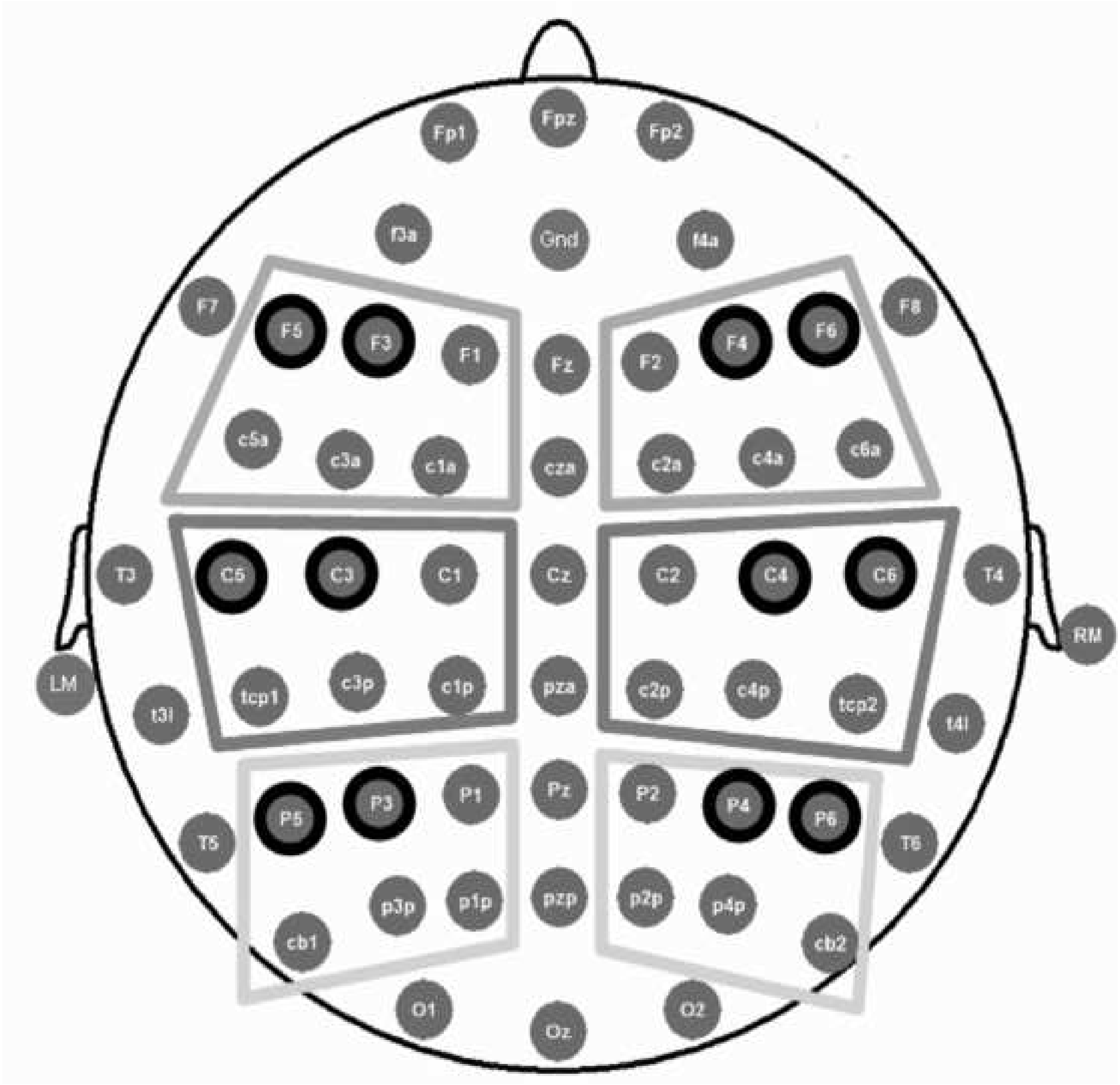
Fp1
Fpz
Fp2
f3a
Gnd
f4a
F7
F5
F3
F1
Fz
F2
F4
F6
F8
c5a
c3a
c1a
cza
c2a
c4a
c6a
T3
C5
C3
C1
Cz
C2
C4
C6
T4
RM
LM
t3l
tcp1
c3p
c1p
pza
c2p
c4p
tcp2
t4l
T5
P5
P3
P1
Pz
P2
P4
P6
T6
cb1
p3p
p1p
pzp
p2p
p4p
cb2
O1
Oz
O2

Table 1

Frequency (in appearances per million), length (in number of letters), number of orthographic neighbors and concreteness scores for the constituents and the compound words, in each experimental condition.

		Frequency			Length			Neighbors		Concreteness		
Condition	example	Word	C1	C2	Word	C1	C2	C1	C2	Word	C1	C2
High-High	Izenburu (title; izen=name, buru=head)	9.1	540.3	580.23	9.03	4.53	4.5	5.17	5.6	5.08	4.89	4.66
High-Low	Eskularru (glove; esku=hand, larru=skin)	9.22	549.18	91.86	9.1	4.53	4.57	4.73	5.47	5.11	5.36	4.61
Low-High	Elizgizon (priest; eliz=church, gizon=man)	8.34	84.73	577.34	9.2	4.7	4.5	4.2	4	5.26	5.28	4.88
Low-Low	Haitzulo (cave; haitz=stone, zulo=hole)	7.48	78.47	74.84	9.67	4.7	4.8	4.93	5.1	5.22	5.45	4.62