Brain & Language 116 (2011) 51–63

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Interplay between acoustic/phonetic and semantic processes during spoken sentence comprehension: An ERP study

Véronique Boulenger[a,*], Michel Hoen[b], Caroline Jacquier[a], Fanny Meunier[a]

[a] Laboratoire Dynamique du Langage, CNRS, Université Lyon 2, UMR 5596, Lyon, France
[b] Stem Cell and Brain Research Institute, INSERM U846, Université Lyon 1, Lyon, France

ARTICLE INFO

Article history:
Accepted 18 September 2010
Available online 20 October 2010

Keywords:
Sentence comprehension
Connected speech
Degraded speech
Event-related potentials
Mismatch Negativity (MMN)
N400

ABSTRACT

When listening to speech in everyday-life situations, our cognitive system must often cope with signal instabilities such as sudden breaks, mispronunciations, interfering noises or reverberations potentially causing disruptions at the acoustic/phonetic interface and preventing efficient lexical access and semantic integration. The physiological mechanisms allowing listeners to react instantaneously to such fast and unexpected perturbations in order to maintain intelligibility of the delivered message are still partly unknown. The present electroencephalography (EEG) study aimed at investigating the cortical responses to real-time detection of a sudden acoustic/phonetic change occurring in connected speech and how these mechanisms interfere with semantic integration. Participants listened to sentences in which final words could contain signal reversals along the temporal dimension (time-reversed speech) of varying durations and could have either a low- or high-cloze probability within sentence context. Results revealed that early detection of the acoustic/phonetic change elicited a fronto-central negativity shortly after the onset of the manipulation that matched the spatio-temporal features of the Mismatch Negativity (MMN) recorded in the same participants during an oddball paradigm. Time reversal also affected late event-related potentials (ERPs) reflecting semantic expectancies (N400) differently when words were predictable or not from the sentence context. These findings are discussed in the context of brain signatures to transient acoustic/phonetic variations in speech. They contribute to a better understanding of natural speech comprehension as they show that acoustic/phonetic information and semantic knowledge strongly interact under adverse conditions.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

One of the most challenging situations that every listener has to deal with is understanding speech. Under ecological conditions, speech is often perceived in acoustically unstable environments, where other conversations, physical noise or reverberations can occur unexpectedly. Even talkers create transient signal instabilities by inserting sudden unpredictable breaks, involuntary voice modulations or noises into their production. Still, our cognitive system is most of the time able to overcome such degradations. When we are listening to someone talking, our brain seems particularly efficient at generating expectancies about the ongoing speech stream from the capture of regularities in the signal. These expectancies seem to be generated at very different, if not at all, levels of speech processing. Many studies have identified clear mechanisms extracting contextual regularities from speech and generating expectancies at levels as various as rhythmic, syntactic, semantic or pragmatic aspects (Obleser & Kotz, 2010; Rothermich, Schmidt-Kassow, Schwartze, & Kotz, 2010; Schmidt-Kassow & Kotz, 2009; see for example Friederici (2002) and Kutas and Federmeier (2007) for reviews). Of course, these expectancies help our system to: (i) proactively anticipate signal characteristics at multiple levels in order to recognize non-awaited events faster and (ii) eventually replace missing or distorted information parts by their expected counterpart if speech signals appear to be too degraded to be efficiently exploited. Multiple higher-level expectancy-generation mechanisms dedicated to semantic or syntactic aspects of the signal, together with the corresponding procedures of violation detection have been well identified. However, despite the crucial importance of lower-level acoustic/phonetic abilities for speech comprehension, the brain mechanisms involved in the real-time detection of sudden acoustic/phonetic distortions within a continuous speech stream remain partially unknown. We actually still need to unravel when and how the brain detects changes of the ongoing speech signal at a low-level, namely at the acoustic/

---

* Corresponding author. Address: Laboratoire Dynamique du Langage CNRS UMR 5596, Institut des Sciences de l'Homme, 14 avenue Berthelot, 69363 Lyon Cedex, France. Fax: +33 (0) 4 72 72 65 90.

E-mail address: Veronique.Boulenger@ish-lyon.cnrs.fr (V. Boulenger).

0093-934X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandl.2010.09.011

phonetic interface, and whether and how this impacts higher-level processes such as for example semantic integration of words into their context, and ultimately speech comprehension. The present study aimed at tackling this issue by investigating whether the brain can extract regularities from connected speech to rapidly form a strong memory trace that can be used as a template to serve fast and automatic detection of transient perturbations in the ongoing speech stream. We also assessed how these early mechanisms at the interface between acoustic and phonetic processes interact with later processes involved in contextual integration. To this aim, we explored the temporal dynamics of cortical responses, as evaluated by the recording of event-related brain potentials (ERPs), associated with the processing of increasingly manipulated portions of speech embedded in sentences.

Previous electrophysiological studies have identified one major evoked component reflecting the detection of any sudden discriminable change in some regular aspect of the ongoing auditory stream, the Mismatch Negativity (MMN; Näätänen & Alho, 1995; Näätänen, Gaillard, & Mäntysalo, 1978). MMN is a fronto-central negative wave peaking between 100 and 250 ms after stimulus onset and thought to index memory traces formed in the supratemporal auditory cortex. It is classically elicited in the so-called ‘‘oddball paradigm” in which an infrequent sound (the ‘‘deviant”) occurs in a series of ‘‘standard” stimuli, irrespective of the subject’s attention or task. MMN has been reported to be insensitive to the predictable occurrence of the deviant within the sequence (Scherg, Vajsar, & Picton, 1989; Sussman, Ritter, & Vaughan, 1998) and to be modulated by the magnitude of the deviance, i.e. the larger the deviance, the larger the MMN amplitude and the shorter its latency (Kujala, Kallio, Tervaniemi, & Näätänen, 2001; Pakarinen, Takegata, Rinne, Huotilainen, & Näätänen, 2007; but see Horvath et al., 2008). Interestingly, it has also been shown that the ‘‘standard” repetitive stimulus does not have to be a simple sound for MMN to be elicited as this response can be observed for transient modifications in sound patterns as complex as speech (Aaltonen, Niemi, Nyrke, & Tuhkanen, 1987; Kraus, McGee, Sharma, Carrell, & Nicol, 1992). Studies on the auditory processing of language have further demonstrated the usefulness of MMN in assessing linguistic processes at different cognitive levels, namely phonological, lexical, semantic and syntactic (for a review, see Pulvermüller & Shtyrov, 2006). For instance, MMN is elicited in response to native compared to non-native phonetic deviants (Dehaene-Lambertz, 1997) and it is modulated by the lexical status of the stimuli (Korpilahti, Krause, Holopainen, & Lang, 2001; Shtyrov & Pulvermuller, 2002). MMN is also sensitive to semantic factors such as the meaning of deviant words (Menning et al., 2005; Shtyrov, Hauk, & Pulvermüller, 2004) and to the grammaticality of word strings (Pulvermüller & Shtyrov, 2003; Shtyrov, Pulvermüller, Näätänen, & Ilmoniemi, 2003). Whether complete sentences can constitute an acoustic context that carries enough regular information (i.e. invariant context) to elicit an MMN whenever a perturbation of the signal occurs is still a matter of debate. It is actually still not known whether the neural system underlying MMN generation can establish natural speech input as a ‘‘standard” or template – just as it does for repetitive tones, syllables or single words – and build up a strong memory trace of this information against which deviants may be compared. The notion of ‘standard’ in oddball paradigms recently moved from the classical view of one acoustic stimulation explicitly embodying the standard stimulus to implicit forms of standards extracted from the stable acoustic aspects of stimuli otherwise varying along different acoustic dimensions (*e.g.* frequency, duration, intensity; Pakarinen, Huotilainen, & Näätänen, 2010). Previous studies have indeed shown that MMN is elicited for deviants that violate complex acoustic regularities such as ‘‘the higher the frequency, the louder the intensity” or ‘‘a long sound is followed by a high sound” (Paavilainen, Jaramillo, Näätänen, & Winkler, 1999; Paavilainen, Simola, Jaramillo, Näätänen, & Winkler, 2001; Saarinen, Paavilainen, Schröger, Tervaniemi, & Näätänen, 1992). Shestakova et al. (2002) also demonstrated MMN response to vowel deviants presented among a sequence of 450 standard vowels each uttered by a different speaker, suggesting that memory traces for specific phoneme categories were formed despite continuous acoustic variation of the speech sounds. Hence, the standard stimuli sequence does not have to be acoustically constant for MMN to be generated as long as some pattern or rule is shared by the standards. This suggests that the brain encodes and transiently stores information about regular interstimulus relationships and then compares incoming sounds to these representations (Ritter, Gomes, Cowan, Sussman, & Vaughan, 1998; Winkler, Cowan, Csépe, Czigler, & Näätänen, 1996). Very recent observations further show that our auditory system is able to form memory traces for regular aspects of complex sounds with an extremely fast and efficient procedure, allowing the extraction of standard portions of sounds only after a few seconds of exposure to novel sounds (Agus, Thorpe, & Pressnitzer, 2010). It therefore appears that automatic sensory processes as those reflected by the MMN may play a role in identifying regular aspects of connected speech signals, allowing the generation of low-level predictions about the ongoing speech stream in order to accurately react to unexpected transient variations. So far however, MMN generation in the context of connected speech processing has not been observed. In the present study, we sought to determine whether spoken sentences are represented in a transient auditory memory as regular, invariant patterns encompassing not only sensory (acoustic) but also higher-level phonetic, categorical information. In other words, we assessed whether the central auditory mechanisms that underlie MMN can extract large-scale ‘‘abstract” regularities in sentences so that any distortions from the established sentence neuronal traces are reflected by an MMN.

In a recent study, Menning et al. (2005) demonstrated that semantic and syntactic deviant spoken sentences among standard semantically and syntactically correct sentences elicited a mismatch response. They suggested that automatic comparison of the input against the expected correct continuation of the sentence provoked an MMN each time the speech signal did not fit this expectation. Recent experiments also suggest that MMN could play a role in speech-in-noise or distorted speech comprehension (Kozou et al., 2005; Muller-Gass, Marcoux, Logan, & Campbell, 2001). For instance, Kozou et al. (2005) reported that the MMN to syllables is differently affected by the type of competing background noise, its amplitude being smaller in the presence of a fluctuating noise such as babble or industrial noise than with a wideband noise. Yet the possibility of a direct involvement of MMN in spoken sentence comprehension has barely been addressed, although recent studies have examined the brain’s response to processing distorted acoustic information in sentential contexts (Aydelott, Dick, & Mills, 2006; Besson, Faita, Czternasty, & Kutas, 1997; Sivonen, Maess, Lattner, & Friederici, 2006). These investigations further allowed addressing the issue of the interaction between early acoustic processes and late semantic integration. Processing of word lexico-semantic information is reflected in the N400, a negative deflection peaking around 400 ms after word onset (Kutas & Hillyard, 1984; see Kutas & Federmeier, 2000 and Lau, Phillips, and Poeppel (2008) for reviews). The N400 is highly sensitive to semantic context: the more words are incongruent with a preceding word or sentence context, the larger the N400 amplitude (Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007). This potential has therefore been proposed to index contextual integration, namely the ease or difficulty (i.e. processing cost) with which words are integrated into their semantic context (Brown & Hagoort, 1993). In this view, the N400 would correspond to combinatorial mechanisms that occur after lexical access. However, an alternative account suggests that the N400 could reflect facilitated access of word lexico-semantic information from long-term memory (Federmeier, 2007; Kutas & Federmeier, 2000). Amplitude of the N400 is indeed modulated by lexical factors such as word frequency (Allen, Badecker, & Osterhout, 2003; Van Petten & Kutas, 1990) and is reduced for incongruent words that share semantic features with expected words (Kutas & Federmeier, 2000; Van Petten, Coulson, Rubin, Plante, & Parks, 1999). This suggests that the N400 cannot be attributed only to post-access processes but that it could also index predictive processes. In other words, semantic context could be used to anticipate and prepare for expected forthcoming words by retrieving their perceptual and semantic features from semantic memory (see Lau et al. (2008) for a review). Although the issue of the exact nature of the neural processes underlying N400 is still debated, it thus seems that the language system would benefit from both integrative and predictive strategies to understand words in context (Kutas & Federmeier, 2000). In a study aimed at examining the effects of acoustic degradation on semantic processes, Aydelott et al. (2006) showed that an early negative peak, labeled "N1 (MMN)" (p. 462), was elicited when sentence-final words, congruent or not with the preceding context, were presented in low-pass filtered context compared to intact context. They interpreted this perceptual effect as evidence that filtered speech set up a particular acoustic context that created a mismatch to the unfiltered target. Their results also revealed that the acoustic degradation modulated the N400, its amplitude being attenuated to incongruent targets in filtered contexts. This suggests that acoustic degradation reduced availability of semantic information and thus produced fewer demands on semantic integration for incongruent words. Sivonen et al. (2006) found comparable results in a study where the first phoneme of sentence-final words was replaced with a cough-noise. A strong N1 response to the onset of the cough was observed, its amplitude being modulated by the duration of the noise (the longer the cough, the larger the N1). This early response was assumed to reflect the automatic detection of the interfering noise which obliterated the word's onset. This was followed by a modulation of the N400 latency when the word was masked with the cough.

Despite these studies suggesting that detection of an acoustic perturbation within a sentence is reflected in the brain by an early negative wave, further compelling evidence is needed to determine whether this component is comparable to the classical MMN elicited to acoustic changes within an auditory stream. This is of particular interest as it would add to previous literature that MMN is involved in language processing at various linguistic levels and that it could constitute an automatic response that may have direct implications in speech comprehension, particularly under adverse conditions. The present study directly addressed this issue by investigating the cortical responses to the early detection of an acoustic/phonetic variation occurring in connected speech and how these processes interact with later stages underlying semantic integration and speech comprehension. We particularly aimed at answering two questions: (i) Does a sudden signal change at the acoustic/phonetic level within a continuous speech stream elicit an MMN, reflecting violation of expectations generated from regularities in the signal? And if so, is MMN amplitude modulated by the magnitude of the manipulation? (ii) Does the early change detection affect contextual integration of words into their context? Participants were engaged in a sentence repetition experiment where acoustic/phonetic (time reversal) and semantic (cloze probability) features were systematically manipulated. We chose to use time reversal to avoid adding an extraneous noise to the target signal which could elicit other confounding effects. Time reversal distorts the temporal structure of speech while preserving its spectral properties (Saberi & Perrott, 1999) and can be seen as an acoustic/phonetic distortion. As an acoustic distortion, it alters the physical nature of the stimulus, for instance the temporal course of a reverberant sound and the perception of its time and intensity (*e.g.* DiGiovanni & Schlauch, 2007; Stecker & Hafter, 2000). As a phonetic distortion, it can give rise to abnormal transitions between phonemes (*e.g.* distortion for rapidly changing sounds such as stop consonants) and to unusual phonemic temporal envelopes (altering the perception of the duration of continuant phonemes; Pellegrino, Ferragne, & Meunier, 2010). Here we hypothesized that an early negative ERP reflecting rapid and automatic detection of the acoustic/phonetic change within spoken sentences should be observed. To precisely assess whether this response matched the well-known MMN reflecting violation of regularities in an auditory sequence, we compared it in terms of spatio-temporal characteristics to an MMN recorded in the same participants during a classical oddball paradigm. We also expected the two types of manipulations (time reversal and cloze probability) to influence late ERPs related to semantic integration of words in their context (N400).

2. Materials and methods

2.1. Participants

Twenty healthy native French speakers aged 18–25 years (mean = 21, SD = 2) participated in the experiment. All were right-handed (mean score Edinburgh inventory = 86, SD = 13; Oldfield, 1971), had no hearing problems (peripheral auditory thresholds below 20 dB HL) and had normal or corrected-to-normal vision. They had no record of neurological diseases and reported no history of drug abuse. All subjects gave their written informed consent to participate in the experiment and were paid for their participation.

2.2. Stimuli

2.2.1. Linguistic oddball experiment

The French consonant–vowel syllable /ba/ was recorded by a French native female speaker (duration = 297 ms, 22 kHz, mono, 16 bits). The syllable could either be kept intact (forward speech) or be reversed along its temporal axis (reversed speech), starting from the onset, using Praat software.

2.2.2. Sentence repetition experiment

Two hundred sentences 7–10 words in length (mean = 8.05, SD = .66) were recorded by the same French native female speaker (22 kHz, mono, 16 bits, adjusted at an equivalent intensity of 60 dB-A). All sentences followed the same global structure: Determiner – Noun 1 – Verb – Determiner – Noun 2 – Preposition – Determiner – Noun 3. All nouns in the sentences were bi-syllabic and Noun 3, always starting with a consonant, constituted the target word. Cloze probability (CP) of the target word within the sentence context, which refers to the probability that this particular word will be produced as being the most likely completion of a sentence fragment (Taylor, 1953), was manipulated. For half of the sentences, the target word had a low-CP (*e.g.* "*Le coureur franchit une rangée de cactus*", literally "The sprinter jumped over a row of cactus") whereas for the other half, CP of the target word was high (*e.g.* "*Le chanteur vend des billets pour son concert*", literally "The singer sells tickets for his concert"). Cloze probability was pre-checked in an offline task where 25 French native participants (different from the participants of the experiment) were asked to read and complete each sentence, from which the last word was omitted, with the first word that came to their mind. Results of this pre-test confirmed that half of the sentences contained a final word with a low-CP ($p < .05$; mean = .016, SD = .051) and the other half a final word with a high-CP ($p > .05$; mean = .68, SD = .21).

The 200 sentences were divided into five lists of 40 items each (20 with low-CP target word and 20 with high-CP target word). Each list contained every sentence only once to avoid repetition effects and was seen by four participants. Final target words were matched for word frequency (mean = 21.78 occurrences per million, SD = 8.28), number of phonological neighbors (mean = 11.81, SD = 4.48) and number of phonemes (mean = 4.87, SD = .09) across lists and between low- and high-CP sentences ($p > .05$) using the French lexical database Lexique (New, Pallier, Brysbaert, & Ferrand, 2004). Within each list, target words could either be kept intact (forward speech) or be reversed along their temporal axis (reversed speech), starting from their onset, using Praat software. The length of the time reversal window varied from 0 (R0; no reversal), 0.5 (R0.5; reversal of half of the first syllable; mean duration = 75 ms), 1 (R1; reversal of the first syllable; mean = 152 ms), 1.5 (R1.5; reversal of the first syllable and half of the second; mean = 262 ms) to 2 syllables (R2; mean = 372 ms). Boundaries between syllables were always taken at the closest zero crossing in the acoustic signal. Edges between normal and reversed portions of speech were smoothed to avoid simple acoustic detection of the transition between normal and reversed speech. Reversal conditions were counterbalanced across lists and participants so that each participant saw each sentence in each of the five reversal conditions. At the end, 10 experimental conditions (5 Time Reversals × 2 Cloze Probability) were thus compared: R0low, R0high, R0.5low, R0.5high, R1low, R1high, R1.5low, R1.5high, R2low and R2high. The order of sentences in the lists was randomized and different for each participant. Fig. 1 shows the example of the sentence "*Le chanteur vend des billets pour son concert*" with the target word "*concert*" in five possible types of time reversals.

2.3. Procedure

Participants sat in an electrically and acoustically shielded chamber in front of a video monitor where they could read instructions of the experiment.

2.3.1. Linguistic oddball experiment

Participants were instructed to watch a silent movie of their own choice and to ignore the auditory stimuli (/ba/) that were presented diotically via headphones at a comfortable listening level (which was kept constant at 60 dB-SPL across subjects). The sounds were presented in a classical oddball paradigm in which a repetitive standard stimulus was replaced at a 15% probability by a deviant with a stimulus onset asynchrony (SOA) of 500 ms. The experiment was divided into two consecutive blocks of 770 stimuli each (660 standards and 110 deviants). In the first block, the intact /ba/ (forward speech) was used as the repetitive standard stimulus and the reversed /ba/ as the occasional deviant, whereas in the second block, the reversed /ba/ served as standard and the intact /ba/ as deviant. Order of blocks was counterbalanced across participants. This experiment lasted about 20 min.

2.3.2. Sentence repetition experiment

Participants were instructed to perform a sentence repetition task, alternating listening and repetition periods. A central fixation cross was presented on the screen at the beginning of each trial. Participants were instructed to attentively listen to the stimuli that were presented diotically via headphones at a comfortable listening level (60 dB-SPL for all subjects). After the end of each sentence (mean length = 2.4 s), the instruction "Repeat" was presented on the screen, prompting participants to repeat the whole sentence they just heard as accurately as possible. Participants were informed that sentences may be more or less intelligible but that they had to repeat what they heard (note that when target words contained large distortions, i.e. R1.5 and R2, most of the participants repeated the sentences with a final word that matched the preceding sentence context). The experimenter categorized the response as either correct or incorrect depending on whether the participants correctly repeated the final word of the sentence (i.e. the target word that could be time-reversed). The next trial was then presented. A training session of five sentences (not belonging to the experimental set) preceded the test phase. A break was pro-

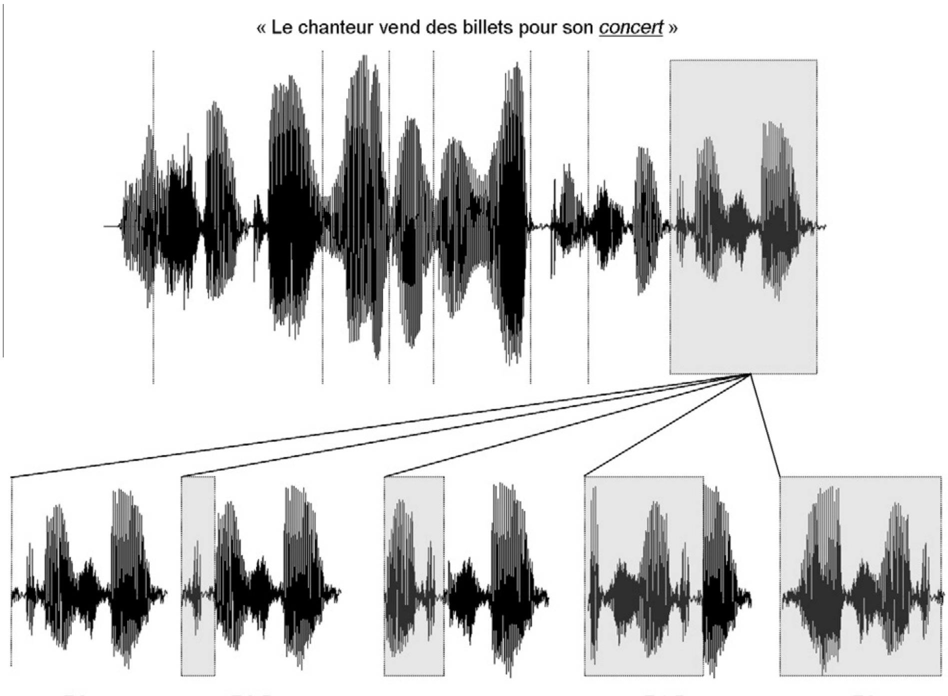

Fig. 1. Example of a stimulus used in the experiment, literally "The singer sells tickets for his concert". For this example, the five types of time reversal (R0, R0.5, R1, R1.5 and R2) were applied to the target word of the same sentence. Dotted vertical lines in the signal indicate frontiers of words within the sentence. Gray rectangles indicate the portions of the word that were time-reversed.

posed to participants halfway through the experiment. Participants were asked to stay relaxed, not move and avoid as much as possible eye movements or blinks throughout the experiment which lasted approximately 45 min.

2.4. EEG recording and pre-processing

EEG was continuously recorded from 32 scalp electrodes (Electro-Cap International, INC., according to the international 10–20 system) using the Biosemi EEG system operating at a sampling rate of 512 Hz, filtered on-line between 1 and 30 Hz and referenced to the nose. Eye movements were monitored by recording horizontal and vertical electro-oculograms (hEOG and vEOG respectively) with a bipolar montage of two electrode pairs: one pair placed above and below the right eye and the other on the temples lateral to the outer canthi. Data were analyzed with BESA software. Raw EEG recordings were first segmented in 700 ms epochs for the linguistic oddball experiment (from 100 ms prior to /ba/ onset to 600 ms after its onset) and in 1000 ms epochs for the sentence repetition experiment (from 100 ms prior to target word onset to 900 ms after its onset). Epochs in which the EEG or EOG exceeded ±150 μV were rejected from further analyses. Seventeen participants provided recordings of satisfactory quality to be included in further analyses.

2.5. ERPs analyses

2.5.1. Linguistic oddball experiment

ERPs were separately averaged for deviant and standard stimuli in each of the two blocks for each participant. Averages were baseline-corrected using the 100 ms pre-stimulus period and re-referenced to a common average reference. Deviant minus standard ERP difference waveforms (MMN) were derived from ERPs elicited by the same syllable (time-reversed /ba/) used as standard and deviant in the two different blocks for each participant (i.e. “identity MMN”; Kujala et al., 2007; see Pulvermüller and Shtyrov (2006) for similar methods). The MMN peak amplitude was quantified by first determining the MMN peak latency from the Fz difference wave as the most negative peak between 200 and 300 ms after stimulus onset. In agreement with most MMN studies (*e.g.* Hahne et al., 2002; Kujala et al., 2001; Sussman et al., 1998; Takegata, Paavilainen, Näätänen, & Winkler, 1999; Ylinen et al., 2009), MMN amplitude was then measured in a 40-ms-window centered at peak latency for each participant. One sample *t*-tests were used to determine whether MMN mean amplitude at Fz significantly differed from zero (i.e. whether a reliable MMN was elicited) and whether it showed polarity inversion at mastoids. ERPs were then re-referenced to the average of the left and right mastoids in order to estimate the full MMN amplitude. To assess the spatial distribution of the MMN, we examined whether it was maximal at frontal sites and whether it was lateralized. Three spatial domains were defined: Frontal (F3, Fz, F4), Central (C3, Cz, C4) and Parietal (P3, Pz, P4). A two-way repeated-measures analysis of variance (ANOVA) was performed with MMN mean amplitude as the dependent variable and Spatial Domain (frontal, central, parietal) and Lateralization (left, midline, right) as within-subjects factors.

2.5.2. Sentence repetition experiment

Average ERPs, aligned to a 100 ms pre-stimulus baseline and re-referenced to a common average reference, were first computed separately for each participant, condition and electrode site. Grand averages were then calculated across all participants. On the basis of our predictions and of visual inspection of the grand mean waveforms, we chose two time-windows for further analysis: an early time-window ranging from 200 to 300 ms after target word onset (i.e. time interval within which MMN typically occurs) and a late time-window ranging from 350 to 550 ms post-stimulus (i.e. time interval related to the N400).

In the early time-window, as for the linguistic oddball experiment, we measured the mean amplitude of evoked activity in each of the 10 conditions in a 40-ms-window centered at the most negative peak latency at Fz for each participant. A two-way repeated-measures ANOVA with mean amplitude of ERPs (referenced to linked mastoids) as the dependent variable and including Time Reversal (R0, R0.5, R1, R1.5, R2) and Cloze Probability (low, high) as within-subjects factors was performed. For effects having more than one degree of freedom, the Greenhouse–Geisser correction (Greenhouse & Geisser, 1959) was applied; in these cases, the reported values of degrees of freedom and *p*-values are corrected values.

The spatio-temporal characteristics of the evoked response to manipulated sentence-final words and of the MMN elicited in the oddball experiment were then compared across participants. To this aim, and given that the MMN is the difference waveform between deviants and standards, the ERP in the R0 condition (which can be seen as a “regular standard”) was subtracted from the ERPs in the four other reversal conditions (which can be seen as “deviants”) for each participant. This was done using a common average reference. The subtraction resulted in four difference waves (“R0.5 minus R0”, “R1 minus R0”, “R1.5 minus R0”, “R2 minus R0”) whose mean amplitude at Fz in a 40-ms window centered at peak latency was tested against zero with one sample *t*-tests across participants. *T*-tests also allowed assessing polarity inversion at mastoids. Comparison of these four difference waves to the linguistic oddball MNN, all re-referenced to linked mastoids (to estimate the full MMN amplitude), involved two steps. First, we directly compared peak latency and mean amplitude of the difference waves to the latency and mean amplitude of the oddball-MMN using *t*-tests. Second, the spatial distribution of the four difference waves to manipulated words was examined using a three-way repeated-measures ANOVA with ERP mean amplitude as the dependent variable. The same spatial domains as the ones defined for the oddball experiment were used: Frontal (F3, Fz, F4), Central (C3, Cz, C4) and Parietal (P3, Pz, P4). The ANOVA included Time Reversal (“R0.5 minus R0”, “R1 minus R0”, “R1.5 minus R0”, “R2 minus R0”), Spatial Domain (frontal, central, parietal) and Lateralization (left, midline, right) as within-subjects factors.

In the late time-window (350–550 ms after target word onset), mean amplitude data were analyzed using a four-way repeated-measures ANOVA with Time Reversal (R0, R0.5, R1, R1.5, R2), Cloze Probability (low, high), Spatial Domain (frontal, central, parietal) and Lateralization (left, midline, right) as within-subjects factors (the Greenhouse–Geisser correction was applied when needed). In case of significant interactions, planned comparisons (LSD test) were computed to evaluate differences between conditions.

All trials were taken into consideration in the statistical analysis regardless of the participant’s response on the repetition task. This was because some of the participants had only very few correct responses in some of the conditions (*e.g.* R1.5 and R2) and a response-contingent averaging would have decreased the signal-to-noise ratio. Note however that ERPs analysis including only correct responses gave similar patterns of results as those reported in the text.

2.6. Behavioral performance assessment

Behavioral accuracy of the 17 participants included in the ERPs analysis was assessed by counting the number of correct and incorrect repetitions of target words. Partial, approximate or semantically-related responses were considered as incorrect. Behavioral results were expressed as comprehension rates for each of the 10 conditions (R0low, R0high, R0.5low, R0.5high, R1low, R1high, R1.5low, R1.5high, R2low and R2high). A two-way repeated-measures ANOVA considering comprehension rates as the dependent variable and including Time Reversal and Cloze Probability as within-subjects factors was performed.

3. Results

3.1. Behavioral results

The two-way ANOVA first revealed a significant main effect of Time Reversal ($F(4, 64) = 301.03$, $p < .001$), conditions R1, R1.5 and R2 eliciting significantly lower comprehension rates than conditions R0 and R0.5 ($p < .01$; Table 1). The three conditions (R1, R1.5 and R2) also significantly differed from each other ($p < .001$) whereas conditions R0 and R0.5 did not. A significant main effect of Cloze Probability was further observed ($F(1, 16) = 148.76$, $p < .001$), indicating higher comprehension rates when target words were predictable from the context (79.8%, SD = 26.7) than when they were not (56.7%, SD = 40). Finally, the interaction between the two factors was significant ($F(4, 64) = 37.42$, $p < .001$): high-CP target words were better recognized and repeated than low-CP target words for time reversals equal to or longer than one syllable (Table 1 and Fig. 2). Performance did not differ between high- and low-CP target words in the conditions R0 and R0.5, suggesting that participants correctly heard and repeated the intact (non-reversed) stimuli and that the reversal of half of the first syllable did not affect word recognition and subsequent word repetition, even when the word was not predictable from the context.

Overall, these results therefore show that time reversal applied to a bi-syllabic target word presented in a sentential context begins to affect its identification and subsequent repetition at reversal sizes as large as one syllable in French, particularly when the word is not predictable from the context. Notably, even when 1 or 1.5 syllables of the bi-syllabic target words were time-reversed, participants were able to retrieve them at rates of 77% and 41% respectively, suggesting that the acoustic/phonetic distortion was somehow overcome by top-down processes ultimately allowing participants to retrieve most of the words.

3.2. ERPs results

3.2.1. Linguistic oddball experiment

Fig. 3 displays the grand-average ERPs to the standard and the deviant stimuli and the corresponding difference waveform at Fz electrode. The difference wave revealed a large negative response, identified as the MMN, peaking at 237 ms from stimulus onset, distributed over fronto-central sites and showing a polarity inversion at mastoids. One sample t-tests confirmed that MMN mean amplitude significantly differed from zero at Fz (i.e. an MMN was elicited; $-2.72\ \mu V$; $t_{16} = -2.13$, $p = .04$) and that it inverted polarity at mastoids ($1.07\ \mu V$; $t_{16} = 4.57$, $p < .001$). The two-way ANOVA (Spatial Domain × Lateralization) revealed a significant main effect of Spatial Domain ($F(2, 32) = 17.69$, $p = .001$): MMN amplitude was maximal over frontal ($-2.72\ \mu V$, SD = 2.06) and central electrodes ($-2.28\ \mu V$, SD = 1.98) compared to parietal sites ($-1.34\ \mu V$, SD = 1.59; $p = .001$). No significant effect of Lateralization was observed nor was there a significant interaction between the two factors.

Table 1
Mean percentage of correct repetition of sentence-final target words (with standard deviations, SD) for each Time Reversal condition (R0, R0.5, R1, R1.5 and R2) and for words with a high or low Cloze Probability within the sentence context.

Time reversal	Cloze probability	% Correct	SD	Mean (%)
R0	High	99.7	1.2	99.4
	Low	99.1	1.9	
R0.5	High	100	–	98.7
	Low	97.3	3.1	
R1	High	90.3	15.8	76.6
	Low	62.9	13.6	
R1.5	High	63.8	21.6	41.3
	Low	18.8	9.8	
R2	High	45.3	22.9	25.3
	Low	5.3	4.5	

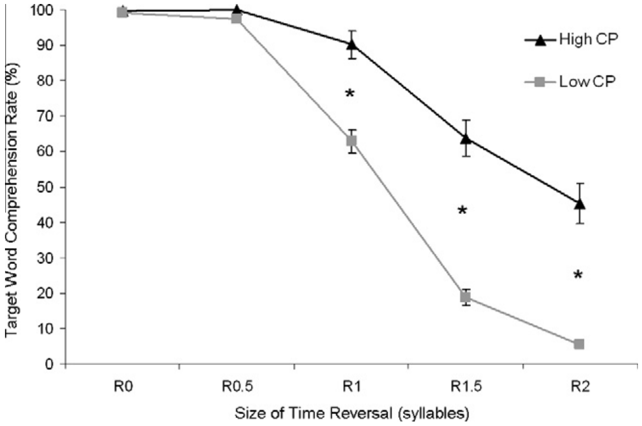

Fig. 2. Comprehension rates (%) for target words in each of the reversed-speech conditions (R0, R0.5, R1, R1.5 and R2) as a function of Cloze Probability (CP) of the word within the sentence context (low or high). (*) indicates a significant difference between conditions ($p < .001$). Error bars are reported.

3.2.2. Sentence repetition experiment

3.2.2.1. Early time-window (40-ms-window centered at peak latency). All reversal conditions (R0.5, R1, R1.5 and R2), irrespective of the cloze probability of target words in the sentences, showed a well-defined negative wave (mean amplitude at Fz = $-3.39\ \mu V$, SD = 3.91) compared to the R0 condition ($-1.54\ \mu V$, SD = 3.92; Fig. 4a). This component had an onset around 180 ms from target word onset, peaked on average at 245 ms at Fz and went back to null voltage-values around 300 ms. Spatial distribution inspection of this ERP showed a large monophasic negative wave maximal over fronto-central anterior sites, with a slight asymmetry in favor of the right hemisphere. The two-way ANOVA (Time Reversal × Cloze Probability) revealed a significant main effect of Reversal on the mean amplitude of this early evoked component ($F(4, 64) = 3.17$, $p = .019$). Planned comparisons showed that mean amplitude in the early time-window was significantly more negative in the R0.5, R1.5 and R2 conditions than in R0 ($p < .02$; Table 2). Mean amplitude in R2 was also significantly more negative than in R1 ($p = .04$); all remaining comparisons remained non-significant. The main effect of Cloze Probability was not significant nor was there a significant interaction between the two factors.

In the early time-window, the processing of time-reversed speech was thus associated with the generation of a frontal negative wave, independently of the actual size of the reversal window. All reversed conditions showed this effect which was absent in the non-reversed control condition.

We then compared the spatio-temporal characteristics of the MMN elicited during the linguistic oddball paradigm to those of the negative wave generated when target words were time-reversed. To this aim, as described in the Methods, we subtracted the evoked response in the R0 condition (“regular standard”) from the response in the four other reversal conditions (“deviants”) for

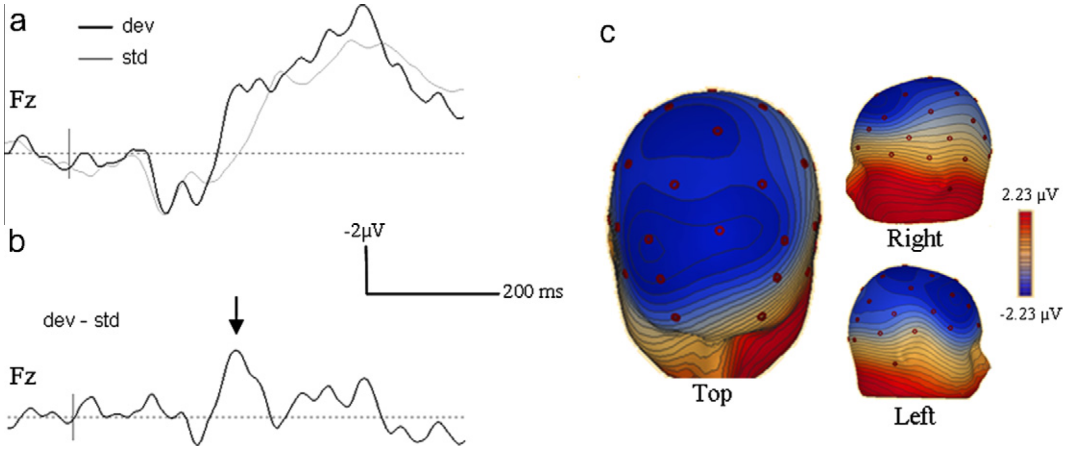

Fig. 3. (a) Grand-average ERPs to the standard (std) and the deviant stimuli (dev) in the linguistic oddball experiment at Fz electrode. (b) Difference waveform (deviant minus standard; "identity MMN") at Fz. (c) Pictures of the 3D voltage interpolation observed at 240 ms for the difference wave, showing the spatial distribution of the MMN.

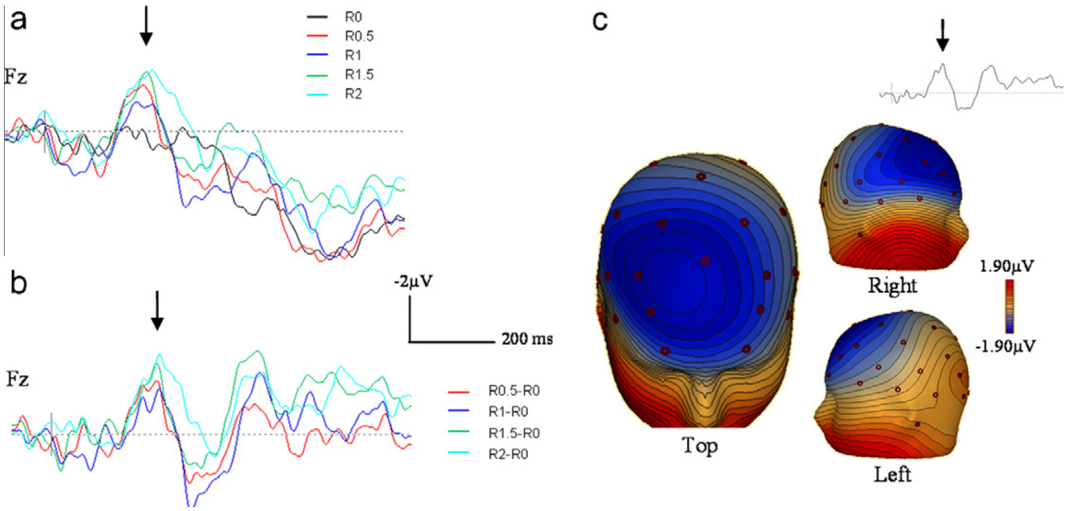

Fig. 4. (a) Grand-average ERPs to target words in the five Time Reversal conditions (R0, R0.5, R1, R1.5 and R2). The arrow indicates the early negative wave (mean latency = 248 ms) that was observed when target words were time-reversed. (b) Grand-average difference wave when activity for the R0 condition was subtracted from activity in each of the other four time-reversed conditions ("R0.5 minus R0", "R1 minus R0", "R1.5 minus R0" and "R2 minus R0"). (c) Pictures of the 3D voltage interpolation observed around 248 ms for the grand-average wave averaged across the four subtraction conditions displayed in (b). The corresponding grand-average difference wave is displayed in the upper right panel.

Table 2

Peak latency and mean amplitude (with SD) at Fz (in a 40-ms-window centered at peak latency) of the ERP to sentence-final words averaged over all participants are reported for each Time Reversal condition (R0–R2) and depending on the high- or low cloze probability of words in sentences. As a reminder, peak latency and mean amplitude of the MMN elicited in the oddball paradigm were 237 ms and −2.72 μV respectively.

Time reversal	Cloze probability	Peak latency (ms)	SD	Mean (ms)	Mean amplitude (μV)	SD	Mean (μV)
R0	High	233	29	239	−0.33	3.73	−1.54
	Low	245	41		−2.74	3.83	
R0.5	High	247	29	232	−3.57	3.31	−3.52
	Low	218	58		−3.46	4.14	
R1	High	242	44	245	−2.49	2.75	−2.44
	Low	249	29		−2.39	4.45	
R1.5	High	252	28	246	−3.42	4.54	−3.46
	Low	241	34		−3.50	4.56	
R2	High	240	28	245	−3.12	2.63	−4.16
	Low	251	33		−5.19	5.24	

each participant. Cloze probability was not taken into account in this comparison as it did not significantly affect ERP amplitude in the first analysis (see ANOVA above). As shown in Fig. 4b, the subtraction resulted in four difference waves ("R0.5 minus R0", "R1 minus R0", "R1.5 minus R0", "R2 minus R0") whose mean amplitudes in a 40-ms window centered at peak latency significantly differed from zero at Fz ($t_{16} < -3$, $p_s < .01$) and which inverted polarity at mastoids ($t_{16} > 2.66$, $p_s < .02$) as attested by one sample *t*-tests. Fig. 4c displays the grand-average wave averaged across the four subtraction conditions and across participants together with its spatial topography. The early component peaked over fronto-central anterior sites around 248 ms from target word onset, with a slight asymmetry in favor of right hemiscalp locations. The surface potential polarity inversion was situated along a circular upper line passing through upper frontal, bilateral temporal and parietal sites. *t*-tests first allowed directly comparing the temporal characteristics (latency and amplitude) of the oddball MMN and of the four difference waves to manipulated words across participants. For all four comparisons, no significant difference was observed between the latency and amplitude of the two evoked components (Table 3). Second, a three-way ANOVA (Time Reversal × Spatial Domain × Lateralization) on mean amplitude of the four difference waves revealed no significant effect of Reversal but a significant main effect of Spatial Domain ($F(2, 32) = 11.13$, $p = .0002$), indicating larger ERP amplitude over frontal (−3.19 μV, SD = 3.06) and central electrodes (−2.55 μV, SD = 2.67) than over parietal electrodes (−1.30 μV, SD = 2.53; $p < .005$). A significant main effect of Lateralization also emerged ($F(2, 32) = 5.32$, $p = .001$), showing larger amplitude of the early negativity along the midline (−2.59 μV, SD = 2.97) and in the right hemisphere (−2.53 μV, SD = 2.68) than in the left hemisphere (−1.93 μV, SD = 2.95; $p < .01$). No significant interaction between the three factors was observed.

Overall, these results therefore indicate that the early negative response elicited when portions of target words were time-reversed strongly mimics the MMN in terms of temporal dynamics and spatial distribution.

3.2.2.2. ERPs – late time-window (350–550 ms). For the ease of visualization, grand-average ERPs were inspected separately for low- and high-CP sentences. For low-CP sentences, a negative wave peaking around 420 ms after word onset and maximal over left fronto-central sites (Cz: peak = −1.86 μV; mean amplitude = −0.14 μV, SD = 3.32) was observed in the R0 condition (Fig. 5a). This ERP most likely corresponds to the N400 reflecting the difficulty of integration of the unpredictable target word within the sentence context. In the four other reversal conditions, a positive shift of the negative wave was observed, especially for the R0.5 and R1 conditions (Cz: peak = 0.44 μV and 1.67 μV respectively; mean amplitude = 1.88 μV, SD = 2.75 and 2.84 μV, SD = 2.11 respectively). For high-CP sentences, a positive wave peaking around 520 ms with a maximum amplitude over centro-parietal sites (Cz: peak = 5.22 μV; mean amplitude = 3.79 μV, SD = 3.59) was observed in R0 (Fig. 5b). This wave shifted towards less positive (more negative) values for the other reversal conditions, particularly R1.5 and R2 (Cz: peak = 2.78 μV and 3 μV respectively; mean amplitude = 0.69 μV, SD = 3.60 and 1.81 μV, SD = 2.76 respectively).

The four-way ANOVA (Time Reversal × Cloze Probability × Spatial Domain × Lateralization) revealed no significant main effects of Reversal or CP but a significant main effect of Spatial Domain ($F(2, 32) = 17.58$, $p < .001$), mean amplitudes being more negative over frontal electrodes (1.36 μV, SD = 3.49) than over central (1.81 μV, SD = 2.99; $p = .002$) and parietal sites (2.15 μV, SD = 2.83; $p = .001$). The main effect of Lateralization was also significant ($F(2, 32) = 15.15$, $p < .001$), indicating more negative amplitudes in the left hemisphere (1.40 μV, SD = 3.10) than in the right hemisphere (1.92 μV, SD = 3.07; $p = .001$) or along the midline (2.01 μV, SD = 3.25; $p = .001$). Interestingly, we found a significant Time Reversal × CP interaction ($F(4, 64) = 2.65$, $p = .041$), showing that predictability of the target words within the sentences affected cortical activity differently depending on the size of the reversal window. Planned comparisons showed that mean ERP amplitude was significantly more negative for low-CP (0.43 μV, SD = 2.73) than for high-CP target words (3.28 μV, SD = 3.66; $p = .002$) only when words were intact (R0). In the other reversal conditions, the comparison between low- and high-CP words remained non-significant. Finally, a significant Time Reversal × CP × Spatial Domain interaction emerged ($F(8, 128) = 2.83$, $p = .006$), indicating that the effect of Time Reversal as a function of CP was more pronounced over frontal than over central and parietal electrodes. For high-CP words, mean amplitudes at frontal sites decreased (i.e. became more negative) as the size of the reversal increased (*e.g.* R0 = 2.94 μV, SD = 4.15 vs. R2 = −0.40 μV, SD = 4.75). Planned comparisons revealed significant differences between all reversal conditions ($p < .05$) except between R0.5 (−0.54 μV, SD = 2.91), R1 (1.18 μV, SD = 3.36) and R1.5 (0.87 μV, SD = 3.77) which gave similar results. For low-CP words, mean amplitudes over frontal electrodes increased (i.e. became more positive) as the size of the reversal increased (*e.g.* R0 = −0.54 μV, SD = 2.91 vs. R1 = 2.99 μV, SD = 3.44). This was confirmed by planned comparisons showing significant differences between all conditions ($p < .05$) except between R1.5 (0.34 μV, SD = 3.31) and R2 (0.89 μV, SD = 4.44). The positive shift was indeed observed for time reversals as long as one syllable; amplitudes again shifted towards more negative values in R1.5 and R2 conditions. The interaction (Time Reversal × CP × Spatial Domain) is illustrated in Fig. 5c where only conditions R0, R0.5 and R1 are represented as for the two other reversal conditions (R1.5 and R2), the distortion was so disruptive that word intelligibility was too low as attested by behavioral performances.

To sum up, particularly over frontal electrodes, mean ERP amplitudes tended to shift towards more positive values when time reversal was applied to low-CP target words, whereas they tended to shift towards more negative values when the distortion affected high-CP words.

4. Discussion

The present study investigated cortical responses to processing transient changes at the acoustic/phonetic level that occurred during auditory sentence processing. We were particularly interested in examining the brain mechanisms underlying early detection of an acoustic/phonetic variation within a continuous speech stream and how these mechanisms interact with those related to contextual integration. Healthy participants were instructed to listen to and repeat sentences whose final target words could be time-reversed and either predictable or not from the context. The lengths of time reversals tested were 0.5, 1, 1.5 or 2 syllables of the bi-syllabic target words.

Behavioral results first showed that when only half of the first syllable of the target word was time-reversed (R0.5), word comprehension rates remained as high as when there was no distortion (R0; 98% vs. 99% respectively), irrespective of word cloze probability within the sentence context. Conversely, for reversals of one or more than one syllable, participants found it harder to retrieve the

Table 3
Peak latency and mean amplitude (with SD) of the MMN elicited in the linguistic oddball paradigm and of the four difference waves “R0.5 − R0”, “R1 − R0”, “R1.5 − R0” and “R2 − R0” elicited in the sentence repetition experiment.

	Peak latency (ms)	SD	Mean amplitude (μV)	SD
MMN	237	26	−2.72	2.22
R0.5 − R0	238	34	−3.30	3.5
R1 − R0	250	28	−2.71	3.05
R1.5 − R0	249	27	−3.80	2.52
R2 − R0	256	39	−4.57	3

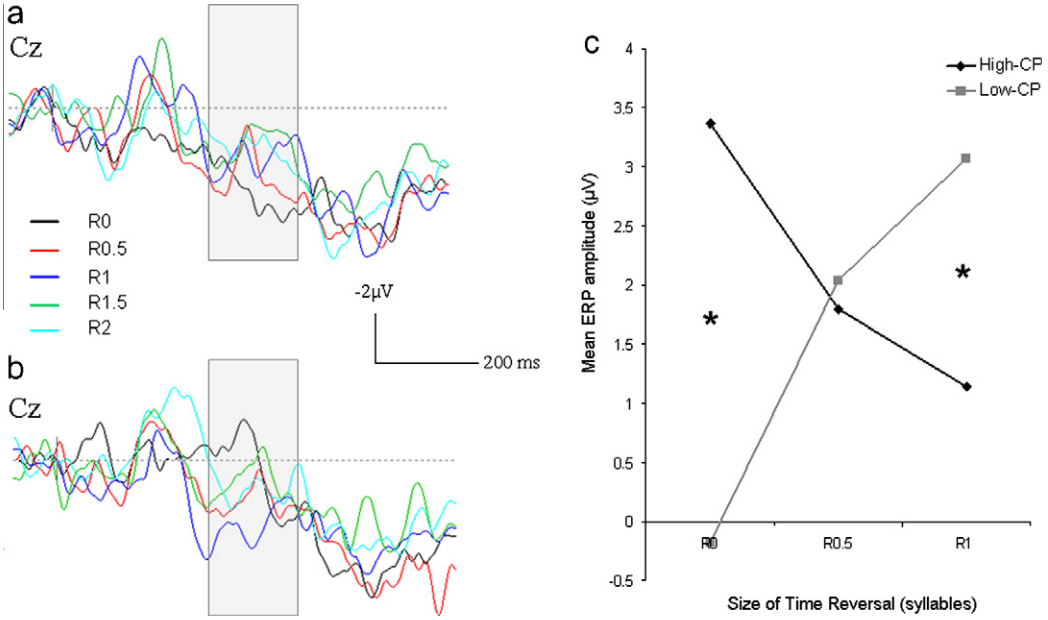

Fig. 5. Grand-average ERPs to target words in the five Time Reversal conditions (R0, R0.5, R1, R1.5 and R2) over Cz for (a) low-CP words and (b) high-CP words. The late time-window (350–550 ms) is represented by the gray rectangle. (c) Illustration of the Reversal × CP × Spatial Domain interaction ($p = .006$). Mean ERP amplitudes (μV) averaged over frontal electrodes (F3, Fz, F4) are displayed for 3 reversal conditions (R0, R0.5 and R1) in which word comprehension was still associated with high comprehension rates. (*) indicates a significant difference between conditions ($p < .001$).

words with scores falling to 77% in R1, 41% in R1.5 and 25% in R2. Interestingly however, word cloze probability strongly affected performance for such large manipulations. Scores actually remained quite high when words had a high-CP, even when three-quarters of the words were reversed (R1: 90%; R1.5: 64%; R2: 45%), whereas they were dramatically reduced when word cloze probability was low (R1: 62%; R1.5: 19%; R2: 5%). Overall, these results provide clear-cut evidence that speech comprehension does not only rely on bottom-up processes but that top-down mechanisms such as activation of lexical and semantic knowledge complement the analysis of acoustic/phonetic features of speech (Davis & Johnsrude, 2007). Such top-down processes allow to some extent maintaining speech intelligibility for efficient comprehension, even when large portions of the signal are distorted (Kiss, Cristescu, Fink, & Wittmann, 2008; Saberi & Perrott, 1999). Previous studies have indeed reported a beneficial effect of semantic context on auditory word recognition under acoustically compromised conditions, suggesting that degraded words within sentences that do not map automatically onto meaning can be reconstructed by reprocessing them in the context of semantic predictability (Obleser & Kotz, 2010; Obleser, Wise, Dresner, & Scott, 2007; Sivonen et al., 2006). The fact that repetition scores for high-CP words were lower than would be expected solely based on their cloze probability however suggests that semantic cues were not sufficient for listeners to reconstruct words but that the quality of the acoustic input plays a crucial role in lexico-semantic processes and speech comprehension.

Second, electrophysiological results revealed that detection of a sudden change in the acoustic/phonetic features of speech sounds embedded in sentences was accompanied by an early fronto-central negativity peaking around 245 ms after target word onset. This ERP was elicited for all reversal conditions, independently of the size of the reversal window and of word cloze probability in the sentences. Time reversal and word cloze probability also affected late evoked potentials recorded over fronto-central and parietal sites from 350 to 550 post-stimulus. In the next sections, we will successively describe and discuss the ERPs in these two time-windows. In a last section, we will finally propose a functional link between the early automatic acoustic/phonetic deviance detection and late semantic integration processes and discuss their neural bases based on previous findings.

4.1. Early negativity to acoustic/phonetic change within a speech stream

When time reversal was applied to the onset of sentence-final target words, an early fronto-central negativity whose amplitude was not modulated by the magnitude of the manipulation was observed. A direct comparison of this negative wave to the Mismatch Negativity (MMN) recorded in a linguistic oddball paradigm to deviant (time-reversed) syllables in a sequence of standard (non-reversed, intact) syllables in the same participants revealed similar spatio-temporal characteristics between the two markers. First, both ERPs showed polarity inversion at mastoids and a fronto-central distribution with maximal amplitude at frontal sites, which is consistent with the scalp topography of the MMN (Alho, Paavilainen, Reinikainen, Sams, & Näätänen, 1986; Giard et al., 1995). Second, latency and amplitude of the two components were very similar, as both peaked around 240 ms after onset of the deviants and had mean amplitude around −3 μV. Analysis of the spatial distribution of the two negativities however revealed that although the oddball MMN was not lateralized, the evoked response to time-reversed words was maximal along the midline and in the right hemisphere. This slight hemiscalp asymmetry favoring right-frontal sites is nevertheless consistent with an MMN interpretation as previous studies have shown that MMN can predominate in one of the two hemispheres depending on stimuli and context (Kujala et al., 2002; Muller-Gass et al., 2001; Shtyrov, Kujala, Ilmoniemi, & Näätänen, 1999; Shtyrov et al., 1998). Using MagnetoEncephaloGraphy (MEG), Kujala et al. (2002) demonstrated that the magnetic counterpart of the MMN (MMNm) was enhanced in the right hemisphere to syllables presented in a word context compared to syllables presented alone. The authors interpreted their finding as reflecting right-hemisphere specialization for the analysis of contextual acoustic information that could be related to right-hemispheric dominance for processing speech prosody. Shtyrov et al. (1998, 1999) also reported that although the left hemisphere is dominant during speech perception, the addition of masking noise causes a shift in the magnetic evoked field from the left to the right hemisphere. The authors suggested that sensory speech perception may be redistributed between the two hemispheres in ecological listening situations involving background noise, with a reinforced contribution of the right hemisphere. In agreement with this, and although analysis of source localizations would need to be carried out using a larger number of electrodes, our results seem to suggest that processing distorted words within a continuous speech stream elicits a slightly right-lateralized fronto-central negativity shortly after the onset of the acoustic/phonetic change. Overall, given the spatio-temporal characteristics of this evoked response, we suggest that it can be labeled an MMN.

In the present study, the early negativity was elicited whenever an acoustic/phonetic change was encountered irrespective of its size. The lack of amplitude modulation as a function of the magnitude of the distortion may seem at odds with previous studies showing that MMN amplitude increases with increasing acoustic difference between the deviant and the standard (Kujala et al., 2001; but see Horvath et al., 2008) and that it is sensitive to the duration of the deviant stimulus (Amenedo & Escera, 2000). However, these studies mostly used non-linguistic short stimuli (*e.g.* tones) or speech segments (*e.g.* syllables), making the comparison with our work rather difficult. Our results at least suggest that this EEG marker may show some degree of speech specificity. It is nevertheless also possible that in our study, MMN amplitude increased slightly with the size of the reversal window but the discriminative power of the current method was insufficient for this effect to come out.

Remarkably, the early negativity was observed even for manipulations as subtle as one half of the first syllable of the word (R0.5) though this did not affect word intelligibility at all. Amplitude and latency of the early ERP in this R0.5 condition did not significantly differ from those observed for larger violations that conversely had a strong behavioral impact. Such a finding suggests that the early negativity we observed in response to manipulated words may reflect fine-discrimination capabilities, regularity and automaticity in the response mechanism that are highly consistent with an MMN interpretation (Näätänen, 2001). It also suggests that this automatic response occurring at a somewhat low-level may not predict higher-level processes and thus intelligibility performance.

The temporal dynamics and scalp distribution of our recorded negativity could also be consistent with an N1. This interpretation is nevertheless unlikely as the observed component peaked later than would have been expected for an auditory word-onset N1 (Rugg & Coles, 1995) and was maximal at Fz whereas the N1 is usually maximal at Cz. In addition, in our experiment, no condition contained physical gap or a clear physical change indicating the onset of the reversal. Instead, participants had to detect a phonotactic violation or a sudden disruption along the temporal axis of the input signal, incompatible with the regularities of natural speech, which may have elicited our negativity. For this reason, and as already observed in other studies using continuous speech without clear boundaries between words, we would not have expected a clear N1 response to emerge at word onsets as these were not physically marked.

Overall, our results therefore suggest that when listening to natural speech, the brain rapidly extracts "abstract" regularities from the continuous signal about speaker's identity (*e.g.* fundamental frequency) as well as about other acoustic/phonetic information, and forms memory traces in the auditory cortex so that a sudden change within the speech stream elicits an MMN. These findings complement previous works by revealing the existence of brain mechanisms involved in the detection of regular patterns or rules among longer units than speech fragments (*e.g.* phonemes and syllables) that further interact with later processes underlying semantic integration. Our results are also corroborated by the study by Agus et al. (2010) who showed that "repeated exposure to a random waveform, up to 2 s long, results in the learning of acoustic details of the waveform". Hence, memory traces for complex arbitrary (periodic) sounds can be formed extremely rapidly even when learning is unsupervised, that is, when participants do not know which ongoing sounds they have to memorize. These traces are long-lasting, as participants retained memories for various noises after a few weeks, and robust to interference from other task-relevant sounds (Agus et al., 2010). Here we show that memory traces also develop for aperiodic long sounds such as sentences and that these traces include large-scale details about acoustic as well as phonetic features of the speech signal. Such an ability to extract abstract patterns seems crucial for speech processing as under ecological conditions, we have to categorize and understand speech sounds that can vary considerably, for instance when they are uttered by different speakers or when they are perceived in noise.

4.2. Late ERPs reflecting semantic integration

In a window ranging from 350 to 550 ms after target word onset, an interaction between time reversal and cloze probability emerged. For low-CP sentences, a fronto-central negativity was observed to intact words (R0low) around 420 ms post-stimulus. This most likely corresponds to the N400 reflecting the difficulty of integration of the unpredictable word into its context. Interestingly, when time reversal was applied to words, this negative wave shifted towards less negative amplitude values, particularly over frontal sites. This was mainly observed in conditions where the manipulation was shorter in duration or equal to the first syllable of the words, whereas for larger reversals which severely reduced comprehension rates, amplitudes tended to return to more negative values. Such a result suggests that although low-CP words were difficult to integrate within the sentences, the acoustic/phonetic change caused them to be less contextually incongruent. In other words, the violation of context-driven expectancies for these words appeared less salient due to the distortion. This seems in agreement with a recent functional Magnetic Resonance Imaging (fMRI) study in which the left inferior frontal gyrus specifically responded to low-predictable sentence-final words, indicating an extra processing effort (corresponding to the N400 as measured by ERP recordings), but only when sentences were intelligible (Obleser & Kotz, 2010). When intelligibility was reduced by spectrally degrading speech, activity in this frontal region decreased, suggesting that sentential integration was compromised.

Conversely, for high-CP sentences, a positive wave peaking around 520 ms after word onset over fronto-central and parietal sites was observed when words were not manipulated. Time reversal then caused a shift of this response towards more negative values (i.e. approaching an N400), amplitudes being the most negative when the size of the deviation was maximal (R2). Hence, although high-CP words were semantically congruent with the context and led to good comprehension rates, acoustic/phonetic change created an uncertainty about these words so that they tended to be processed as low-CP words. This indicates that comprehending distorted speech, even when it matches semantic expectations built up from context, is more demanding and recruits more neuronal resources – as evidenced by the shift towards negative amplitudes, particularly over frontal regions – than comprehending normal predictable speech which is effortless. This again agrees with the study by Obleser and Kotz (2010) who found no specific inferior frontal activation during processing of high-predictable sentence-final words. Altogether, these observations stress the involvement of fronto-parietal neural systems in the comprehension of speech under adverse conditions. Fronto-parietal networks are known to be involved in reorienting mechanisms, including anticipatory procedures used to direct attention based on goals and expectations as well as detection procedures allowing reorientation of attention towards behaviorally relevant stimuli (see Corbetta, Patel, & Shulman, 2008 for a review). Functional connectivity in fronto-parietal circuits has been shown to increase as a function of predictability of words in sentential contexts when words were only moderately intelligible (Obleser et al., 2007; see also Sharp, Turkheimer, Bose, Scott, & Wise, 2010). In a study of the auditory continuity illusion effect, Shahin, Bishop, and Miller (2009) further demonstrated that frontal regions were activated by missing speech information. This suggests that frontal regions contain high-level representations of expected information that drive top-down modulations of sensory processing via fronto-parietal networks (Desimone & Duncan, 1995) and eventually replace information when it is missing. Parietal regions on the other hand would be more generally involved in the reallocation of attentional resources, either under the pressure of top-down (expectancy-based) controls originating from prefrontal regions or under the influence of relevant but non-expected sensory inputs that automatically capture attention.

Overall, our findings corroborate the study by Aydelott et al. (2006) who found reduced N400 to incongruent words in degraded (filtered) contexts. As mentioned in the introduction, the authors proposed that the acoustic degradation reduced availability of semantic information present in the context such that semantic integration of incongruent words was less demanding. Accordingly, in the present study, manipulation of the acoustic/phonetic features of sentence-final words produced an ambiguity about these words, low-CP words being less semantically incongruent and high-CP words becoming somewhat incongruent with the preceding context. Our results also seem to corroborate both integrative and lexical accounts of the N400 (Kutas & Federmeier, 2000; Lau et al., 2008). The language system could have used context to activate relevant information for expected words. When these words were actually encountered but they were strongly distorted, the brain was unable to match information from activated lexical candidates with actual input, therefore eliciting an N400 reflecting effortful word integration into context. By contrast, when manipulated incongruent words completed the sentences, incompatibility with information from activated expected candidates may have been smaller than when incongruent words were intact, thus reducing the N400 amplitude.

4.3. Functional link between early acoustic/phonetic and late semantic processes

One intriguing observation when looking for the neural correlates of most of the deviance-detection associated components identified in ERP experiments is that they seem to engage specialized, differently localized systems that however share a common functional architecture: loops engaging the frontal and temporal cortices as well as basal ganglia nuclei. Neural generators of the MMN have been identified in the auditory cortices, but also seem to engage a larger frontal-basal comparator network including (pre)frontal cortices as well as the thalamus and hippocampus (Alho, 1995; Giard, Perrin, Pernier, & Bouchet, 1990; Rinne, Alho, Ilmoniemi, Virtanen, & Näätänen, 2000). Generation of the N400, evoked to the detection of mismatching semantic information, has also been assumed to involve a fronto-temporal network mainly engaging the left middle temporal and inferior frontal gyri (Lau et al., 2008; Van Petten & Luka, 2006) or medial temporal structures close to the hippocampus (McCarthy, Nobre, Bentin, & Spencer, 1995; Nobre & McCarthy, 1995). Interestingly the involvement of such frontal-temporal-basal loops has also been evidenced for the extraction of regularities in the rhythmic and syntactic domains (Friederici, Rüschemeyer, Hahne, & Fiebach, 2003; Opitz & Friederici, 2003; see Kotz, Schwartze, and Schmidt-Kassow (2009) for a review). One hypothesis is that reverberation of information in fronto-temporo-basal loops is associated with the processing of regularities and generation of expectancies that can occur at the different levels of speech information processing, namely from the acoustic/phonetic up to higher levels such as semantic or pragmatic contextual integration. A growing body of research indeed suggests that the brain can exploit various constraining information (*e.g.* morpho-syntactic, lexico-semantic) during sentence and discourse comprehension to make predictions about upcoming events (Federmeier, 2007; Kotz et al., 2009; Lau, Stroud, Plesch, & Phillips, 2006; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). As to the MMN, it has been proposed that this early automatic response results from a comparison between the auditory input encoded in the auditory cortex with a memory trace embodied in top-down predictions generated in prefrontal regions (Garrido, Kikner, Stephan, & Friston, 2009; Winkler, 2007). When predictions are not met, MMN response is observed that would reflect a process updating predictive models. Similar mechanisms have been assumed to account for the N400: during speech comprehension, lexico-semantic representations of words are activated in the middle temporal cortex. Such activation is facilitated by the predictive context (N400 effect), a top-down process mediated by the inferior frontal cortex (DeLong, Urbach, & Kutas, 2005; Federmeier et al., 2007; Lau et al., 2008). The fact that the very same general neuronal architecture involving fronto-temporal loops underlies encoding of different domain-specific types of information (*e.g.* acoustic, phonetic, rhythmic, semantic) would explain why 'deviance waves' (*e.g.* ERP markers located over fronto-central regions) are observed so frequently in domains as various as speech comprehension, music processing or face recognition. The idea that very general, basic information processing mechanisms could serve as a basis for apparently more complex cognitive mechanisms will certainly deserve extended research efforts in the future (Näätänen, Astikainen, Ruusuvirta, & Huotilainen, 2010).

5. Conclusions

In the present study we investigated the electrophysiological correlates of understanding reversed speech. Early detection of a time reversal applied to words embedded in sentences elicited a fronto-central negativity that spatio-temporally matched the well-known MMN. Acoustic/phonetic change then affected semantic integration of words into their context differently when these words were predictable or not from the context. We suggest that in ecological listening conditions, the MMN response may be involved in detecting transient acoustic/phonetic perturbations of the signal that violate the regularities of speech and cause it to be full of irrelevant noise. This would enhance the use of top-down contextual information that can correct for these noisy or missing bits of information and help with the final comprehension of an acoustically imperfect message. Our study therefore provides important findings regarding natural speech comprehension as it demonstrates that acoustic/phonetic information and semantic knowledge strongly interact when processing speech-in-noise. Future work will be dedicated to the better understanding of the dynamics and functional links between early and late ERP components during degraded speech comprehension.

Acknowledgments

We would like to thank two anonymous Reviewers for their efforts and help in improving this manuscript. We would also like to thank Emmanuel Ferragne for his help in the construction of the stimuli. This research was supported by a European Research Council grant to the SpiN Project (No. 209234) and by a grant from the French National Research Agency (ANR).

References

Aaltonen, O., Niemi, P., Nyrke, T., & Tuhkanen, J. M. (1987). Event related brain potentials and the perception of a phonetic continuum. *Biological Psychology, 24*, 197–207.

Agus, T. R., Thorpe, S. J., & Pressnitzer, D. (2010). Rapid formation of robust auditory memories: Insights from noise. *Neuron, 66*, 610–618.

Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear and Hearing, 16*(1), 38–51.

Alho, K., Paavilainen, P., Reinikainen, K., Sams, M., & Näätänen, R. (1986). Separability of different negative components of the event-related potential associated with auditory stimulus processing. *Psychophysiology, 23*, 613–623.

Allen, M., Badecker, W., & Osterhout, L. (2003). Morphological analysis in sentence processing: An ERP study. *Language and Cognitive Processes, 18*, 405–430.

Amenedo, E., & Escera, C. (2000). The accuracy of sound duration representation in the human brain determines the accuracy of behavioural perception. *European Journal of Neuroscience, 12*, 2570–2574.

Aydelott, J., Dick, F., & Mills, D. L. (2006). Effects of acoustic distortion and semantic context on event-related potentials to spoken words. *Psychophysiology, 43*(5), 454–464.

Besson, M., Faita, F., Czternasty, C., & Kutas, M. (1997). What's in a pause: Event-related potential analysis of temporal disruptions in written and spoken sentences. *Biological Psychology, 46*, 3–23.

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience, 5*, 34–44.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron, 58*, 306–324.

Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research, 229*(1–2), 132–147.

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport, 8*, 919–924.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117–1121.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.

DiGiovanni, J. J., & Schlauch, R. S. (2007). Mechanisms responsible for differences in perceived duration for rising-intensity and falling-intensity sounds. *Ecological Psychology, 19*(3), 239–264.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*, 491–505.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research, 1146*, 75–84.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*(2), 78–84.

Friederici, A. D., Rüschemeyer, S. A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. *Cerebral Cortex, 13*, 170–177.

Garrido, M. I., Kikner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology, 120*, 453–463.

Giard, M., Lavikainen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., et al. (1995). Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: An event-related potential and dipole-model analysis. *Journal of Cognitive Neuroscience, 7*, 133–143.

Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance: A topographic event-related potential study. *Psychophysiology, 27*, 627–640.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95–111.

Hahne, A., Schroger, E., & Friederici, A. D. (2002). Segregating early physical and syntactic processes in auditory sentence comprehension. *Neuroreport, 13*, 305–309.

Horvath, J., Czigler, I., Jacobsen, T., Maess, B., Schröger, E., & Winkler, I. (2008). MMN or no MMN: No magnitude of deviance effect on the MMN amplitude. *Psychophysiology, 45*(1), 60–69.

Kiss, M., Cristescu, T., Fink, M., & Wittmann, M. (2008). Auditory language comprehension of temporally reversed speech signals in native and non-native speakers. *Acta Neurobiologicae Experimentalis, 68*, 204–213.

Korpilahti, P., Krause, C. M., Holopainen, I., & Lang, A. H. (2001). Early and late mismatch negativity elicited by words and speech-like stimuli in children. *Brain and Language, 76*, 332–339.

Kotz, S. A., Schwartze, M., & Schmidt-Kassow, M. (2009). Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. *Cortex*.

Kozou, H., Kujala, T., Shtyrov, Y., Toppila, E., Starck, J., Alku, P., et al. (2005). The effect of different noise types on the speech and non-speech elicited mismatch negativity. *Hearing Research, 199*, 31–39.

Kraus, N., McGee, T., Sharma, A., Carrell, T., & Nicol, T. (1992). Mismatch negativity event-related potential elicited by speech stimuli. *Ear and Hearing, 13*, 158–164.

Kujala, A., Alho, K., Valle, S., Sivonen, P., Ilmoniemi, R. J., Alku, P., et al. (2002). Context modulates processing of speech sounds in the right auditory cortex of human subjects. *Neuroscience Letters, 331*, 91–94.

Kujala, T., Kallio, J., Tervaniemi, M., & Näätänen, R. (2001). The mismatch negativity as an index of temporal processing in audition. *Clinical Neurophysiology, 112*(9), 1712–1719.

Kujala, T., Tervaniemi, M., & Schroger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological cnosiderations. *Biological Psychology, 74*, 1–19.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463–470.

Kutas, M., & Federmeier, K. D. (2007). Event-related brain potential (ERP) studies of sentence processing. In G. Gaskell (Ed.), *Oxford handbook of psycholinguistics* (pp. 385–406). Oxford: Oxford University Press.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*, 161–163.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantic: (De)constructing the N400. *Nature Reviews Neuroscience, 9*(12), 920–933.

Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language, 98*, 74–88.

McCarthy, G., Nobre, A. C., Bentin, S., & Spencer, D. D. (1995). Language-related field potentials in the anterior medial temporal lobe: I. Intracranial distribution and neural generators. *Journal of Neuroscience, 15*(2), 1080–1089.

Menning, H., Zwitserlood, P., Schoning, S., Hihn, H., Bolte, J., Dobel, C., et al. (2005). Pre-attentive detection of syntactic and semantic errors. *NeuroReport, 16*, 77–80.

Muller-Gass, A., Marcoux, A., Logan, J., & Campbell, K. (2001). The intensity of masking noise affects the mismatch negativity to speech sounds in human subjects. *Neuroscience Letters, 299*, 197–200.

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology, 38*, 1–21.

Näätänen, R., & Alho, K. (1995). Mismatch negativity – A unique measure of sensory processing in audition. *International Journal of Neuroscience, 80*, 317–337.

Näätänen, R., Astikainen, P., Ruusuvirta, T., & Huotilainen, M. (2010). Automatic auditory intelligence: An expression of the sensory-cognitive core of cognitive processes. *Brain Research Review, 64*(1), 123–136.

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective attention effect on evoked potential reinterpreted. *Acta Psychologica, 42*, 313–329.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments and Computers, 36*(3), 516–524.

Nobre, A. C., & McCarthy, G. (1995). Language-related field potentials in the anterior-medial temporal lobe. 2. Effects of word type and semantic priming. *Journal of Neuroscience, 15*(2), 1090–1098.

Obleser, J., & Kotz, S. A. (2010). Expectancy constraints in degraded speech modulate the language comprehension network. *Cerebral Cortex, 20*(3), 633–640.

Obleser, J., Wise, R. J., Dresner, M. A., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *Journal of Neuroscience, 27*, 2283–2289.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97–113.

Opitz, B., & Friederici, A. D. (2003). Interactions of the hippocamapl system and the prefrontal cortex in learning language-like rules. *Neuroimage, 19*, 1730–1737.

Paavilainen, P., Jaramillo, M., Näätänen, R., & Winkler, I. (1999). Neuronal populations in the human brain extracting invariant relationships from acoustic variance. *Neuroscience Letters, 265*(3), 179–182.

Paavilainen, P., Simola, J., Jaramillo, M., Näätänen, R., & Winkler, I. (2001). Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN). *Psychophysiology, 38*, 359–365.

Pakarinen, S., Huotilainen, M., & Näätänen, R. (2010). The mismatch negativity (MMN) with no standard stimulus. *Clinical Neurophysiology, 121*, 1043–1050.

Pakarinen, S., Takegata, R., Rinne, T., Huotilainen, M., & Näätänen, R. (2007). Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) of the auditory event-related potential (ERP). *Clinical Neurophysiology, 118*(1), 177–185.

Pellegrino, F., Ferragne, E., & Meunier, F. (2010). A speech oddity: Phonetic transcription of reversed speech. In *Proceedings of Interspeech*.

Pulvermüller, F., & Shtyrov, Y. (2003). Automatic processing of grammar in the human brain as revealed by the mismatch negativity. *Neuroimage, 20*, 159–172.

Pulvermüller, F., & Shtyrov, Y. (2006). Language outside the focus of attention: The mismatch negativity as a tool for studying higher cognitive processes. *Progress in Neurobiology, 79*(1), 49–71.

Rinne, T., Alho, K., Ilmoniemi, R. J., Virtanen, J., & Näätänen, R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity sources. *Neuroimage, 12*, 14–19.

Ritter, W., Gomes, H., Cowan, N., Sussman, E., & Vaughan, H. G. Jr., (1998). Reactivation of a dormant representation of an auditory stimulus feature. *Journal of Cognitive Neuroscience, 10*, 605–614.

Rothermich, K., Schmidt-Kassow, M., Schwartze, M., & Kotz, S. A. (2010). Event-related potential responses to metric violations: Rules versus meaning. *NeuroReport, 21*(8), 580–584.

Rugg, M. D., & Coles, M. G. H. (1995). *Electrophysiology of mind–event-related brain potentials and cognition*. Oxford: Oxford University Press.

Saarinen, J., Paavilainen, P., Schröger, E., Tervaniemi, M., & Näätänen, R. (1992). Representation of abstract attributes of auditory stimuli in the human brain. *NeuroReport, 3*, 1149–1151.

Saberi, K., & Perrott, D. R. (1999). Cognitive restoration of reversed speech. *Nature, 398*, 760.

Scherg, M., Vajsar, J., & Picton, T. (1989). A source analysis of the late human auditory evoked potentials. *Journal of Cognitive Neuroscience, 1*, 336–355.

Schmidt-Kassow, M., & Kotz, S. A. (2009). Attention and perceptual regularity in speech. *NeuroReport, 20*, 1643–1647.

Shahin, A. J., Bishop, C. W., & Miller, L. M. (2009). Neural mechanisms for illusory filling-in of degraded speech. *Neuroimage, 44*, 1133–1143.

Sharp, D. J., Turkheimer, F. E., Bose, S. K., Scott, S. K., & Wise, R. J. (2010). Increased frontoparietal integration after stroke and cognitive recovery. *Annals of Neurology*, *PMID*: 20687116.

Shestakova, A., Brattico, E., Huotilainen, M., Galunov, V., Soloviev, A., Sams, M., et al. (2002). Abstract phoneme representations in the left temporal cortex: Magnetic mismatch negativity study. *NeuroReport, 13*(14), 1813–1816.

Shtyrov, Y., Hauk, O., & Pulvermüller, F. (2004). Distributed neuronal networks for encoding category-specific semantic information: The mismatch negativity to action words. *European Journal of Neuroscience, 19*, 1083–1092.

Shtyrov, Y., Kujala, T., Ahveninen, J., Tervaniemi, M., Alku, P., Ilmoniemi, R. J., et al. (1998). Background acoustic noise and the hemispheric lateralization of speech processing in the human brain: Magnetic mismatch negativity study. *Neuroscience Letters, 251*(2), 141–144.

Shtyrov, Y., Kujala, T., Ilmoniemi, R. J., & Näätänen, R. (1999). Noise affects speech-signal processing differently in the cerebral hemispheres. *NeuroReport, 10*(10), 2189–2192.

Shtyrov, Y., & Pulvermuller, F. (2002). Neurophysiological evidence of memory traces for words in the human brain. *NeuroReport, 13*, 521–525.

Shtyrov, Y., Pulvermüller, F., Näätänen, R., & Ilmoniemi, R. J. (2003). Grammar processing outside the focus of attention: An MEG study. *Journal of Cognitive Neuroscience, 15*, 1195–1206.

Sivonen, P., Maess, B., Lattner, S., & Friederici, A. D. (2006). Phonemic restoration in a sentence context: Evidence from early and late ERP effects. *Brain Research, 1121*, 177–189.

Stecker, G. C., & Hafter, E. R. (2000). An effect of temporal asymmetry on loudness. *Journal of the Acoustical Society of America, 107*(6), 3358–3368.

Sussman, E., Ritter, W., & Vaughan, H. G. Jr., (1998). Predictability of stimulus deviance and the mismatch negativity. *NeuroReport, 9*, 4167–4170.

Takegata, R., Paavilainen, P., Näätänen, R., & Winkler, I. (1999). Independent processing of changes in auditory single features and feature conjunctions in humans as indexed by the mismatch negativity. *Neuroscience Letter, 26*, 109–112.

Taylor, W. L. (1953). "Cloze" procedure: A new tool for measuring readability. *Journalism Quarterly, 30*, 415–433.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31*(3), 443–467.

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25*(2), 394–417.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition, 18*, 380–393.

Van Petten, C., & Luka, B. J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language, 97*(3), 279–293.

Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology, 21*(3–4), 147–163.

Winkler, I., Cowan, N., Csépe, V., Czigler, I., & Näätänen, R. (1996). Interactions between transient and long-term auditory memory as reflected by the mismatch negativity. *Journal of Cognitive Neuroscience, 8*, 403–415.

Ylinen, S., Uther, M., Latvala, A., Vepsalainen, S., Iverson, P., Akahane-Yamada, R., & Naatanen, R. (2009). Training the brain to weight speech cues differently: a study of finnish second-language users of english. *Journal of Cognitive Neuroscience, 22*(6), 1319–1332.