

Similarities in the neural signature for the processing of behaviorally categorized and uncategorized speech sounds

Carine Signoret,^{1,2} Etienne Gaudrain^{3,4} and Fabien Perrin¹

¹Auditory Cognition and Psychoacoustics Team, Lyon Neuroscience Research Center, CNRS UMR5292, INSERM U1028, Université de Lyon, Lyon, France

²Linnaeus Centre HEAD, The Swedish Institute for Disability Research, Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

³Centre for the Neural Basis of Hearing, Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK

⁴Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK

Keywords: beta activity, EEG, ERP, ERSP, lexical processing

Abstract

Recent human behavioral studies have shown semantic and/or lexical processing for stimuli presented below the auditory perception threshold. Here, we investigated electroencephalographic responses to words, pseudo-words and complex sounds, in conditions where phonological and lexical categorizations were behaviorally successful (categorized stimuli) or unsuccessful (uncategorized stimuli). Data showed a greater decrease in low-beta power at left-hemisphere temporal electrodes for categorized non-lexical sounds (complex sounds and pseudo-words) than for categorized lexical sounds (words), consistent with the signature of a failure in lexical access. Similar differences between lexical and non-lexical sounds were observed for uncategorized stimuli, although these stimuli did not yield evoked potentials or theta activity. The results of the present study suggest that behaviorally uncategorized stimuli were processed at the lexical level, and provide evidence of the neural bases of the results observed in previous behavioral studies investigating auditory perception in the absence of stimulus awareness.

Introduction

The extent to which linguistic processes could occur for stimuli presented below the perception threshold has been a topic of considerable debate. In the visual modality, linguistic processing of stimuli presented below categorization threshold could occur at the phonological and semantic levels (for reviews, see Kouider *et al.*, 2007; Van den Bussche *et al.*, 2009). In the auditory modality, recent studies show evidence of lexical processing for stimuli presented below categorization (Kouider & Dupoux, 2005), and below detection (Signoret *et al.*, 2011) thresholds. Semantic processing of auditory stimuli has also been observed below the auditory categorization threshold (Daltrozzo *et al.*, 2011).

The neural correlates of linguistic processing of stimuli presented below the categorization threshold have rarely been investigated in the auditory modality. In an functional magnetic resonance imaging study by Kouider *et al.* (2010), participants performed a lexical decision task on target items preceded by masked primes that could not be categorized. The result showed that word repetition between the prime and the target reduced the activity in the insula and in the left superior temporal gyrus, suggesting that this area is involved in early processing stages of lexical access. However, this study did not address whether lexical processing above and below the categorization threshold shared similar mechanisms.

The aim of the present study was to investigate the electroencephalographic (EEG) correlates of speech sound processing, above and below the categorization threshold, by testing whether the processing of stimuli presented in these two conditions shared common neural signatures. Silences, words, pseudo-words and complex sounds were used. Words, pseudo-words and complex sounds were 'energetically matched', that is they had the same long-term spectra and the same envelope fluctuations (see Signoret et al., 2011 for details). These auditory stimuli were presented at two intensity levels of presentation: one at which participants could perform a categorization task at both the lexical (differentiating between word and pseudo-word) and the phonological (differentiating between pseudoword and complex sound) levels; and the other at which they were unsuccessful in the categorization task at the phonological and lexical levels, similarly to Daltrozzo et al. (2011). EEG methodology was used, and particular attention was dedicated to the spectral response for lexical access described by Pulvermüller et al. (1996). In Pulvermüller et al.'s (1996) study, responses recorded in the left

Correspondence: Dr C. Signoret, ²Department of Behavioral Sciences and Learning, as above. E-mail: carine.signoret@liu.se

Received 20 July 2012, revised 14 November 2012, accepted 16 November 2012

hemisphere revealed a depression of spectral power in the 14–33-Hz frequency range after pseudo-words but not after words, confirming results observed in the visual modality by Lutzenberger *et al.* (1994). In studies comparing words and pseudo-words (Eulitz *et al.*, 2000; Krause *et al.*, 2006; Shahin *et al.*, 2009), this decrease in power was interpreted as a breakdown of cognitive processing induced by a stimulus for which no lexical representation is available, thus reflecting a failure in lexical access. We used this marker to investigate the neural differences between the processing of stimuli differing in their phonological or lexical contents, or both. We hypothesized that if lexical processing for stimuli presented above and below the categorization threshold shared common neural signature, then a similar decrease would be observed after pseudo-words and complex sounds, but not after words above as well as below the categorization threshold.

Materials and methods

Participants

Fourteen participants (mean age 23.43 ± 3.93 years, median value of 22.31 years, min = 20 years, max = 33 years, five women), right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971), were included in the study. All participants were native speakers of French, did not report any hearing problems or any history of neurological disease, and presented low auditory detection thresholds (American National Standard Institute, 2004). All participants provided written informed consent, and the study was conducted in accordance with the principles laid down in the Declaration of Helsinki and approved by the local Ethics Committee (CPP Sud Est II, n°2007-010-2).

Materials

Three types of stimuli were used: words; pseudo-words; and complex sounds. Three-hundred and sixty words were selected from a French database (Lexique 2; New et al., 2004, 2007). All words were common, singular and monosyllabic nouns, and had a frequency of occurrence higher than 1 per million. Three-hundred and sixty pseudowords were created differing only from the words by their absence of lexical content but matched on all other properties. Like the words, the pseudo-words were monosyllabic, phonotactically legal and pronounceable, but meaningless¹ combinations of sounds in French. Overall, the pseudo-words contained all the phonemes of the words, as done in previous studies investigating speech perception (Daltrozzo et al., 2011; Signoret et al., 2011). The average number of phonemes of the pseudo-words was carefully matched to that of the words (2.3 for the words against 2.4 for the pseudo-words; $t_{718} = -1.52$, P > 0.125). The words, pseudo-words and complex sounds had an average duration of 517.78 ms (\pm 101.71 ms), 548.13 ms $(\pm$ 86.06 ms) and 532.02 ms $(\pm$ 91.52 ms), respectively. Despite this slight difference in duration between words and pseudo-words, the phoneme rate was not different for the two types of stimuli: the average phoneme duration was 237.1 ms for the words and 240.2 ms for the pseudo-words ($t_{718} = 0.60$, P > 0.33). Phoneme co-occurrence frequencies were different on average (the pseudo-words having a mean lower frequency), but identical in 89% of the instances. The words (e.g. 'sage', 'miel') and pseudo-words (e.g./mi₃/,/lEl/) were uttered by the same female speaker and recorded separately (32 bits, 44,100 Hz). Three-hundred and sixty non-phonological complex sounds were created from the words and pseudo-words using the algorithm *Fonds sonores* (Perrin & Grimault, 2005; Hoen *et al.*, 2007). This method is similar, at least in its principles, to other methods used in neuroimaging studies (e.g. Davis & Johnsrude, 2003; Giraud *et al.*, 2004): the phase spectrums of words and pseudo-words were randomized, but their average spectral content and temporal envelope (up to 60 Hz) were preserved. All stimuli were equalized to the same dB-A-weighting (dB-A) level. The A-weighting was chosen because it roughly mimics the type of filtering caused by the ear canal and the middle ear, and provides a reasonable approximation of the sound level reaching the inner ear.

To select the intensity levels at which participants were or were not able to behaviorally categorize the auditory stimuli (as done in Kouider & Dupoux, 2005; and Daltrozzo et al., 2011), eight other participants (mean age 23.72 ± 1.17 years, median value of 22.09 years, min = 21.20 years, max = 32.87 years, five women) took part in two pretest experiments. To obtain a better view of the processing involved at each intensity level, these participants first performed a detection experiment (presence/absence), in which silences were randomly presented in 3.25% of the trials between words, pseudo-words and complex sounds. These stimuli were randomly presented at five intensity levels in both the inaudible and audible ranges (from 0 to 20 dB-A by 5-dB steps). The second experiment consisted of a lexical decision task (words/pseudowords), in which words and pseudo-words were randomly but equiprobably presented at five intensity levels (from 0 to 20 dB-A by 5-dB steps). The results were analysed with the Signal Detection Theory method (Macmillan & Creelman, 2005) as an objective threshold measure. At 5 dB-A, the participants could detect the stimuli 51.50% of the time (d' = 1.94), but could not categorize the stimuli as words or pseudo-words (d' = 0.12, not significantly different from zero, $t_7 = 0.74$, P = 0.48). At 20 dB-A, the participants could detect the stimuli 97.86% of the time (d' = 3.84), and were able to categorize the stimuli as words or pseudo-words 64.50% of the time (d' = 0.81, significantly different from zero, $t_7 = 3.57$, P = 0.009). Based on these results, we retained two levels of presentation for the electrophysiological study: (i) 5 dB-A, at which the stimuli were not successfully categorized (labeled 'uncategorized' stimuli condition); and (ii) 40 dB-A (i.e. 20 dB-A above the loudest level of the pretest), at which the stimuli were expected to be successfully categorized (labeled 'categorized' stimuli condition). To ensure that the presentation levels effectively induced the same behavioral performance of categorization as those observed in the pretest, the participants had to provide a motor response in 10% of the trials in the electrophysiological study (see below).

Design and procedure

Each participant sat in a sound-attenuated electromagnetically shielded booth while the EEG was recorded. A white fixation cross was presented in the middle of the screen 100–500 ms (by 100-ms steps) before the presentation of a silence or a stimulus and until its end. Before the start of a new trial, the participants were shown a black screen for 1200 ms to allow them to blink. Each participant was presented with 180 stimuli from each type (word, pseudo-word, complex sound) at each presentation level (categorized stimuli at 40 dB-A; uncategorized stimuli at 5 dB-A) and 180 silences (14.29% of the trials), forming a total of 1260 trials that were ran-

 $^{^{1}}$ In a pretest, five other participants (mean age 26 ± 2.1 years, two women) evaluated the phonological similarities of the pseudo-words to words. The participants had to determine if the pronounced pseudo-words sounded like a word and, if so, they were instructed to write down the corresponding word. All pseudo-words for which words were reported from at least two participants were eliminated.

domly presented in four blocks. Each stimulus was presented only once to a participant (i.e. the same stimuli could not be presented to the same participant at the two presentation levels), but all stimuli were presented at the two presentation levels across participants. The participants were instructed to keep their eyes opened and to focus on the white cross. They were not informed that some stimuli were presented at a low intensity level. For each trial, participants had to categorize the presented stimulus as word, pseudo-word, complex sound or silence. In 10% of the trials, a question mark replaced the fixation cross after the stimulus offset. In these cases, the participants had to provide a motor response on the keyboard as quickly and as accurately as possible by pressing Mot to answer 'words', PseudoMot to answer 'pseudo-words', Son to answer 'complex sounds' and Silence to answer 'silence'. The response buttons were attributed to s, f, h or k buttons on the keyboard, and were counterbalanced between participants. All other response buttons of the keyboard were hidden with a paper mask. All conditions were equally and randomly evaluated. The behavioral responses were analysed in order to determine whether lexical categorization processing was successful in each experimental condition (at 40 dB-A and at 5 dB-A presentation levels). The trials with motor responses were not included in the EEG analyses.

Apparatus

Words and pseudo-words were recorded (32 bits, 44.1 kHz) using a Røde NT1 microphone, a Behringer Ultragain preamplifier, and a VxPocket V2 Digigram soundcard. The mean level of presentation was calibrated (American National Standard Institute, 1995) to reach 80 dB-A in a standard artificial ear (Larson Davis AEC101 and 824). All stimuli were played through a soundcard (Creative Sound-blaster Audigy 2) followed by an analog attenuator (TDT PA4, one for each channel) that applied a fixed 40-dB attenuation. This attenuation was analog rather than digital to prevent acoustical distortion at low levels of presentation. All stimuli were binaurally presented to participants through comfortable headphones (Sennheiser HD 250 Linear II) connected to a headphone buffer (TDT HB6).

EEG signals from 60 Ag–AgCl electrodes (positioned according to the International 10–20 system) referenced to the nose, were amplified using the Brain Quick SD64 Micromed system and sampled at 512 Hz (16 bits) using an analog bandpass filter of 0.05– 128 Hz. Horizontal electrooculograms (EOGs) were monitored via bipolar electrodes positioned at the outer canthi of both eyes, and vertical EOGs were monitored via bipolar electrodes positioned below and above the left eye. A ground electrode was placed at FPz site and the impedance at all electrodes was kept below 5 k Ω .

EEG analysis

Both event-related potentials (ERPs) and event-related spectral perturbations (ERSPs) were analysed using custom scripts written in Matlab R2008b (The MathWorks) and the Eeglab 6.01b toolbox (Delorme & Makeig, 2004). Signals containing non-stereotypical artifacts, including high-amplitude, high-frequency muscle noise and electrode cable movements, were rejected. Stereotypical artifacts such as eye movements and eye blinks were corrected with an extended Infomax independent component analysis (Lee *et al.*, 1999) implemented in Eeglab. In this analysis, the data were separated into 60 independent components per participant. One independent component representing non-brain artifacts was removed per participant based on a visual inspection of their scalp topographies, time courses and frequency spectra. EEG data were reconstructed from the remaining component activations (Delorme & Makeig, 2004). The data from one participant were excluded from the analysis because of large artifacts in the signal. For each participant, an average of 37 trials has been removed for each condition (leaving a mean of 125 trials for statistical analyses in each condition). All statistical analyses were based on ANOVAS. If P < 0.05 after applying the Greenhouse–Geisser correction, the interaction was considered significant; *post hoc* tests (Fisher's LSD) were applied to significant interactions.

For the ERP analysis, the reconstructed signal was segmented from -200 to +1000 ms around the stimulus onsets, and the baseline was corrected from 200 ms before the stimulus onset. For each electrode, grand average waveforms were computed based on the condition of presentation (categorized, uncategorized, silence) and/or on the nature of the stimulation (word, pseudo-word, complex sound). Four classical auditory potentials were evoked by the categorized stimuli (Fig. 1): N1 (50-200 ms); P2 (200-450 ms); N2 (450-700 ms); and P3 (700-1000 ms). As observed in previous studies presenting auditory complex stimuli (e.g. Woods & Elmasian, 1986; Simanova et al., 2010), classical auditory potentials were delayed compared with those observed in studies where pure-tones are presented. To highlight processing differences related to the categorization task, the mean amplitudes of individual ERPs in each time window were analysed using two-way ANOVAS with repeated measures of the Condition of Presentation (categorized, uncategorized, silence) and Electrode Position (Fz, Cz, Pz). Classical auditory components are expected to show maximal amplitudes in the categorized condition for N1 and P2 at Cz, for N2 at Fz, and for P3 at Pz (see Picton et al., 1974; Mazaheri & Picton, 2005). The results of the ANOVAS confirmed that the expected pattern was indeed observed: amongst these three electrodes, mean amplitudes in the categorized condition were maximal at Cz for N1 component for nine participants and for P2 component for nine participants, at Fz for N2 component for eight participants, and at Pz for P3 component for seven participants. To obtain the topographic picture as a function of the phonological and/or the lexical status of the categorized stimuli, the mean amplitudes of individual ERPs were analysed using threeway ANOVAS with repeated measures of the Nature of Stimulation (word, pseudo-word, complex sound), Lateral Position (left, right) and Antero-Posterior Position (anterior, posterior). For this purpose, four quadrants (of 11 electrodes each) were defined: a right anterior site; a left anterior site; a right posterior site; and a left posterior site.²

For the ERSP analysis, the reconstructed signal was segmented from -1000 to +2000 ms around the stimulus onsets. A trial-by-trial time-frequency analysis was computed with complex Morlet wavelets using 5 cycles at 4 Hz (i.e. 1250 ms) expanding linearly to 12 cycles at 48 Hz (i.e. 250 ms), generating 200 spectra of 177 points each (thus between 4 and 48 Hz) between -301.9 and 1293.9 ms. For each electrode, the signals were averaged based on the condition of presentation (categorized, uncategorized, silence) and/or on the nature of the stimulation (word, pseudo-word, complex sound). The ERSPs were observed in five spectro-temporal windows relative to the baseline (Fig. 2): theta (θ : 4–7 Hz) between 0 and 500 ms; alpha (a: 8-12 Hz) between 0 and 500 ms; beta (\beta: 15-30 Hz) between 150 and 400 ms; early gamma (γ_a : 30-42 Hz) between 0 and 100 ms; and late gamma ($\gamma_b{:}$ 30–48 Hz) between 150 and 450 ms. As in the ERP analysis, the average power of the five spectro-temporal windows was analysed using two-way ANOVAS with repeated measures of the Condition of Presentation (categorized, uncategorized,

²The quadrants contained the following electrodes: left anterior with FP1, AF3, AF7, F7, F5, F3, F1, FT7, FC5, FC3, FC1; right anterior with FP2, AF8, AF4, F8, F6, F4, F2, FT8, FC6, FC4, FC2; left posterior with O1, PO7, PO3, P7, P5, P3, P1, TP7, CP5, CP3, CP1; right posterior with O2, PO8, PO4, P8, P6, P4, P2, TP8, CP6, CP4, CP2.



FIG. 1. (A) ERPs (N1, P2, N2 and P3) are presented as a function of Condition of Presentation (categorized, uncategorized and silence) for each Electrode (Fz, Cz, Pz). The blue curve represents the categorized condition, the magenta curve represents the uncategorized condition, and the gray curve represents the silence condition. The standard error is shown for each curve. Only the categorized condition presented classical auditory ERPs. (B) ERPs in the categorized condition are presented as a function of Nature of Stimulation (word, pseudo-word and complex sound) from one electrode in each quadrant (F5, F6, P5, P6). The red curve represents the categorized corditions are presented as a function in transparency. On the standard error is shown for each curve. (C) The ERP differences between conditions are presented, showing the 95% confidence interval in transparency. On the left, the differences between each Condition of Presentation (categorized and silence) are shown in Cz. On the right, the differences between each Nature of Stimulation (word, pseudo-word and complex sound) are shown in P6. The color used for each difference is the mix of the two colors corresponding to the two operands of the difference in (A) or (B).

silence) and Electrode Position (Fz, Cz, Pz). The ERSPs were also analysed as a function of the nature of the stimuli for the categorized and uncategorized conditions. As previous studies have observed, alpha (8-12 Hz), low-beta (15-20 Hz) and/or high-beta (20-30 Hz) lateralized power differences over the frontal and/or temporal regions for linguistic and non-linguistic stimuli (Lutzenberger et al., 1994; Pulvermüller et al., 1996; Palva et al., 2002; Luo et al., 2005; Krause et al., 2006; Shahin et al., 2009), four electrode groups were selected: F4, F6 and FC6 (right frontal); F3, F5 and FC5 (left frontal); T8, TP8 and FT8 (right temporal); and T7, TP7 and FT7 (left temporal). Alpha (8-12 Hz; 0-500 ms), low-beta (15-20 Hz; 200-450 ms) and high-beta (20-30 Hz; 150-400 ms) windows were analysed using four-way ANOVAS with repeated measures of the Condition (categorized, uncategorized), Nature of Stimulation (words, pseudowords, complex sounds), Lateral Position (left, right) and Fronto-temporal Position (frontal, temporal).

Results

Behavioral data

The behavioral analyses (Table 1) showed that silences were correctly categorized as 'silence' in 89.06% of the cases. Stimuli were correctly categorized in 85.74% of the cases when presented at

40 dB-A, and in 16.33% of the cases when presented at 5 dB-A. The participants detected a stimulus when they did not categorize it as a silence. In these cases, one out of three chances of finding the correct response (word, pseudo-word or complex sound) was expected as chance level. Therefore, the participants' categorization performance for the detected 5 dB-A stimuli was not different from the chance level for words, pseudo-words and complex sounds (*t*-tests individual chance level were $t_{13} = -1.27$, P = 0.23; $t_{13} = 0.67$, P = 0.51; $t_{13} = 1.53$, P = 0.15, respectively). The stimuli presented at 5 dB-A were discriminated from silences (d' = 1.30) but never successfully categorized (as usually reported in studies investigating unconscious processing; for a review, see Kouider & Dehaene, 2007). This result replicated the results of the pretest where the participants were not able to perform the categorization task above the objective discrimination threshold in the 5 dB-A condition.

EEG data

ERPs: as a function of the condition of presentation

The Condition of Presentation had a significant effect on the amplitude of N1 ($F_{2,24} = 6.69$, P = 0.005), P2 ($F_{2,24} = 12.12$, P < 0.001), N2 ($F_{2,24} = 29.27$, P < 0.001) and P3 ($F_{2,24} = 11.57$, P < 0.001). As shown in Fig. 1A, for any of the ERPs and elec-



FIG. 2. ERSP activity as a function of Condition of Presentation. Left panel: activity maps (average of rectangles in the right panel) for each frequency band (theta, alpha, beta, early gamma and late gamma). Right panel: ERSPs as a function of Condition of Presentation (categorized, uncategorized and silence) at Fz, Cz and Pz sites.

TABLE 1. Behavioral results

Stimuli → Responses ↓	Uncategorized (5 dB-A)			Categorized (40 dB-A)			
	Word	Pseudo-word	Complex sound	Word	Pseudo-word	Complex sound	Silences
Word	10.4% (14.3%)	8.8% (9.4%)	5.6% (9.4%)	79.9% (21.0%)	13.6% (7.5%)	0.0% (0.0%)	4.0% (7.1%)
Pseudo-word	19.1% (16.7%)	17.0% (19.4%)	7.5% (7.5%)	18.3% (18.7%)	85.0% (7.8%)	0.4% (1.6%)	3.1% (4.1%)
Complex sound	21.9% (16.5%)	16.0% (14.1%)	21.8% (22.4%)	1.3% (3.5%)	0.9% (2.2%)	92.0% (25.6%)	3.8% (3.8%)
Silence	48.7% (27.3%)	58.2% (21.3%)	65.0% (22.4%)	0.4% (1.6%)	0.4% (1.6%)	7.6% (25.7%)	89.1% (13.1%)

Average percentage of responses to each type of stimulus over the 14 participants. The numbers in bold correspond to the correct categorization. The numbers in brackets are the standard deviation.

trodes, only the categorized condition evoked classical auditory ERPs and therefore differed from the silence condition (P < 0.007) and from the uncategorized condition (P < 0.003). The ERPs in the uncategorized condition were not different from those obtained in the silence condition (P > 0.126).

ERPs: as a function of the nature of the categorized stimuli

The ERPs for one electrode of each quadrant are shown in Fig. 1B. N1 response differentiated between the Nature of Stimulation at posterior sites (interaction with Antero-Posterior Position: $F_{2,24} = 8.56$, P = 0.002), where N1 response was greater for complex sounds than for pseudo-words (P < 0.005) and words (P < 0.001), and greater for pseudo-words than for words (P < 0.001). The Nature of the Stimulation affected also P2 ($F_{2,24} = 4.13$, P = 0.029) and N2 ($F_{2,24} = 10.05$, P < 0.001) responses across all quadrants; greater responses were observed for speech sounds (i.e. words and pseudo-words) than for non-speech sounds (i.e. complex sounds; P < 0.001). P2 and N2 responses were also modulated by the electrode position for some stimuli (Nature of Stimulation \times Lateral Position: $F_{2,24} = 4.58$, P = 0.021 for P2; $F_{2,24} = 6.23$, P = 0.007 for N2): P2

and N2 responses were greater for complex sounds in the left than in the right hemisphere (P < 0.001); and N2 response was greater for pseudo-words in the left than in the right hemisphere (P < 0.012). P3 responses (Nature of Stimulation × Antero-Posterior Position: $F_{2,24} = 13.60$, P < 0.001) were greater for words and pseudo-words than for complex sounds at posterior sites (P < 0.001).

ERSPs: as a function of the condition of presentation

The ERSPs as a function of the condition of presentation are shown in Fig. 2. The ERSPs observed in the lower frequency bands revealed that the power for the categorized condition was different from both the uncategorized and the silence conditions, while the power for the uncategorized and the silence conditions did not differ. This pattern was observed at all recording sites in the theta frequency band ($F_{2,24} = 49.56$, P < 0.001), with a greater increase in power for the categorized condition than for the uncategorized or the silence conditions (P < 0.001), while the latter two conditions did not differ (P = 0.72). All participants demonstrated this pattern. In the alpha frequency band, this pattern was observed at Fz and Cz sites ($F_{4,48} = 5.60$, P < 0.001), with a smaller decrease in power for

the categorized condition than for the uncategorized or the silence conditions ($P \le 0.001$), while the latter two conditions did not differ (P > 0.50). All participants demonstrated this pattern. Finally, in the beta frequency band, this pattern was observed at Fz site ($F_{4,48} = 3.17$, P = 0.022), with a greater decrease in power for the categorized condition than for the uncategorized or the silence conditions (P < 0.030), while the latter two conditions did not differ (P > 0.232). Nine of 13 participants demonstrated this pattern.

The ERSPs observed in the higher frequency bands revealed a different pattern: the power was different between the uncategorized condition and the silence condition. In the beta frequency band, this difference in power was observed at Pz site. The decrease in power was smaller for the silence condition than for the uncategorized or the categorized conditions (P < 0.005). This decrease in power was also smaller for the uncategorized condition than for the categorized condition (P = 0.017). Nine of 13 participants demonstrated this pattern. In the early gamma frequency band at all recording sites ($F_{2,24} = 8.63$, P = 0.003), a greater increase in power was observed for the silence condition than for the uncategorized or the categorized conditions (P < 0.004), while the categorized and the uncategorized conditions did not differ (P = 0.568). Twelve of 13 participants demonstrated this pattern. No power differences were observed in the late gamma (150-450 ms) at any site.

ERSPs: as a function of the nature of the stimuli

In the alpha frequency band, no effect of the Nature of Stimulation was observed. The decrease in power depended on both the Condition and the Fronto-temporal Position ($F_{1,12} = 11.10$, P = 0.006). A greater decrease in power for the uncategorized condition than for the categorized condition was observed at frontal sites (P < 0.001) only. The effect of Condition also depended on the Lateral Position ($F_{1,12} = 4.84$, P = 0.048). The decrease in power for the categorized and the uncategorized conditions was lateralized on the right hemisphere (P < 0.001), where this decrease in power was greater for the uncategorized condition than for the categorized condition (P < 0.048).

In the low-beta (15–20 Hz) frequency band (Fig. 3), the decrease in power was modulated by the Nature of Stimulation as a function of the sites (Nature of Stimulation × Lateral Position × Fronto-temporal Position: $F_{2,24} = 6.31$, P = 0.006). At temporal sites, a greater decrease in power was observed on the left hemisphere for pseudowords and for complex sounds than for words (P < 0.048 and P < 0.017, respectively), while there was no difference in power between pseudo-words and complex sounds (P > 0.353). On the right hemisphere, the low-beta power was equivalent for all stimuli (P > 0.056). Importantly, this difference in low-beta power between lexical (i.e. words) and non-lexical (i.e. pseudo-words and complex sounds) stimuli did not depend on whether the stimuli were or were not categorized, as none of the interactions involving the factor Condition was significant (all P > 0.269).

In the high-beta (20–30 Hz) frequency band, the difference in power between lexical and non-lexical stimuli was observed in the left regions (Condition × Nature of Stimulation × Lateral Position: $F_{2,24} = 6.34$, P = 0.006). In the left regions, the decrease in power was greater for categorized pseudo-words than for categorized words (P = 0.026). In addition, a greater decrease in power was observed for uncategorized complex sounds than for uncategorized words (P = 0.003) and for uncategorized pseudo-words (P = 0.016). In the right regions, the high-beta power was equivalent for all stimuli (P > 0.096).

Discussion

The results of this study show different EEG responses to behaviorally categorized stimuli, uncategorized stimuli and silences. First, the results suggest that uncategorized stimuli were not processed because they evoked neither potentials nor theta activity. However, uncategorized stimuli displayed neural high-frequency responses that were more similar to the responses observed for behaviorally categorized stimuli than to those observed for silences, notably reflected by a decrease in power in the beta (20–30 Hz) frequency band at



FIG. 3. Low-beta activity for categorized and uncategorized words, pseudowords and complex sounds. Top and bottom panels: low-beta power decrease in percent of baseline as a function of Nature of Stimulation (words, pseudowords and complex sounds), Condition (categorized, uncategorized), Lateral Position (right, left), averaged across temporal electrodes (T7, TP7, FT7/T8, TP8, FT8) and participants. The box limits represent the upper and lower quartiles, and the dotted line indicates the median. The thick solid line represents the mean. The boxplot whiskers extend to the most extreme data points within \pm 1.5 times the interquartile range. The outliers, i.e. data points outside this range, are plotted directly. Attached to the mean are shown the 95% confidence intervals for the differences between type of stimulus. The color indicates with which other type of stimulus the difference is made; the direction depends on the sign of the difference. Middle panel: maps of ERSPs as a function of Condition (categorized, uncategorized) and Nature of Stimulation (words, pseudo-words and complex sounds) of the low-beta frequency band.

posterior site. This suggests that uncategorized stimuli were partially processed. This assumption was confirmed by the fact that the categorized and uncategorized stimuli yielded similar left hemispheric differences between lexical (words) and non-lexical sounds (pseudowords and complex sounds) in the low-beta (15–20 Hz) frequency band, demonstrating that behaviorally uncategorized stimuli were processed, probably at the lexical level.

Cerebral differences associated to the presence/absence of auditory stimuli

Classical auditory N1-P2-N2-P3 components were observed for categorized stimuli, but not for uncategorized stimuli or silences. This result is not surprising as ERP amplitudes decrease when intensity is reduced (Hillyard *et al.*, 1971; Polich *et al.*, 1986) to the psychophysical threshold or near it (Suzuki *et al.*, 1976; Musiek *et al.*, 2005). The increase in theta power for categorized stimuli may be associated with the evoked responses, as Klimesch *et al.* (2007) have shown that ERPs reflect the superposition of an increase in evoked theta frequency band over a decrease in alpha frequency band.

The increasing theta power could be linked to the sensory coding of the stimulus (see review in Buzsáki, 2006), but not the decreasing alpha power. Indeed, a posterior decrease in alpha power was recorded after all stimuli and silences, and a greater fronto-central decrease in alpha power was observed for uncategorized stimuli and silences. If alpha power was not directly linked to sensory coding, it might rather be dependent of attentional or 'alertness' processes (see review in Klimesch et al., 2006). In our study, the posterior decrease in alpha power likely indicates that participants had a constant level of alertness during the task (Klimesch et al., 1996), while the frontal decrease in alpha power is probably linked to the expectation of the stimulus (Yamagishi et al., 2003). Supporting this latter interpretation, the combined probability of occurrence of barely audible stimuli at 5 dB-A and silences was higher than the probability of occurrence of clearly audible stimuli at 40 dB-A, the former stimuli becoming more expected.

A greater decrease in beta (20–30 Hz) power was observed for categorized and uncategorized stimuli than for silences in posterior regions; the biggest decrease being observed for categorized stimuli. This decrease could be linked to the perception of the stimulus, as differences between perceived and unperceived stimuli have previously been observed in the beta frequency band (Gail *et al.*, 2004; Luo *et al.*, 2005; Palva *et al.*, 2005; Gaillard *et al.*, 2009). However, in the present study, it was not possible to distinguish the contribution of awareness from the processing of the stimulus energy in the differential brain responses to categorized and uncategorized stimuli. Based on these considerations, it cannot be excluded that variations in beta power could also indicate that even if the stimuli presented at 5 dB-A were not behaviorally categorized, they were nevertheless partially detected.

An early, transient increase in gamma power was observed only for silences, followed by a decrease in gamma power for all stimuli and silences. This decrease in gamma power is classically observed after attended stimuli (Fell *et al.*, 1997; see review in Herrmann & Knight, 2001), suggesting that participants had a constant level of alertness during the task for all conditions (as previously suggested for the posterior alpha power). However, the early increase in gamma power in response to silences is more difficult to explain. Because this early increase in gamma power is neither observed for categorized nor uncategorized stimuli, this could reflect the early increase in gamma power for categorized stimulis. Thus, the absence of early increase in gamma power for categorized stimuli could indicate that the participants were

awaiting the offset of the word recognition process (about 250 ms after stimulus onset; see Pulvermüller *et al.*, 2005) in order to correctly perform the categorization task. When the participants detected that a stimulus was actually presented, they required more information about the stimulus to categorize it as word, pseudo-word or complex sound. Interestingly, the absence of early increase in gamma power for uncategorized stimuli could confirm that these stimuli were not processed as silences, and that participants were probably waiting for more information regarding the stimulus.

Lexical processing of categorized stimuli

Hemispheric differences were observed between the three types of stimuli, when they were successfully categorized. Differences observed for P2 and N2 are reminiscent of the speech specialization of the left hemisphere (Démonet et al., 1992; Zatorre et al., 1992). As the lexical stimuli used were all monosyllabic (as defined in Lexique 2), and because the participants had to perform the task as quickly as possible, the words might have been lexically categorized as words before the end of word occurred (McClelland & Elman, 1986; Marslen-Wilson, 1987). This explanation could be supported by the fact that an effect of lexicality was observed for ERP responses as early as the N1 time window (50-200 ms). As reported in the visual modality in previous studies (comparing words and non-words in Martin et al., 2006; or comparing words and pseudowords in Coch & Mitra, 2010), ERP amplitudes observed in similar early time window (before 240 ms) seem to be influenced by topdown information regarding lexicality.

Most importantly, the results of the present study provide support in the auditory modality for the results obtained in the visual modality by Lutzenberger et al. (1994) regarding the link between the beta frequency band and lexical processing. Indeed, we observed a greater decrease in low-beta (15-20 Hz) power for non-lexical stimuli (pseudo-words and complex sounds) than for lexical stimuli (words) in the left hemisphere. Thus, our data support the assumption that a reduction of high-frequency cortical responses after pseudo-words and after complex sounds reflects a breakdown of cognitive processing induced by an uncommon stimulus for which no lexical representation is available, like the failure of lexical access described in previous studies (Lutzenberger et al., 1994; Pulvermüller et al., 1996). As pseudo-words used less frequent phoneme pairs than words (and although this only affected 11% of the pair instances), one may suggest that the decrease in low-beta power could also reflect more strenuous linguistic processing at a pre-lexical level. However, the decrease in low-beta power for complex sounds (which do not have any phonological content) was similar to that observed for pseudowords. This would tend to suggest that the differences observed in the low-beta frequency band between lexical and non-lexical stimuli are not due to pre-lexical differences. Moreover, significant differences between words and pseudo-words, but also between words and complex sounds, in the low-beta and/or high-beta frequency bands, are also more compatible with an explanation in terms of lexicality effect (Pulvermüller et al., 1999; Palva et al., 2002; Bastiaansen & Hagoort, 2006; Shahin et al., 2009).

Lexical processing of uncategorized stimuli

The most important result obtained in the present study is that lowbeta (15–20 Hz) and high-beta (20–30 Hz) differences between lexical and non-lexical stimuli did not significantly depend on whether the stimuli were successfully categorized or not. While the uncategorized stimuli evoked neither classical auditory potentials nor theta

activity, the results suggest that they induced low-beta power differences according to the lexical nature of the stimuli in the left temporal hemisphere. In the same way, the uncategorized stimuli induced also high-beta power differences according to the phonological nature of the stimuli in the left hemisphere. These findings would support the hypothesis that behaviorally uncategorized stimuli elicited neural responses similar to behaviorally categorized stimuli, i.e. that behaviorally uncategorized stimuli could be categorized without awareness. Such a pattern of neural responses, specific to the linguistic content of the stimuli, cannot be explained by differences in intensity or duration (as the observed effects appeared between 150 and 450 ms, and that all stimuli have an average duration of more than 500 ms). These findings indicate that the behaviorally uncategorized stimuli were actually partially processed at the phonological level and, most likely, at the lexical level. This result is in accordance with studies done in the visual modality suggesting that the lexical or semantic processing can occur without awareness (for a meta-analysis, see Van den Bussche et al., 2009). This result is also in accordance with behavioral studies done in the auditory modality showing lexical processing without awareness (Daltrozzo et al., 2011; Signoret et al., 2011), and suggests that lexical processing is concomitant with a variation of low-beta (15-20 Hz) power near the left temporal regions, independently of any awareness of the category of the stimulus.

Finally, the results of the present study showed that lexical processing could be observed for stimuli that are not behaviorally categorized, but partially detected. Although the choice of a lexical task as a measure of awareness is largely accepted in studies investigating linguistic processing without awareness (for reviews, see Kouider *et al.*, 2007; Van den Bussche *et al.*, 2009), further research would be required to assess whether lexical processing could also occur in experimental conditions involving a detection task, as recently suggested in a behavioral study showing lexical processing without subjective detection (Signoret *et al.*, 2011).

Acknowledgements

The work described in this paper was supported by a doctoral grant from the Ministère de l'Éducation Nationale et de la Recherche of France, the grant ACI 'Junior Research Team', the grant ANR 'PICS' and the United Kingdom-Medical Research Council (grant G9900369 and programme MC_A060-5PQ70). We thank Olivier Bertrand for helpful comments in the preparation of this study, Arnaud Delorme for EEGLAB expertise, and Rachel Ellis as well as two anonymous reviewers for helpful comments on this paper.

Abbreviations

EEG, electroencephalography; EOG, electrooculography; ERP, event-related potential; ERSP, event-related spectral perturbation.

References

- American National Standard Institute (1995) S3.7-1995 (R2003), Methods for Coupler Calibration of Earphones. American National Standard Institute, New York.
- American National Standard Institute (2004) ANSI S3.6-2004, Specifications for Audiometers. American National Standard Institute, New York.
- Bastiaansen, M.C.M. & Hagoort, P. (2006) Oscillatory neuronal dynamics during language comprehension. Prog. Brain Res., 159, 179–196.
- Buzsáki, G. (2006) *Rhythms of the Brain*. Oxford University Press, Oxford, UK.
- Coch, D. & Mitra, P. (2010) Word and pseudoword superiority effects reflected in the ERP waveform. *Brain Res.*, **1329**, 159–174.
- Daltrozzo, J., Signoret, C., Tillmann, B. & Perrin, F. (2011) Subliminal semantic priming in speech. *PLoS ONE*, **6**, e20273.
- Davis, M.H. & Johnsrude, I.S. (2003) Hierarchical processing in spoken language comprehension. J. Neurosci., 23, 3423–3431.

- Delorme, A. & Makeig, S. (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods, **134**, 9–21.
- Démonet, J.F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.L., Wise, R., Rascol, A. & Frackowiak, R. (1992) The anatomy of phonological and semantic processing in normal subjects. *Brain*, **115**(Pt 6), 1753–1768.
- Eulitz, C., Eulitz, H., Maess, B., Cohen, R., Pantev, C. & Elbert, T. (2000) Magnetic brain activity evoked and induced by visually presented words and nonverbal stimuli. *Psychophysiology*, **37**, 447–455.
- Fell, J., Hinrichs, H. & Röschke, J. (1997) Time course of human 40 Hz EEG activity accompanying P3 responses in an auditory oddball paradigm. *Neurosci. Lett.*, 235, 121–124.
- Gail, A., Brinksmeyer, H.J. & Eckhorn, R. (2004) Perception-related modulations of local field potential power and coherence in primary visual cortex of awake monkey during binocular rivalry. *Cereb. Cortex*, 14, 300–313.
- Gaillard, R., Dehaene, S., Adam, C., Clémenceau, S., Hasboun, D., Baulac, M., Cohen, L. & Naccache, L. (2009) Converging intracranial markers of conscious access. *PLoS Biol.*, 7, e61.
- Giraud, A.L., Kell, C., Thierfelder, C., Sterzer, P., Russ, M.O., Preibisch, C. & Kleinschmidt, A. (2004) Contributions of sensory input, auditory search and verbal comprehension to cortical activity during speech processing. *Cereb. Cortex*, 14, 247–255.
- Herrmann, C.S. & Knight, R.T. (2001) Mechanisms of human attention: event-related potentials and oscillations. *Neurosci. Biobehav. Rev.*, 25, 465– 476.
- Hillyard, S.A., Squires, K.C., Bauer, J.W. & Lindsay, P.H. (1971) Evoked potential correlates of auditory signal detection. *Science*, **172**, 1357– 1360.
- Hoen, M., Meunier, F., Grataloup, C., Pellegrino, F., Grimault, N., Perrin, F., Perrot, X. & Collet, L. (2007) Phonetic and lexical interferences in informational masking during speech-in-speech comprehension. *Speech Comm.*, 49, 905–916.
- Klimesch, W., Doppelmayr, M., Russegger, H. & Pachinger, T. (1996) Theta band power in the human scalp EEG and the encoding of new information. *NeuroReport*, **7**, 1235–1240.
- Klimesch, W., Hanslmayr, S., Sauseng, P., Gruber, W., Brozinsky, C.J., Kroll, N.E.A., Yonelinas, A.P. & Doppelmayr, M. (2006) Oscillatory EEG correlates of episodic trace decay. *Cereb. Cortex*, 16, 280–290.
- Klimesch, W., Sauseng, P. & Hanslmayr, S. (2007) EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res. Rev.*, 53, 63–88.
- Kouider, S. & Dehaene, S. (2007) Levels of processing during non-conscious perception: a critical review of visual masking. *Philos. T. Roy. Soc. B*, 362, 857–875.
- Kouider, S. & Dupoux, E. (2005) Subliminal speech priming. Psychol. Sci., 16, 617–625.
- Kouider, S., Dehaene, S., Jobert, A. & Le Bihan, D. (2007) Cerebral bases of subliminal and supraliminal priming during reading. *Cereb. Cortex*, **17**, 2019–2029.
- Kouider, S., de Gardelle, V., Dehaene, S., Dupoux, E. & Pallier, C. (2010) Cerebral bases of subliminal speech priming. *NeuroImage*, 49, 922–929.
- Krause, C.M., Grönholm, P., Leinonen, A., Laine, M., Säkkinen, A.-L. & Söderholm, C. (2006) Modality matters: the effects of stimulus modality on the 4- to 30-Hz brain electric oscillations during a lexical decision task. *Brain Res.*, **1110**, 182–192.
- Lee, T.W., Girolami, M. & Sejnowski, T.J. (1999) Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput.*, **11**, 417–441.
- Luo, H., Husain, F.T., Horwitz, B. & Poeppel, D. (2005) Discrimination and categorization of speech and non-speech sounds in an MEG delayedmatch-to-sample study. *NeuroImage*, 28, 59–71.
- Lutzenberger, W., Pulvermüller, F. & Birbaumer, N. (1994) Words and pseudowords elicit distinct patterns of 30-Hz EEG responses in humans. *Neurosci. Lett.*, **176**, 115–118.
- Macmillan, N.A. & Creelman, C.D. (2005) *Detection Theory: a User's Guide*, 2nd edn. Lawrence Erlbaum Associates, Mahwah NJ.
- Marslen-Wilson, W.D. (1987) Functional parallelism in spoken word-recognition. Cognition, 25, 71–102.Martin, C.D., Nazir, T., Thierry, G., Paulignan, Y. & Demonet, J.-F. (2006)
- Martin, C.D., Nazir, T., Thierry, G., Paulignan, Y. & Demonet, J.-F. (2006) Perceptual and lexical effects in letter identification: an event-related potential study of the word superiority effect RID C-2688-2008 RID B-3227-2008. *Brain Res.*, **1098**, 153–160.
- Mazaheri, A. & Picton, T.W. (2005) EEG spectral dynamics during discrimination of auditory and visual targets. *Cogn. Brain Res.*, 24, 81–96.
- McClelland, J.L. & Elman, J.L. (1986) The TRACE model of speech perception. Cogn. Psychol., 18, 1–86.

- Musiek, F.E., Froke, R. & Weihing, J. (2005) The auditory P300 at or near threshold. J. Am. Acad. Audiol., 16, 698–707.
- New, B., Pallier, C., Brysbaert, M. & Ferrand, L. (2004) Lexique 2: a new French lexical database. *Behav. Res. Meth. Ins. C.*, 36, 516–524.
- New, B., Brysbaert, M., Veronis, J. & Pallier, C. (2007) The use of film subtitles to estimate word frequencies. *Appl. Psycholing*, 28, 661–677.
- Oldfield, R.C. (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Palva, S., Palva, J.M., Shtyrov, Y., Kujala, T., Ilmoniemi, R.J., Kaila, K. & Näätänen, R. (2002) Distinct gamma-band evoked responses to speech and non-speech sounds in humans. J. Neurosci., 22, RC211.
- Palva, S., Linkenkaer-Hansen, K., Näätänen, R. & Palva, J.M. (2005) Early neural correlates of conscious somatosensory perception. J. Neurosci., 25, 5248–5258.
- Perrin, F. & Grimault, N. (2005) Fonds Sonores (Version 1.0) [Sound samples] [WWW Document]. Available at http://olfac.univ-lyon1.fr/unite/equipe-02/FondsSonores.html
- Picton, T.W., Hillyard, S.A., Krausz, H.I. & Galambos, R. (1974) Human auditory evoked potentials. I: Evaluation of components. *Electroen. Clin. Neuro.*, 36, 179–190.
- Polich, J., Ehlers, C.L., Otis, S., Mandell, A.J. & Bloom, F.E. (1986) P300 latency reflects the degree of cognitive decline in dementing illness. *Electroen. Clin. Neuro.*, 63, 138–144.
- Pulvermüller, F., Eulitz, C., Pantev, C., Mohr, B., Feige, B., Lutzenberger, W., Elbert, T. & Birbaumer, N. (1996) High-frequency cortical responses reflect lexical processing: an MEG study. *Electroen. Clin. Neuro.*, 98, 76–85.

- Pulvermüller, F., Lutzenberger, W. & Preissl, H. (1999) Nouns and verbs in the intact brain: evidence from event-related potentials and high-frequency cortical responses. *Cereb. Cortex*, 9, 497–506.
- Pulvermüller, F., Shtyrov, Y. & Ilmoniemi, R. (2005) Brain signatures of meaning access in action word recognition. J. Cogn. Neurosci., 17, 884–892.
- Shahin, A.J., Picton, T.W. & Miller, L.M. (2009) Brain oscillations during semantic evaluation of speech. *Brain Cogn.*, 70, 259–266.
- Signoret, C., Gaudrain, E., Tillmann, B., Grimault, N. & Perrin, F. (2011) Facilitated auditory detection for speech sounds. *Front. Psychol.*, **2**, 176.
- Simanova, I., van Gerven, M., Oostenveld, R. & Hagoort, P. (2010) Identifying object categories from event-related eeg: toward decoding of conceptual representations. *PLoS ONE*, 5, e14465.
- Suzuki, T., Yamamoto, K., Taguchi, K. & Sakabe, N. (1976) Reliability and validity of late vertex-evoked response audiometry. *Audiology*, 15, 357– 369.
- Van den Bussche, E., Van den Noortgate, W. & Reynvoet, B. (2009) Mechanisms of masked priming: a meta-analysis. *Psychol. Bull.*, 135, 452–477.
- Woods, D.L. & Elmasian, R. (1986) The habituation of event-related potentials to speech sounds and tones. *Electroen. Clin. Neuro.*, 65, 447–459.
- Yamagishi, N., Callan, D.E., Goda, N., Anderson, S.J., Yoshida, Y. & Kawato, M. (2003) Attentional modulation of oscillatory activity in human visual cortex. *NeuroImage*, 20, 98–113.
- Zatorre, R.J., Evans, A.C., Meyer, E. & Gjedde, A. (1992) Lateralization of phonetic and pitch discrimination in speech processing. *Science*, 256, 846– 849.