



HAL
open science

Spectrotemporal processing drives fast access to memory traces for spoken words

Alessandro Tavano, Sabine Grimm, Jordi Costa Faidella, Lavinia Slabu, Erich Schroger, Carles Escera

► **To cite this version:**

Alessandro Tavano, Sabine Grimm, Jordi Costa Faidella, Lavinia Slabu, Erich Schroger, et al.. Spectrotemporal processing drives fast access to memory traces for spoken words. *NeuroImage*, 2012. hal-01895794

HAL Id: hal-01895794

<https://hal.science/hal-01895794v1>

Submitted on 15 Oct 2018

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

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

Spectrotemporal processing drives fast access to memory traces for spoken words

A Tavano^{1,2,*}, S Grimm^{3,4}, J Costa Faidella^{3,4}, L Slabu^{3,4}, E. Schroger¹, C. Escera^{3,4}

Affiliations:

¹ BioCog – Cognitive Incl. Biological Psychology, Institute for Psychology, University of Leipzig, 14-20 Seeburgstr., 04103 Leipzig, Germany;

² Child Neuropsychiatry Unit, Research Institute “E. Medea”, 20 via D. L. Monza, 23842 Bosisio Parini, Lecco, Italy;

³ Cognitive Neuroscience Research Group, Department of Psychiatry and Clinical Psychobiology, University of Barcelona, P. Vall d’Hebron 171, 08035 Barcelona, Catalonia – Spain;

⁴ Institute for Brain, Cognition and Behavior (IR3C), University of Barcelona, P. Vall d’Hebron 171, 08035 Barcelona, Catalonia – Spain;

* Corresponding author: Alessandro Tavano, Ph.D., BioCog - Cognitive Incl. Biological Psychology, Institute for Psychology, University of Leipzig, 14-20 Seeburgstr., 04103 Leipzig,

Germany; Tel.: +49 (0) 341-973 59 75, Fax.: +49 (0) 341-973 59 69, E-mail: tavano@uni-leipzig.de

Keywords: Mismatch Negativity, Spectrotemporal processing, Lexical Access, Sensory memory, Long-term memory

Abstract

The Mismatch Negativity (MMN) component of the event-related potentials is generated when a detectable spectrotemporal feature of the incoming sound does not match the sensory model set up by preceding repeated stimuli. MMN is enhanced at frontocentral scalp sites for deviant words as compared to acoustically similar deviant pseudowords, suggesting that automatic access to long term memory traces for spoken words contributes to MMN generation. Does spectrotemporal feature matching drive also automatic lexical access? To test this, we recorded human auditory event-related potentials (ERPs) to disyllabic spoken words and pseudowords within a passive oddball paradigm. We first aimed at replicating the word-related MMN enhancement for Spanish language, thereby adding to the available cross-linguistic evidence (e.g., Finnish, English). We then probed its resilience to spectrotemporal perturbation by inserting short (20 ms) and long (120 ms) silent gaps between first and second syllables of deviant and standard stimuli. A significantly enhanced, frontocentrally distributed MMN to deviant words was found for stimuli with no gap. The long gap yielded no deviant word MMN, showing that expectations on word forms in a given language influence deviance detection processes. Crucially, the insertion of a short gap was sufficient to cancel deviant word MMN enhancement at frontocentral sites. We propose that spectrotemporal point-wise matching constitutes a core mechanism for fast serial computations in audition and language, bridging sensory and long-term memory systems.

1. Introduction

Audition and speech/language perception have common neural mechanisms for the analysis of acoustic stimuli (Hickok and Poeppel, 2007). The Mismatch Negativity (MMN) component of the event-related potentials indexes violation of auditory as well as speech feature regularities (Kirmse et al., 2008; Naatanen and Alho, 1995; Naatanen et al., 2001; Pakarinen et al., 2009; Schroger, 2007). The MMN system collects contextually extracted regularities to build a sensory memory trace from event history, defining a neural model of the frequent stimulus (the “standard”, Widmann et al., 2004; Winkler, 2007). Such model is used to devise concrete predictions of the spectrotemporal characteristics of incoming auditory stimuli on a point by point, real-time basis (Grimm and Schroger, 2007). A MMN response is generated to infrequent sounds carrying a deviant spectrotemporal feature, and the resulting prediction error prompts regularity model update (Garrido et al., 2008; Winkler, 2007). In a breakthrough study of the audition/speech interface, prototypical speech sounds in one’s native language delivered as deviant stimuli in a passive oddball paradigm elicited a significantly larger MMN than deviant speech sounds that were not prototypical members of that language (Naatanen et al., 1997). It was argued that MMN enhancement to language-specific deviants reflected the additional activation of long-term memory traces of native speech sounds (Diaz et al., 2008; Naatanen et al., 1997). This prompted the use of MMN amplitude modulation in language research as a neurophysiological correlate of access to memory traces for existing words compared to phonemically similar pseudowords (Pulvermuller et al., 2001; Pulvermuller et al. 2004; Shtyrov and Pulvermuller, 2002). The word-related MMN enhancement effect entails both quantitative and qualitative aspects. Deviant words are assumed to be processed to a significantly greater extent or to have stronger underlying memory circuit connections than deviant pseudowords (Shtyrov et al., 2011). In addition, a frontocentral scalp potential distribution is the usual finding for deviant word MMN, while deviant pseudoword MMN has a centroparietal peak maximum – see Shtyrov and Pulvermuller (2002). Using a cross-linguistic paradigm, Jacobsen and colleagues found that MMN amplitudes were not modulated by the contrast between language-familiar and language-unfamiliar deviant words; rather, it was the different extent of processing of familiar

vs. unfamiliar standard words, with familiar words engaging more processing resources (Jacobsen et al., 2004). Therefore, in investigating language-related regularity violation the contribution of both standard and deviant stimuli needs to be considered. The question now arises as to what type of information contained in a word's lexical structure interacts with sensory model update processes within the MMN system. On the one hand, it has been argued that pre-attentive word-related MMN enhancement may be underlined by implicit access to semantic information (Shtyrov et al., 2004). This line of research has been especially fruitful in the study of action words (e.g., action verbs), whose differences in motor meaning have been found to be mirrored by the distinct cortical somatotopies of motor actions, thus establishing a direct mapping between automatic access to action word meaning and action-specific sensorimotor circuits (Hauk et al., 2006; Pulvermuller et al., 2005; for a review, see Pulvermuller and Fadiga, 2010). This stance would suggest that potentially the entire lexical structure of a memory trace for a word - from phonetic word form to semantic content - becomes available to automatic search processes (Marslen-Wilson and Tyler, 1980). On the other hand, it is not known how the mechanisms underlying MMN elicitation manage to access stored lexical information in the first place. One possibility is that the very same mechanisms that contribute to establishing a sensory memory trace also drive lexical access. As mentioned, the high temporal resolution skills and sensitivity to contextual changes of the MMN system (Bertoli et al., 2001) have it constantly generating feed-forward hypotheses about the spectrotemporal fine structure of incoming stimuli, in the aim of reducing prediction error (Garrido et al., 2008; Grimm and Schroger, 2007; Todd and Robinson, 2010; Winkler et al., 2009). The MMN system would be equally active at all (sampled) temporal intervals in comparing incoming stimuli to a logically prior neural model of the standard. This way, each deviant feature would be linked to the time point at which it was sampled, emerging as a specific spectrotemporal relationship (Grimm and Schroger, 2007). If long-term memory information about the phonetic/phonemic form of words is sufficiently rich to allow point-by-point spectrotemporal processing in real time, then automatic lexical access might too reflect the high spectrotemporal resolution and sensitivity of the MMN system. We set out to test this hypothesis by recording the electroencephalogram (EEG) of young adult Spanish participants, using a passive orthogonal oddball design matching the physical features of critical second syllables of deviant words in the context of standard pseudowords, and deviant pseudowords in the context of standard words. We selected CVCV (Consonant-Vowel-Consonant-Vowel) stimuli, using stop consonants to design the critical second syllable structure. Spanish language prosodic and phonemic rules do not allow lengthening (i.e., *gemination*) of syllable initial stop consonants (Gerfen, 2002, Hualde et al., 2011). We assumed that participants with Spanish as native language would be highly sensitive to the effects of even minimally detectable silent gaps (spectrotemporal perturbation) when inserted before the second syllable initial stop consonant. Two gap lengths were selected: 20 and 120 milliseconds. Previous studies showed that the insertion of a 20-ms silent gap in a noise-masked sound is reliably detected by the MMN system of healthy participants in passive experimental settings (Bertoli et al., 2001). If fast lexical access relies on a point-wise spectrotemporal matching scrutiny procedure, there should be no word MMN enhancement in the 20-ms gap condition, as no such gap can be assumed to be coded in the lexical entry for that word. Accordingly, we predicted that in the 120-ms condition no word MMN enhancement should be present. The insertion of a 120-ms gap silent gap perceptually breaks word object unity, and thereby represents a manifest violation of top-down language perceptual constraints. We were interested in verifying whether in such case bottom-up deviance detection processes would still be operative (Nahum et al., 2008).

2 Methods

Participants

Fourteen healthy right-handed (handedness assessed with the criterion < 20 points on the Edinburgh 1 Handedness Inventory; Bryden, 1977), Spanish native speakers (3 males, 11 females; age range 20 – 31 years, mean: 23) volunteered to participate in the EEG experiment for paid compensation. All participants self-reported normal hearing and no history of neurologic illness or substance abuse; vision was normal or corrected-to-normal. Two participants had to be excluded because of excessive movement artifacts during EEG recording. Of the remaining 12 participants, eight were monolingual Spanish native speakers and four were early (< 2 years of age) Spanish/Catalan native bilinguals, with Spanish as their main family, life and study language. It is worth noticing at this point that the distribution of phonetic/phonotactic constraints on syllable-initial voiceless and voiced plosives is the same in both Spanish and Catalan languages, so that no language interference is expected in native bilingual individuals (Hualde et al., 2011). Since the effects of silence gap insertion on spoken word perception were not known, following an anonymous reviewer's suggestion we ran a separate behavioral experiment with different subjects to obtain measures of perceived naturalness for experimental word stimuli. Fifteen right-handed Spanish native speakers (5 males, 10 females; age range 20 – 38 years, mean. 25.5) participated for paid compensation, self-reported normal hearing and no history of neurologic illness or substance abuse, and had normal or corrected-to-normal vision. Fourteen were Spanish/Catalan native bilinguals, with Spanish as their main family, life and study language; one was a monolingual Spanish native speaker. All participants signed a written informed consent to participate in the experiment, which had been approved by the Ethical Committee of the University of Barcelona and was run in accordance with the Helsinki Declaration on Research Involving Human Subjects.

EEG experimental design

Three separate experimental recordings (corresponding to no gap, 20-ms gap, 120-ms gap conditions) were performed with every participant in one recording session. Each condition was organized into two blocks, counterbalanced across participants. Each block lasted about 15 minutes, and was followed by short breaks of 2 to 5 minutes. For the no gap condition, the Spanish words [foka] (*seal*, aquatic mammal) and [gota] (*drop*) were selected and the corresponding pseudowords [fota] and [goka] were created (see the Stimuli section below). The phonemic contrast (silent velar [k] vs. silent alveolar [t]) between standard and deviant stimuli determining the elicitation of the MMN event-related component was therefore weighted on the stop consonant of the second, critical syllable. Words and pseudowords were administered in an orthogonal paradigm to match the physical characteristics of the first syllable within block, and of the second syllable contrast within condition. Thus, the first block included word [foka] as deviant, and pseudoword [fota] as standard; the second block included pseudoword [goka] as deviant and word [gota] as standard. Therefore, by maintaining the physical contrast [ta] vs. [ka] identical and at the same time rotating the functional role of critical syllables across conditions, eventual differences in amplitude and/or topography can be solely attributable to the functional role each critical syllable subserves at any given time (see Pulvermuller et al., 2001, Shtyrov and Pulvermuller, 2002). In the second condition, an intersyllabic silent gap of 20-ms duration was added to both deviant and standard stimuli (for a graphic illustration of standard and deviant stimuli in no gap and 20-ms gap conditions for word [foka] as deviant and pseudoword [fota] as standard, see Figure 1). In the third condition, the intersyllabic silent gap was extended to 120-ms duration. Conditions were administered in a pseudorandom order. For each block, 160 deviant stimuli ($P = 0.2$) and 640 standards for a total of 800 stimuli were delivered with a constant Stimulus Onset Asynchrony of 1200 ms (Inter Stimulus Intervals = 700, 680 and 580 ms, for no gap, 20-ms gap and 120-ms gap conditions, respectively). Stimuli were organized in pseudorandom

continuous sequences with the additional constraint that at least two standard stimuli should separate a succession of two deviant stimuli, to allow the continuous building up of a neural model of the standard.

Experimental stimuli

Multiple repetitions of experimental words ([foka], [gota]) and non experimental words ([bota], *shoe*, and [boka], *mouth*) were recorded as spoken by a young adult female Spanish native speaker on a computer with a sampling rate of 44.1 kHz, in a room shielded from background noise. The software Praat (Boersma and Weenink, 2011) was used for stimulus analysis and editing. For the experiment, we selected the two initial syllables from the most acoustically similar (as for the spectrographical representation of vowel F0, F1 and F2 formant frequencies) initial syllables of experimental words ([fo] from [foka], and [go] from [gota]). Onto these we crossed-spliced the second, critical syllables extracted from the most acoustically similar second syllables of non experimental word exemplars ([ta] from [bota], and [ka] from [boka]), to avoid differential acoustic effects from first syllable context. We chose to operate a careful cross-splicing at the nearest zero crossing before the burst release of the second syllable initial stop consonant to avoid click effects as much as possible. All experimental stimuli were root-mean-square normalized as to obtain the same average sound intensity level. In the process, all syllables were individually adjusted to a length of 250 ms. Therefore, the divergence point between contextual first and critical second syllables started 250 ms after sound onset in no gap condition. Short and long intersyllabic gaps were inserted before the burst release of the word medial stop consonant (see Figure 1). Experimental pseudowords [fota] and [goka] were checked for having no meaning in Spanish, and therefore no lexical properties such as frequency of occurrence. The experimental word [gota] was found to be more frequent than [foka] on a logarithmic class of frequency count ([gota] = 13, for [foka] = 17), where numbers represent the base-2 exponent of the relative occurrence of the most frequent word in Spanish, the preposition [de] (Quasthoff, 2002; see also Alameda and Cuetos 1995, for a similar count). As higher frequency of occurrence might play a role in generating word related enhancement effect (Shtyrov et al., 2011), we employed the less frequent word ([foka]) as deviant word to control for possible confounding frequency-of-use effects on word-related MMN elicitation.

EEG recording and data processing

The Electroencephalographic activity was recorded (ANT Software bv., Netherlands) during auditory stimulation using a 33 Ag/AgCl ring-electrode extended 10-20 system mounted on a cap (EasyCap, Herrsching, Germany), against a common on-line reference at the tip of the nose, in an electrically and acoustically shielded room. Horizontal and vertical electro-oculograms were recorded by means of two additional bipolar electrodes. EEG and EOG activity was continuously recorded (passband 0.01-100 Hz, 512 Hz sampling rate). Participants were asked to refrain from excessive blinking and instructed to watch a silenced movie and completely ignore the auditory stimuli presented to them binaurally via headphones at a fixed comfortable intensity value of 75 dB SPL. EEG recordings were subject to off-line filtering (passband 1-20 Hz, Kaiser window FIR filter, length 3711), and baseline correction for epochs starting 100 ms before the divergence point, which was located 250 ms from sound onset in no gap condition, 270 ms in the 20-ms gap condition, and 370 ms in the 120-ms gap condition; epochs ended 350 ms thereafter (epoch length = -100..250 ms in all cases). The divergence point acted as the time-locking event onset across conditions. Epochs with voltage variation exceeding 80 μ V at any EEG or EOG channel were discarded from

further analysis. At least 90 accepted deviant sweeps were averaged per condition, per participant. Epochs corresponding to standard and deviant stimuli were analyzed and grand averaged separately. The MMN component was highlighted in the difference waves obtained by subtracting the response to the standard stimulus from that to the deviant stimulus. Deviant-minus standard difference responses were re-referenced to the linked mastoid leads, to sum the frontally measured response and its polarity reversed deflection below the Sylvian fissure (Schroger, 1998). The latencies at Fz of grand-average peak maxima for deviant word and deviant pseudoword difference waves were determined in each condition between 100 and 200 ms after the divergence point. They were: 420 ms from sound onset for word deviant and 393 ms for pseudoword deviant (i.e., 170 and 143 ms from the divergence point) for the no gap condition; 426 (mean between two peaks at 412 and 438 ms) and 426 (i.e., both 156 ms from the divergence point) for the 20-ms gap condition; 516 and 512 (i.e., 146 and 142 ms from the divergence point) for the 120-ms gap condition. We then averaged these measures in condition-wise manner (i.e., across the two blocks of each condition) to obtain common latency values with respect to the divergence point. They were: 406 ms (156 ms from the divergence point) for the no gap condition; 426 ms (156 ms from the divergence point) for the 20-ms gap condition; 514 ms (144 ms from divergence point) for the 120-ms gap condition. A 30-ms long window was calculated around these common latency values to obtain mean amplitudes for standard and deviant ERPs as well as for re-referenced MMN, per condition and participant. EEG/ERP data were analyzed using routines from EEProbe, Release Version 3.3.148 (ANT Software BV, Enschede, The Netherlands, www.ant-neuro.com) and the open source Matlab toolbox EEGLAB, Release Version 10.2.5.5a (www.sccn.ucsd.edu/eeglab; Delorme and Makeig, 2004).

Behavioral experiment

Participants were required to judge the naturalness of word stimuli on a six-point scale (from 1 = unnatural, to 6 = fully natural). Both experimental ([foka], [gota]) and non experimental words ([bota], [boka]) were used. Cross-splicing was applied to non experimental words. Eight different silent gap insertion levels were chosen: no gap, and 5, 10, 15, 20, 40, 80, 120 milliseconds. The resulting 32 experimental stimuli were replicated ten times in a pseudo-random order, and delivered with a constant Stimulus Onset Asynchrony of 1200 ms. The experiment lasted about 10 minutes.

Statistical Analysis

To ascertain the contribution of deviant and/or standard stimuli to MMN amplitude modulation

across conditions, repeated-measures (RM), three-way Analyses of Variance with factors Gap (three levels: no gap, 20-ms gap, 120-ms gap), Lexicality (two levels: word, pseudoword), and

Deviancy (two levels: deviant, standard) were run on ERP amplitudes at Fz and at left and right

mastoid sites. Since the topography of preattentive lexical access as indexed by MMN has been shown to vary considerably (Shtyrov and Pulvermuller, 2002), an electrode grid w 475 as selected including the responses from nine leads located at left and right frontal, central and parietal scalp sites (F3/C3/P3, Fz/Cz/Pz, F4/C4/P4) and two additional frontotemporal electrodes (F7, F8) to map the presence of MMN using a series of one-sample, two-tailed T-tests. To assess the impact of gap length on lexical status of deviant stimulus across conditions, a repeated-measures ANOVA with factors Lexicality (two levels pertaining to the lexical status of the deviant stimulus: word, pseudoword), Frontality (three levels: midline frontal, central, parietal electrodes), and Gap was run on MMN amplitudes. Given that we

were specifically interested in the topographic effects of lexical status of deviant stimulus on MMN amplitudes within each experimental condition, RM ANOVAs were run condition-wise with factors Lexicality, Frontality, and Laterality (three levels: left, central, right) using the core nine electrodes of the electrode-grid. The overall significance of scalp potential distribution differences was checked by vector-scaling each data point at a given electrode using the root mean square of all data points at all electrodes, on a single-subject, grand-average basis, using a window-of-interest-based procedure based on MMN windows as defined within each experimental condition (Picton et al., 2000). Since statistically congruent results were obtained in all cases, only results obtained from unscaled voltage data are reported. Huynh-Feldt correction was used whenever the sphericity assumption was violated. Behavioral data were analyzed by calculating the median values across word types for each gap level, and then applying the Friedmann test for non-parametric repeated-measures designs. Results with $p \leq 0.05$ were declared significant. Bonferroni correction was applied to multiple comparisons. Statistical analysis was run using PASW Statistics 18, Release Version 18.0.0 (SPSS, Inc., 2009, Chicago, IL, www.spss.com).

Results

Standard and deviant ERPs

The ERPs of twelve different stimuli were analyzed: deviant word vs. standard pseudoword and

deviant pseudoword vs. standard word in three Gap conditions: 501 no gap, 20-ms gap and 120-ms gap (see Figure 2). A significant three-way interaction (Gap x Lexicality x Deviancy) was found at Fz: $F(2, 22) = 4.017$, $\varepsilon = 1.000$, partial $\eta^2 = 0.267$, observed power = 0.655, $p = 0.033$. Follow-up two way (Lexicality x Deviancy) Anovas showed a significant interaction in the no gap condition: $F(1, 11) = 5.530$, $\varepsilon = 1.000$, partial $\eta^2 = 0.335$, observed power = 0.573, $p = 0.038$. When broken down by Deviancy, deviant word stimuli were more negative than deviant pseudoword stimuli: mean amplitude of deviant word stimuli = $-0.994 \mu\text{V}$; deviant pseudoword stimuli = $-0.413 \mu\text{V}$, $p = 0.002$. When broken down by Lexicality, only the block with deviant word stimuli showed a significant difference between standard and deviant stimuli: mean amplitude of deviant word stimuli = $-0.994 \mu\text{V}$; standard pseudoword stimuli = $-0.364 \mu\text{V}$, $p = 0.002$. In the 20-ms gap condition, a main effect of Deviancy was detected: $F(1, 11) = 9.784$, $\varepsilon = 1.000$, partial $\eta^2 = 0.471$, observed power = 0.812, $p = 0.010$, showing that overall deviants (mean = $-0.939 \mu\text{V}$) were significantly more negative than standards (mean = $-0.658 \mu\text{V}$). In the 120-ms gap condition, a significant interaction was detected: $F(1, 11) = 5.644$, $\varepsilon = 1.000$, partial $\eta^2 = 0.339$, observed power = 0.582, $p = 0.037$. When broken down by Lexicality, only the block with deviant pseudoword stimuli showed a significant difference between standard and deviant stimuli: mean amplitude of deviant pseudoword stimuli = $-0.332 \mu\text{V}$; standard word stimuli = $0.274 \mu\text{V}$, $p = 0.005$. No other significant effects were found. Results were similar at right mastoid site. A significant three-way interaction (Gap x Lexicality x Deviancy) was found on polarity-reversed mean amplitudes: $F(2, 22) = 4.148$, $\varepsilon = 1.000$, partial $\eta^2 = 0.274$, observed power = 0.670, $p = 0.030$. Follow-up two-way (Lexicality x Deviancy) Anovas showed a significant interaction in the no gap condition: $F(1, 11) = 12.800$, $\varepsilon = 1.000$, partial $\eta^2 = 0.538$, observed power = 0.902, $p = 0.004$. When broken down by Deviancy, deviant word stimuli were more positive than deviant pseudoword stimuli: mean amplitude of deviant word stimuli = $0.895 \mu\text{V}$; deviant pseudoword stimuli = $0.332 \mu\text{V}$, $p = 0.009$. When broken down by Lexicality, only the block with deviant word stimuli showed a significant difference between standard and deviant stimuli: mean amplitude of deviant word stimuli = $0.895 \mu\text{V}$; standard pseudoword stimuli = $-0.041 \mu\text{V}$, $p = 0.001$. In the 20-ms gap condition, a main effect of Deviancy was detected: $F(1, 11) = 9.128$, $\varepsilon = 1.000$, partial $\eta^2 = 0.453$, observed power = 0.785, $p = 0.012$, showing

that overall deviants (mean = $-0.939 \mu\text{V}$) were significantly more negative than standards (mean = $-0.658 \mu\text{V}$). No significant effects were found in the 120-ms gap condition. At left mastoid site, the expected main effect of Deviancy was present: $F(1, 11) = 15.910$, $\varepsilon = 1.000$, partial $\eta^2 = 0.591$, observed power = 0.952, $p = 0.002$, showing that overall deviants (mean = $1.007 \mu\text{V}$) were significantly more positive than standards (mean = $0.665 \mu\text{V}$).

Mismatch Negativity

Deviant-minus-Standard difference waveforms along the midline electrodes are presented in Figure 3. The presence of MMN is typically detected at Fz, and this is confirmed for both deviant word MMN and deviant pseudoword MMN in the no gap condition: deviant word MMN $t(1,11) = -3.901$, $p = 0.002$; deviant pseudoword MMN $t(1,11) = -2.441$, $p = 0.033$. A similar result was found at Cz: deviant word MMN $t(1,11) = -3.188$, $p = 0.009$; deviant pseudoword MMN $t(1,11) = -3.118$, $p = 0.009$. However, at Pz MMN was present only for deviant pseudoword stimuli: $t(1,11) = -2.712$, $p = 0.020$. In the 20-ms gap condition, MMN was found for deviant word stimuli at all midline leads: Fz, $t(1,11) = -2.564$, $p = 0.026$; Cz, $t(1,11) = -2.448$, $p = 0.032$; Pz, $t(1,11) = -2.184$, $p = 0.050$. Deviant pseudoword MMN displayed instead a bilateral frontotemporal distribution: F7, $t(1,11) = -2.384$, $p = 0.036$; F8, $t(1,11) = -2.982$, $p = 0.012$. In the 120-ms gap condition, no significant deflection at any electrode site was found for deviant word MMN, while deviant pseudoword MMN showed a scattered centroparietal distribution: CP5, $t(1,11) = -2.354$, $p = 0.038$; CP6, $t(1,11) = -2.194$, $p = 0.050$. Table 1 reports significance at T-tests as well as means and standard errors of the mean for MMN in all conditions. Across conditions, a significant three-way interaction (Gap x Lexicality x Frontality) was detected at Fz: $F(4, 44) = 5.039$, $\varepsilon = 0.783$, partial $\eta^2 = 0.314$, observed power = 0.944, $p = 0.002$. Follow-up two-way (Lexicality x Frontality) Anovas showed a significant interaction in the no gap condition: $F(2, 22) = 9.802$, $\varepsilon = 0.772$, partial $\eta^2 = 0.471$, observed power = 0.967, $p = 0.001$. When broken down by Frontality, a significant difference was found only at frontal sites: mean amplitude of deviant word MMN = $-1.223 \mu\text{V}$; deviant pseudoword MMN = $-0.365 \mu\text{V}$, $p = 0.036$ (see Figure 4). When broken down by Lexicality, significant differences were found across the Frontality dimension only for deviant word MMN: mean amplitude at Fz = $-1.223 \mu\text{V}$; mean amplitude at Cz = $-0.946 \mu\text{V}$, Fz vs. Cz, $p = 0.013$; mean amplitude at Pz = $-0.384 \mu\text{V}$; Fz vs. Pz, $p < 0.001$, Cz vs. Pz, $p = 0.004$. In the 20-ms gap condition, a main effect of Frontality was detected: $F(2, 22) = 4.714$, $\varepsilon = 0.792$, partial $\eta^2 = 0.300$, observed power = 0.729, $p = 0.020$, showing that across blocks Cz (mean = $-0.536 \mu\text{V}$) was more negative than Pz (mean = $-0.231 \mu\text{V}$, $p = 0.040$). No significant effects were found in the 120-ms gap condition. As for the presence of within-condition topographic effects, a tendency to significance was found for the three-way (Lexicality x Frontality x Laterality) interaction in the no gap condition: $F(4, 44) = 2.536$, $\varepsilon = 0.804$, partial $\eta^2 = 0.187$, observed power = 0.670, $p = 0.053$. The two-way interaction (Lexicality x Frontality) was instead significant: $F(2, 22) = 10.115$, $\varepsilon = 0.746$, partial $\eta^2 = 0.479$, observed power = 0.926, $p = 0.003$. At pairwise comparisons it resulted in deviant word MMN being more negative than deviant pseudoword MMN at frontal electrode sites: deviant word MMN = $-1.162 \mu\text{V}$, deviant pseudoword MMN = $-0.386 \mu\text{V}$, $p = 0.033$. In the 20-ms gap condition, a main effect of Frontality was also deemed significant: $F(2, 22) = 6.184$, $\varepsilon = 0.799$, partial $\eta^2 = 0.360$, observed power = 0.844, $p = 0.007$, suggesting that across blocks central electrode locations (mean = $-0.509 \mu\text{V}$) were more negative than parietal ones (mean = $-0.192 \mu\text{V}$, $p = 0.025$). No other significant differences were found.

The picture provided by behavioral data is consistent with the literature on silent gap insertion as well as with the present ERP/MMN results. There was a statistically significant difference in word naturalness perception depending on the length of gap insertion: $\chi^2(7) = 93.128$, $p = 0.000$. At pairwise comparisons, the 120-ms and 80-ms conditions were significantly different

from all other gap lengths except the 40-ms condition, which in turn was significantly 579 different from no gap and 5-ms conditions. This suggests the presence of a continuum in perceived naturalness of gap-inserted words whose effects become manifest when word object unity is broken but can be detected already at 10-ms silent gap length (see Figure 5).

Discussion

We investigated automatic access to memory traces for spoken words and pseudowords in the human brain as evidenced by the MMN component of the event-related potentials. A word-related enhanced MMN effect at frontal sites had been previously documented for other languages (Pulvermuller et al., 2001, 2004; Shtyrov and Pulvermuller, 2002). We replicated this finding for Spanish language: deviant words in the context of standard pseudowords elicited a significantly larger, frontocentrally distributed MMN than deviant pseudowords in the context of standard words (Shtyrov and Pulvermuller, 2002). For the first time, we were able to document a word-related MMN enhancement in a Romance language, further extending the effect's reliability across language families. ERP results at Fz suggest that in the no gap condition MMN amplitude differences were determined by differences in the lexical status of deviant ERPs only, with no significant contribution from standard ERPs (Pulvermuller et al., 2001). The lexicality effect for deviant stimuli was mirrored at right mastoid location (see Figure 2). Lexical access involves retrieval and selection processes, as the incremental nature of word recognition progressively activates competing entries in the mental lexicon. Functional activation studies have documented the existence of a bilateral pattern connecting sensory areas to frontal integrative, supramodal regions specific for spoken word recognition (Gagnepain et al., 2008; Marinkovic et al., 2003). The latter have been proposed to code for properties such as word frequency and density of competitors (Prabhakaran et al., 2006). This picture nicely fits with our finding of a bilateral, frontocentral distribution for deviant word MMN, which we envisage as an electric brain response correlate of automatic lexical access (see Figure 4, left column). Although a prominence of right- over left-sided MMN distribution was detectable at visual inspection, the effect of laterality showed only a tendency to significance. We probed the sensitivity of deviant word MMN enhancement to the insertion of intersyllabic silent gaps in both deviant and standard word/pseudoword stimuli. Crucially, a 20-ms gap was sufficient to cancel any word-related MMN enhancement at frontocentral sites, suggesting that no additional activation of long-term memory traces was integrated within the MMN system. This is confirmed by the ERP analysis, which highlighted only a main effect of deviancy across blocks. Deviant word MMN to stimuli containing a 20-ms gap was broadly distributed along the midline electrode positions, resembling pseudoword deviant MMN in the no gap condition (see Figure 4, central column). It has been suggested that when participants are presented with pseudoword stimuli, lexical search processes are mainly driven by phonetic/phonemic unit re-analysis (Zaehle et al., 2008). We hypothesize that a similar process may take place when participants are presented with word stimuli containing a short but detectable spectrotemporal perturbation. As for 20-ms gap deviant pseudoword MMN, a significant negativity enhancement was evident bilaterally at frontotemporal sites only, hinting at a selective involvement of associative speech areas. In this case, pseudoword stimuli added two sources of difficulty for speech/language recognition processes, and may have possibly required deeper articulatory reanalysis processes (Vaden, Piquado, and Hickok, 2011). No significant MMN differences were found in the 120-ms gap condition.

The pattern of results across conditions cannot be the consequence of gap insertion hindering deviance detection *per se*. Deviance detection processes were fully operative in both no gap and 20- ms gap condition as far as deviant word MMN is concerned (see Table 1). However, the absence of deviant word MMN in the 120-ms gap condition implies that overtly breaking

word-object unity – as defined by language-specific phonotactic rules – amounts to overriding deviance detection processes. Behavioral results corroborate this assumption, as words containing a 120-ms long gap were perceived as significantly less natural than words with either no gap or a 20-ms long gap (see Figure 5). As for ERPs at Fz, a significant difference between the 631 deviant and standard stimuli is present only when the deviant is a pseudoword. We take this finding to suggest the presence of a top-down influence of language-mediated expectations on word forms onto fast perceptual processes (Nahum et al. 2008). Once the process of pre-attentive access to the mental lexicon begins, high-level perceptual rules on what constitutes a word in a given language would assume primacy with respect to low-level ones (Davis and Johnsrupe, 2007; Nahum et al., 2008). Hence, breaking a high-level perceptual rule would override the detection of low-level deviant spectrotemporal events (Nahum et al., 2008). Our results argue for a distinction to be made between spectrotemporal matching to a regularity model within sensory memory, which was not significantly disturbed by gap insertion in the 20-ms gap condition, and matching to a long-term memory trace, which instead was markedly disturbed by gap insertion in the 20-ms gap condition. Spectrotemporal point-wise matching is advantageous for discriminative perception within sensory memory (Bertoli et al., 2001; Bitterman et al., 2008; Desjardins et al., 1999a), and its characterizing *modus operandi* appears to be feed-forward modeling from auditory event history statistics (Bendixen et al., 2009; Grimm and Schroger, 2007). Regularities collected from repeated standard events constitute the model and the logical prior for comparison. In matching to a long-term memory trace, a functional dissociation takes place: the incoming deviant stimulus is the prior which triggers lexical search, but the model of comparison, strengthened by acquisition history and frequency of use, is the stored word form. Stored word forms would resist update if confronted with occasional bottom-up, low-level deviant sensory evidence. For this reason, the specific trait of spectrotemporal matching applied to long-term memory can be better understood as triggering a feed-back response (Kiebel et al., 2009). However, in some specific contexts this may not hold: for example, in implicitly picking up a different accent in a new place of living, repeated exposure to phonetically deviant stimuli might change the stored long-term word form, instantiating a learning-related feed-forward mechanism, this time from sensory to long-term memory (Hickok et al., 2011). As a limitation to our study, our findings should be considered in light of the fact that we used a linguistically coded feature of Spanish language, i.e. the illegality of syllable initial stop consonant lengthening, to make a type of spectrotemporal perturbation (silent gap insertion) relevant for *both* the MMN and the linguistic systems. We would not warrant the inference that every detectable spectrotemporal disturbance would invariably impair lexical access regardless of its relevance within a given language system: different language systems may tolerate variable degrees and types of low-level disturbance (Obleser and Kotz, 2011).

Conclusions

Our results suggest that word forms are stored as detailed phonetic/phonemic representations governed by rule-based, top-down expectations on their constituents. Overall, spectrotemporal point-wise matching may dynamically serve as a procedural mechanism for fast, automatic instantiation of serial computations in audition and language, driving communication within and between memory systems.

Disclosure Statement

The authors are not aware of any conflict of interest that might be perceived as affecting the objectivity of this study.

Aknowledgments

This work was supported by grants from the Program Consolider-Ingenio 2010 (CSD2007-00012) from the Spanish Ministry of Science and Innovation, the Catalan Government (2009SGR-11) and an ICREA Academia Distinguished Professorship awarded to C. Escera, by a DFG – Reinhart- Koselleck Project grant assigned to E. Schroger, and by a grant from the Italian Ministry of Health (Ricerche correnti 2008, IRCCS “E.Medea”, Bosisio Parini, Lecco) assigned to A. Tavano. The authors would like to thank Francisco Diaz-Santaella (University of Barcelona, Catalonia, Spain) and Dr. Iria San Miguel (University of Leipzig, Germany) for help with programming, Dr. Alexandra Bendixen (University of Leipzig, Germany) for advice in performing vector-scaling and Philipp Runhau (University of Leipzig, Germany) for advice on generalized eta squared. Finally, we would like to thank Professor Pilar Prieto (Universitat Pompeu Fabra, Barcelona, Catalonia, Spain) for discussion on Spanish/Catalan phonological rules. EEG filtering routines and voltage map calculations were run with the aid of two EEGlab plug-ins written by Andreas Widmann at the University of Leipzig, Germany (<https://www.uni-leipzig.de/~biocog/content/widmann/eeglabplugins/>).

References

- Alameda, J.R., Cuetos, F. 1995. Diccionario de frecuencias de las unidades linguisticas del castellano. Servicio de Publicaciones de la Universidad de Oviedo, Oviedo, Spain.
- Bendixen, A., Schroger, E., Winkler, I. 2009. I heard that coming: event-related potential evidence for stimulus-driven prediction in the auditory system. *J. Neurosci.* 29, 8447-8451.
- Bertoli, S., Heimberg, S., Smurzynski, J., Probst, R. 2001. Mismatch negativity and psychoacoustic measures of gap detection in normally hearing subjects. *Psychophysiology* 38, 334-342.
- Bitterman, Y., Mukamel, R., Malach, R., Fried, I., Nelken, I. 2008. Ultra-fine frequency tuning revealed in single neurons of human auditory cortex. *Nature* 451, 197-201.
- Boersma, P., Weenink, D. 2011. Praat: doing phonetics by computer [Computer program]. <http://www.praat.org/>
- Bryden, M. P. 1977. Measuring handedness with questionnaires. *Neuropsychologia* 15, 617-624.
- Chandrasekaran, B., Krishnan, A., Gandour, J. T. 2007. Mismatch negativity to pitch contours is influenced by language experience. *Brain Res.* 1128, 148-156.
- Davis, M. H., Johnsruide, I. S. 2007. Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hear. Res.* 229, 132-147.
- Dehaene-Lambertz, G. 1997. Electrophysiological correlated of categorical phoneme perception in adults. *Neuroreport*, 8, 919-924.
- Desjardins, R. N., Trainor, L. J., Hevenor, S. J., Polak, C. P. 1999. Using mismatch negativity to measure auditory temporal resolution thresholds. *Neuroreport* 10, 2079-2089.
- Diaz, B., Baus, C., Escera, C., Costa, A., Sebastian-Galles, N. 2008. Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proc. Nat. Acad. Sci. U S A* 105, 16083-16088.
- Gagnepain, P., Chetelat, G., Landeau, B., Dayan, J., Eustache, F., Lebreton, K. 2008. Spoken word memory traces within the human auditory cortex revealed by repetition priming and functional magnetic resonance imagine. *J. Neurosci.* 28, 5281-5289.
- Garrido, M. I., Friston, K. J., Kiebel, S. J., Stephan, K. E., B 717 aldeweg, T., Kilner, J. M. 2008. The functional anatomy of the MMN: A DCM study of the roving paradigm. *Neuroimage* 42, 936– 944.
- Gerfen, C. 2002. Andalusian codas. *Probus* 14, 247-277.

- Grimm, S., Schroger, E. 2007. The processing of frequency deviations within sounds: evidence for the predictive nature of the Mismatch Negativity (MMN) system. *Rest. Neurol. Neurosci.* 25, 241-249.
- Hauk, O., Shtyrov, Y., Pulvermuller, F. 2006. The sound of actions as reflected by mismatch negativity: rapid activation of cortical sensory-motor networks by sounds associated with finger and tongue movements. *Eur. J. Neurosci.* 23, 811-821.
- Hickok, G., Poeppel, D. 2007. The cortical organization of speech processing. *Nature Rev. Neurosci.* 8, 393-402.
- Hickok, G., Houde, J., Rong, F. 2011. Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69, 407-422.
- Hualde, J.I., Simonet, M., Nadeu, M. 2011. Consonant lenition and phonological recategorization. *Laboratory Phonology 2*, 275-300.
- Jacobsen, T., Horvath, J., Schroger, E., Lattner, S., Widmann, A., Winkler, I. 2004. Pre-attentive auditory processing of lexicality. *Brain Lang.* 88, 54-67.
- Kiebel, S. J., von Kriegstein, K., Daunizeau, J., Friston, K. J. 2009. Recognizing Sequences of Sequences. *PLoS Comput. Biol.* 5, e1000464.
- Kirmse, U., Ylinen, S., Tervaniemi, M., Vainio, M., Schroger, E., Jacobsen, T. 2008. Modulation of the mismatch negativity (MMN) to vowel duration changes in native speakers of Finnish and German as a result of language experience. *Int. J. Psychophysiol.* 67, 131-143.
- Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., Halgren, E. 2003. Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron* 38, 487-497.
- Marslen-Wilson, W.D., Tyler, L.K. 1980. The temporal 743 structure of spoken language understanding. *Cognition* 8, 1-71.
- Naatanen, R., Alho, K. 1995. Mismatch negativity - a unique measure of sensory processing in audition. *Int. J. Neurosci.* 80, 317-337.
- Naatanen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huottilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., Alho K. 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432- 434.
- Naatanen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., Winkler, I. 2001. "Primitive intelligence" in the auditory cortex. *Trends Neurosci.* 24, 283-288.
- Nahum, M., Nelken, I., Ahissar, M. 2008. Low-level information and high-level perception: The case of speech in noise. *PLoS Biol.* 6, e126.
- Obleser, J., Kotz, S. A. 2011. Multiple brain signatures of integration in the comprehension of degraded speech. *Neuroimage* 55, 713-723.
- Pakarinen, S., Lovio, R., Huottilainen, M., Alku, P., Naatanen, R., Kujala, T. 2009. Fast multifeature paradigm for recording several mismatch negativities (MMNs) to phonetic and acoustic changes in speech sounds. *Biol. Psychol.* 82, 219-226.
- Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson, R. Jr., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., Taylor, M. J. 2000. Guidelines for using human event related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology* 37, 127-152.
- Prabhakaran, R., Blumstein, S. E., Myers, E. B., Hutchison, E., Britton, B. 2006. An event-related fMRI investigation of phonological-lexical competition. *Neuropsychologia* 44, 2209-2221.
- Pulvermuller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Alho, K., Martinkauppi, S., Ilmoniemi, R. J., Naatanen, R. 2001. Memory traces for words as revealed by the mismatch negativity. *Neuroimage* 14, 607-616.

- Pulvermuller, F., Shtyrov, Y., Kujala, T., Naatanen, R. (2004). Word-specific cortical activity as revealed by the mismatch negativity. *Psychophysiology* 41, 106-112.
- Pulvermuller, F., Shtyrov, Y., Ilmoniemi, R. 2005. Brain signatures of meaning access in action word recognition. *J. Cogn. Neurosci.* 17, 884-892.
- Pulvermuller, F., Fadiga, L. 2010. Active perception: sensorimotor circuits as a cortical basis for language. *Nature Rev. Neurosci.* 11, 351-360.
- Quasthoff, U. 2002. Deutscher Wortschatz im Internet [Online database, Spanish language Corpus], <http://corpora.informatik.uni-leipzig.de/>, Leipzig, Germany: University of Leipzig.
- Schroger, E. 1998. Measurement and interpretation of the mismatch negativity. *Behav. Res. Meth.Instr. Comput.* 30, 131-145.
- Schroger, E. 2007. Mismatch negativity - A microphone into auditory memory. *J. Psychophysiol.* 21, 138-146.
- Shtyrov, Y., Pulvermuller, F. 2002. Neurophysiological evidence of memory traces for words in the human brain. *Neuroreport* 13, 521-525.
- Shtyrov, Y., Hauk, O., Pulvermuller, F. 2004. Distributed neuronal networks for encoding category specific semantic information: the mismatch negativity to action words. *Eur. J. Neurosci.* 19, 1083-1092.
- Shtyrov, Y., Kimppa, L., Pulvermuller, F, Kujala, T. 2011. Event-related potentials reflecting the frequency of unattended spoken words: A neuronal index of connection strength in lexical memory circuits? *Neuroimage* 55, 658-668.
- Todd, J., Robinson, J. 2010. The use of conditional inference to reduce prediction error--a mismatch negativity (MMN) study. *Neuropsychologia* 48, 3009-3018.
- Vaden, K. I., Piquado, T., Hickok, G. 2011. Sublexical Properties of Spoken Words Modulate Activity in Broca's Area but Not Superior Temporal Cortex: Implications for Models of Speech Recognition. *J. Cogn. Neurosci.* 23, 2365-2374.
- Widmann, A., Kujala, T., Tervaniemi, M., Kujala, A., Schroger, E. 2004. From symbols to sounds: visual symbolic information activates sound representations. *Psychophysiology* 41, 709-715.
- Winkler, I. 2007. Interpreting the mismatch negativity (MMN). *Psychophysiology* 21, 60-69.
- Winkler, I., Denham, S. L., Nelken, I. 2009. Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends Cogn. Sci.* 13, 532-540.
- Zaehle, T., Geiser, E., Alter, K., Jancke, L., Meyer, M. 2008. Segmental processing in the human auditory dorsal stream. *Brain Res.* 1220, 179-190.

Table 1. Presence of MMN across selected e 801 electrodes in all conditions

	Word deviant MMN	t	p-Value	Mean (SE)	Pseudoword deviant MMN	t	p-Value	Mean (SE)
No gap	F7	-2.650	0.023	-.74 (.279)	F7	-1.289	0.224	-.29 (.227)
	F3	-3.424	0.006	-1.08 (.316)	F3	-2.455	0.032	-.42 (.171)
	Fz	-3.901	0.002	-1.22 (.314)	Fz	-2.441	0.033	-.37 (.150)
	F4	-5.652	0.000	-1.18 (.209)	F4	-2.613	0.024	-.37 (.142)
	F8	-7.976	0.000	-.83 (.105)	F8	-1.699	0.117	-.28 (.165)
	C3	-3.176	0.009	-.89 (.279)	C3	-3.424	0.006	-.71 (.208)
	Cz	-3.188	0.009	-.95 (.297)	Cz	-3.188	0.009	-.75 (.235)
	C4	-5.232	0.000	-.87 (.166)	C4	-2.433	0.033	-.41 (.167)
	P3	-2.472	0.031	-.45 (.183)	P3	-2.668	0.022	-.45 (.169)
	Pz	-1.896	0.085	-.38 (.202)	Pz	-2.712	0.020	-.55 (.202)
	P4	-2.324	0.040	-.33 (.143)	P4	-2.594	0.025	-.39 (.149)
	20-ms gap	F7	-1.796	0.100	-.39 (.215)	F7	-2.384	0.036
F3		-2.559	0.027	-.63 (.245)	F3	-1.704	0.116	-.44 (.258)
Fz		-2.564	0.026	-.68 (.264)	Fz	-1.689	0.119	-.45 (.264)
F4		-2.426	0.034	-.57 (.236)	F4	-1.912	0.082	-.44 (.228)
F8		-2.014	0.069	-.40 (.199)	F8	-2.982	0.012	-.59 (.198)
C3		-2.708	0.020	-.66 (.242)	C3	-1.708	0.116	-.37 (.216)
Cz		-2.448	0.032	-.68 (.277)	Cz	-1.404	0.188	-.39 (.280)
C4		-2.612	0.024	-.54 (.205)	C4	-2.080	0.062	-.42 (.202)
P3		-2.600	0.025	-.44 (.170)	P3	.489	0.635	.09 (.183)
Pz		-2.184	0.050	-.55 (.254)	Pz	.441	0.668	.09 (.207)
P4		-1.663	0.125	-.39 (.232)	P4	.342	0.739	.05 (.137)
120-ms gap		F7	-.853	0.412	-.26 (.301)	F7	-.283	0.782
	F3	-1.358	0.202	-.55 (.403)	F3	-1.561	0.147	-.43 (.274)
	Fz	-1.305	0.219	-.58 (.448)	Fz	-1.733	0.111	-.58 (.332)
	F4	-1.309	0.217	-.55 (.421)	F4	-1.779	0.103	-.51 (.288)
	F8	-1.027	0.327	-.36 (.355)	F8	-.962	0.357	-.26 (.271)
	C3	-1.933	0.079	-.63 (.328)	C3	-2.560	0.027	-.56 (.219)
	Cz	-2.063	0.064	-.74 (.358)	Cz	-1.917	0.082	-.51 (.266)
	C4	-1.723	0.113	-.57 (.330)	C4	-1.711	0.115	-.44 (.258)
	P3	-.989	0.344	-.24 (.239)	P3	-1.675	0.122	-.35 (.211)
	Pz	-1.209	0.252	-.30 (.246)	Pz	-1.449	0.175	-.39 (.267)
	P4	-1.089	0.299	-.28 (.258)	P4	-1.971	0.074	-.41 (.210)

Significant p values are highlighted in bold characters.

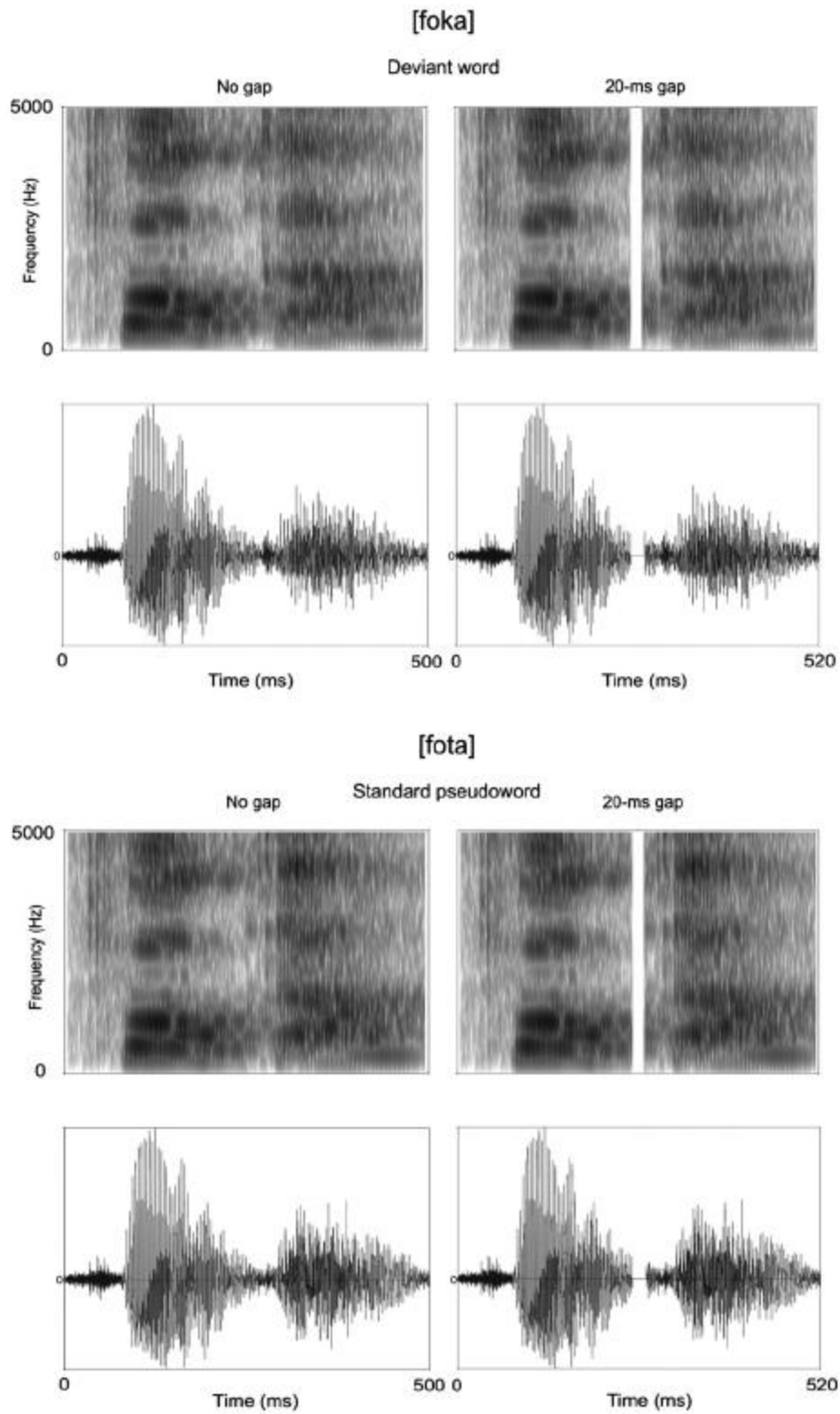


Figure 1. Exemplar spectrograms (upper row) and waveforms (lower row) of deviant word [foka] and standard pseudoword [fota] for no gap and 20-ms experimental conditions.

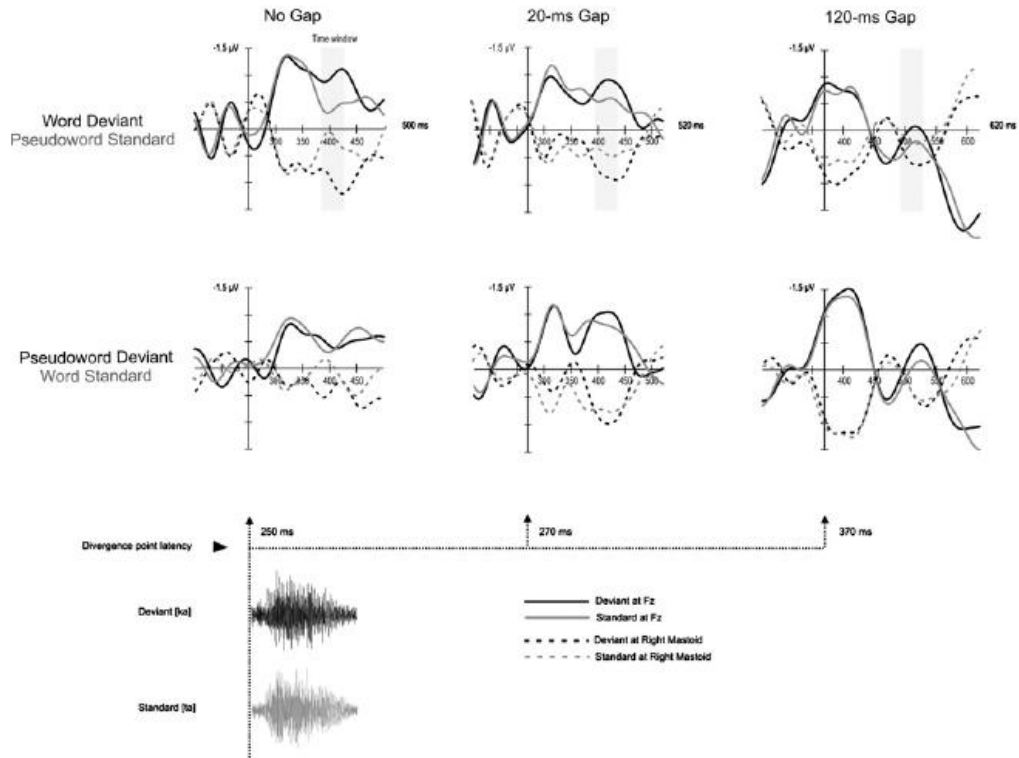


Figure 2. ERPs at Fz (continuous line) and Right Mastoid (dashed line) for each Experimental Condition (no Gap, 20-ms gap, 120-ms gap), across Type of Deviant (Word versus Pseudoword).

Figure 3. Difference Waves at midline electrodes for each Experimental Condition (no Gap, 20-ms gap, 120-ms gap), across Type of Deviant (Word versus Pseudoword).

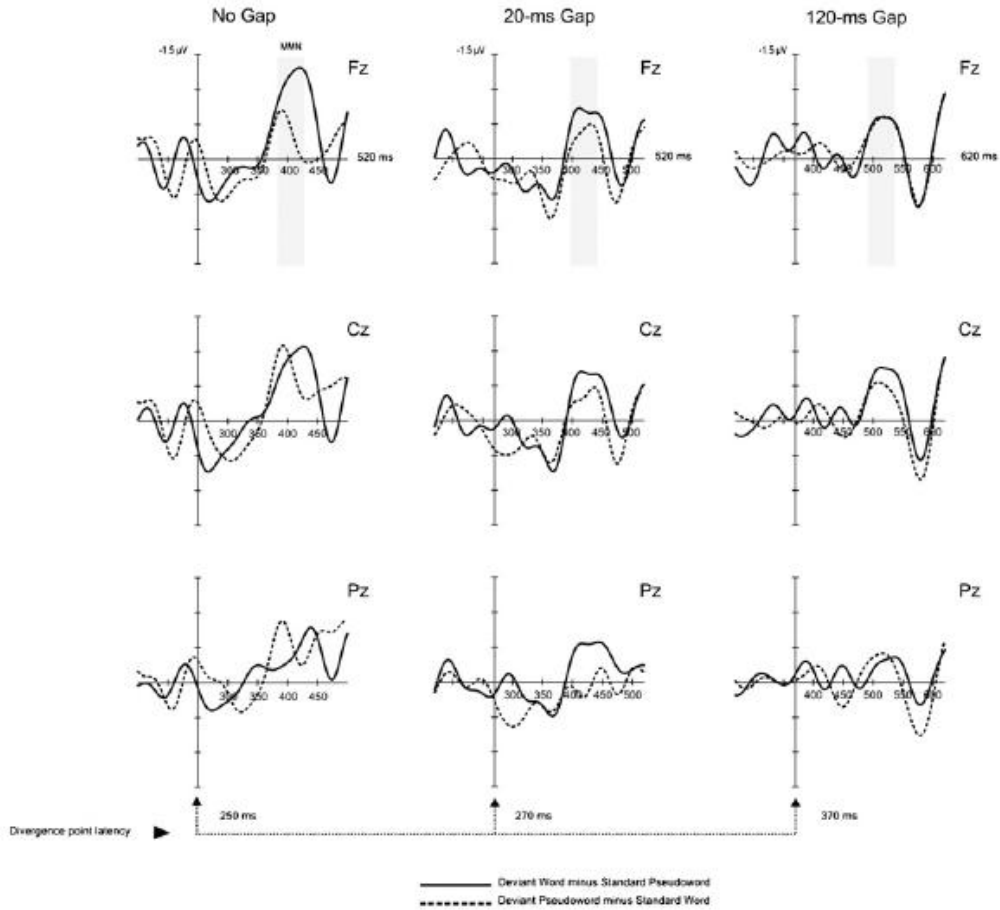


Figure 3. Difference Waves at midline electrodes for each Experimental Condition (no Gap, 20-ms gap, 120-ms gap), across Type of Deviant (Word versus Pseudoword).

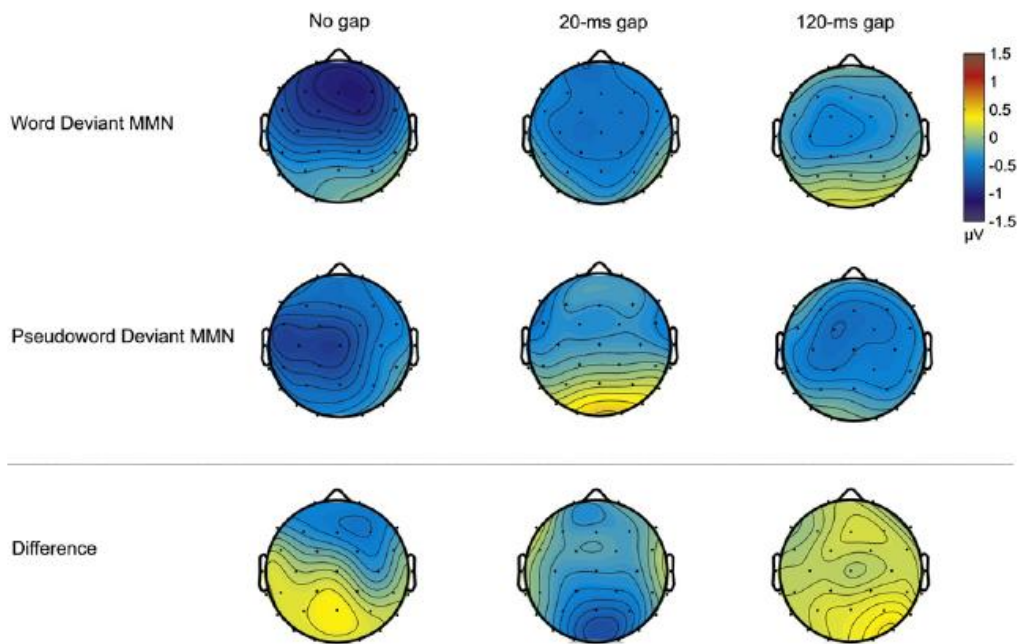


Figure 4. Normalized scalp voltage maps of MMN for each experimental condition (no Gap, 20-ms gap, 120-ms gap), across Type of Deviant (Word versus Pseudoword). Maps of the difference between Type of Deviant blocks are also shown.

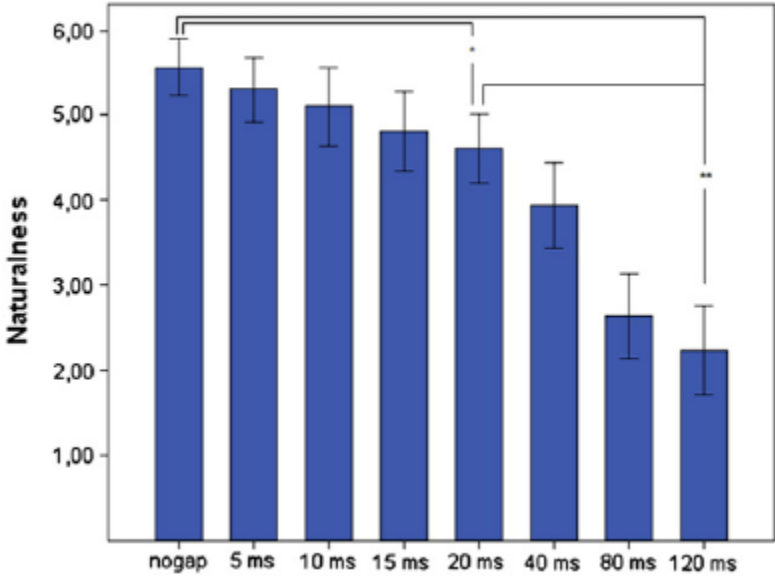


Figure 5. Naturalness perception judgments (from 1 = unnatural to 6 = fully natural) according to silent gap length.