HAL
open science

On the locus of talker-specificity effects in spoken word recognition: an ERP study with dichotic priming

Sophie Dufour, Dierdre Bolger, Stéphanie Massol, Phillip J. Holcomb, Jonathan Grainger

To cite this version:

Sophie Dufour, Dierdre Bolger, Stéphanie Massol, Phillip J. Holcomb, Jonathan Grainger. On the locus of talker-specificity effects in spoken word recognition: an ERP study with dichotic priming. Language, Cognition and Neuroscience, 2017, 32 (10), pp.1273-1289. 10.1080/23273798.2017.1335421 . hal-01764693

HAL Id: hal-01764693

https://amu.hal.science/hal-01764693

Submitted on 15 Apr 2021

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

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

On the locus of talker-specificity effects in spoken word recognition:

An ERP study with dichotic priming

Sophie Dufour[1,2], Dierdre Bolger[2], Stephanie Massol[3], Phillip J. Holcomb[4], & Jonathan Grainger[2,5]

1 : Aix-Marseille Université, CNRS, LPL, UMR 7309, 13100 Aix-en-Provence, France

2 : Brain and Language Research Institute, Aix-Marseille Université, Aix-en-Provence, France

3 : Dpto. Psichologia Basica, Clinica y Psichobiologia, Universitat Jaume I, Castellon de la Plana, Spain

4 : San Diego State University, San Diego, USA

5 : Laboratoire de Psychologie Cognitive, Aix-Marseille Université & CNRS

Sophie Dufour
Laboratoire Parole et Langage
Aix-Marseille Université, CNRS
5, avenue Pasteur
13604 Aix-en-Provence
France
Email: sophie.dufour@lpl-aix.fr

Abstract: We used event-related potentials to examine the precise moment at which talker-specific information comes into play during spoken word recognition. Dichotic repetition priming was examined with primes presented in the left unattended ear and targets presented in the right attended ear, and we manipulated both word frequency and talker identity. A clear repetition priming effect was observed in an early time-window spanning 100-200 ms post-target onset, and the effect continued after target word offset in a time-window between 650 and 800 ms. Crucially, we observed that talker change caused a diminution in repetition priming only in the N400 time-window, and only for low frequency words but not for high frequency words. Together, our findings suggest that spoken word recognition relies primarily on abstract representations, and that talker-specific information mainly affects later stages of this process, namely lexical selection.

Key words: spoken word recognition; abstract representations; talker-specificity effects; dichotic priming; ERPs

A major issue in spoken word recognition research concerns the process by which the incoming speech signal is mapped onto representations of words stored in long-term memory. Clearly, the way in which the human word recognition system deals with spoken language is remarkable, since word recognition occurs effortlessly and with few errors under a wide variety of conditions that could be disruptive. Indeed, a given word is never spoken twice in exactly the same way and presents acoustic and phonetic differences according to the age, gender, regional accent, emotional state or speech rate of the talker. Each word can thus be materialized by an infinite number of different sound patterns that listeners have to associate with a unique lexical entry.

How do listeners deal with the inherent variability of the speech signal? The view of a mental lexicon encoding words in a format that does not retain the surface details of the inputs has been for a long-time the basic assumption (e.g. Marslen-Wilson, 1990; McClelland & Elman, 1986; Norris, 1994). As a result most models of spoken word recognition assume that variation in the speech signal is treated as noise that is stripped away before making contact with lexical representations. In this view, the listener first engages in a normalization process and thus converts the speech signal into a sequence of discrete segments, for instance phonemes, removing all acoustic details deemed irrelevant for identification. The result of this first analysis is then projected onto abstract phonological representations consisting of a string of discrete symbols, which does not include details about how these words are pronounced. In accordance with the normalization hypothesis, the pioneering studies that have examined the impact of variability on spoken word processing have shown performance costs both in terms of decrease in accuracy and increase in Reaction Times (RTs) when the speech signal is rendered highly variable by the use of multiple talkers (e.g. Creelman, 1957; Mullennix, Pisoni & Martin, 1989). For example, Mullennix et al. (1989) found that identification of

English words was both faster and more accurate when a single talker was used than when multiple talkers were used, thus suggesting that increased variability in the speech signal imposes a higher demand on the normalization process, thereby deteriorating word recognition.

However a growing body of research has emerged showing that variation in the surface form of words are in fact retained in memory and consequently influence spoken word processing (e.g. Bradlow, Nygaard & Pisoni, 1999; Palmeri, Goldinger & Pisoni, 1993; see Nygaard & Pisoni, 2008; Pufahl & Samuel, 2014 for reviews). A striking demonstration comes from studies using the long-term repetition priming paradigm (e.g. Dufour & Nguyen, 2014; McLennan & Luce, 2005; McLennan & González, 2012). These studies showed that, under some circumstances, the repetition priming effect – i.e. decrease in reaction times (RTs) when a word is encountered for the second time– is smaller when two different talkers are used between the first and second presentation. This talker specificity effect has thus strong theoretical implications since it challenges abstractionist models and in particular the assumption that words are encoded in a manner that does not retain talker-specific information. Talker-specificity effects lend support to another class of models, namely exemplar based models of spoken word recognition (Goldinger, 1998). The strongest versions of these models propose a view of the mental lexicon that is radically opposed to that postulated by abstractionist models, in that they assume that the lexicon consists of multiple episodic traces including perceptual and contextual details associated to each individual occurrence of a word (Goldinger, 1998; Hintzman, 1986, 1988), thus encoding talker-specific details. As a result, an imperfect acoustic match between the first and the second presentation of a word diminishes the repetition priming effect, because it is not the same form-based representation that is re-activated in memory during the second presentation. However a

review of all the above-mentioned studies using the repetition priming paradigm reveals that the story is more complex than it seems, in that the repetition priming effect is sometimes affected by talker variation and is sometimes insensitive to talker variation. Below we describe some of the key results that illustrate this discrepancy.

In a series of long-term repetition priming experiments, McLennan and Luce (2005) presented participants with two blocks of stimuli, the first consisting of the primes, and the second consisting of the targets. During the second block, some of the words from the first block were repeated, either with the same talker as in the first block, or with a different talker. The ease of discrimination between words and nonwords was varied in a lexical decision task, and the delay between the presentation of stimuli and the participant's response was varied in a shadowing task. In the lexical decision task, a greater repetition priming effect was observed for words repeated by the same talker than for words repeated by a different talker but only when the nonwords were wordlike. In contrast, no talker-specificity effect was observed when the nonwords were unwordlike, that is when the discrimination between words and nonwords was relatively easy. In the shadowing task, a talker-specificity effect was found when participants had to wait for a response cue to repeat words (delayed shadowing task), but not when participants had to repeat words immediately upon hearing them. According to McLennan & Luce (2005), these findings argue in favor of hybrid models of spoken word recognition which combine both abstract and detailed representations (Grossberg & Myers, 2000; Grossberg & Stone, 1986), and with an access to detailed representations occurring when processing is slow and effortful. When the nonwords used in a lexical decision task are wordlike, processing is slow which allows talker information to interact with abstract lexical information. Also, the additional processing time in the delayed shadowing task leaves time for listeners to process information related to the talker, and thus talker-specificity effects can

emerge. Compatible with this claim, subsequent studies have shown that talker-specificity effects are more likely to emerge with foreign-accented words (McLennan & González, 2012), with dysarthric speech (Mattys & Liss, 2008) or yet with words of low frequency (Dufour & Nguyen, 2014; see also Luce, Charles-Luce & McLennan, 1999 and Luce and Lyons, 1998) that take longer to process than native-accented words, healthy speech and words of high frequency.

Together the above-mentioned studies suggest that both abstract and talker-specific representations co-exist, and thus are both susceptible to influence spoken word recognition. Crucially, they suggest that when variability in the speech signal comes from differences among talkers, spoken word recognition primarily relies on abstract representations, with talker-specificity effects occurring relatively late during processing. Clear evidence for an influence of talker-specific representations on spoken word recognition has thus been found when processing is relatively slow and effortful[1]. As we have seen, a considerable amount of effort has been devoted to identify the precise circumstances under which talker-specificity effects are more likely to occur (Dufour & Nguyen, 2014; Mattys & Liss, 2008; McLennan & González, 2012; McLennan and Luce, 2005). After having achieved significant progress in this area, the challenge now is to determine the precise moment in time at which talker-specific information comes into play and interacts with abstract representations, thereby influencing spoken word recognition. This was the aim of the present study.

The short-term priming paradigm in which primes and targets are presented in close temporal succession was used. We used this kind of priming because the effects observed with this paradigm have largely been interpreted within the framework of abstractionist

models (see Norris, McQueen & Cutler, 2002, and Bowers, 2000, for visual word recognition). For example, the facilitation that occurs when spoken primes and targets are phonologically related has been interpreted as reflecting the repeated activation of the same form-based representation either at a pre-lexical or at a lexical level, or at both levels of processing (i.e. Grainger & Holcomb, 2015; McQueen & Sereno, 2005; Norris et al., 2002). Short-term priming thus appears as a promising methodological tool to examine how talker-specific information interacts with abstraction representations during spoken word recognition, and crucially to examine if activation of common abstract phonological representations can be modulated when two exemplars of the same word (e.g., the word "DRESS" pronounced by two talkers) are presented as prime and target words, thereby influencing the magnitude of the repetition priming effect. To the best of our knowledge no study has yet examined whether short-term priming effect is influenced by a talker change between the prime and the target words. Furthermore, we used a variant of short-term priming that enables presentation of prime and target stimuli in even closer temporal proximity – the dichotic priming paradigm (Dupoux, Kouider, & Mehler, 2003) - that arguably reflects even more automatic processing than when primes are presented in the clear just before targets.

As Reaction Times (RTs) provide a single measure of the various mental operations involved between word onset and participant's response, it is useful to use more fine-grained measurements to determine the precise moment at which talker-specificity effects take place. We therefore recorded Event-Related Potentials (ERPs) because the millisecond-by-millisecond resolution provided by this technique allows us to study spoken word processing as it unfolds over time. Moreover, specific ERP components have been associated with distinct stages of spoken-word recognition. For example, and particularly relevant for the present study, the P350, a positive wave peaking around 350 ms from word onset, has been

interpreted as reflecting activation of lexical word-forms (Dufour, Brunellière & Frauenfelder, 2013; Friedrich, Kotz, Friederici, & Alter, 2004; Friedrich, Kotz, Friederici, & Gunter, 2004; Friedrich, Schild, Röder, 2009), and the N400, a negative wave peaking around 400 ms from word onset, has been interpreted as reflecting activation of lexical word forms (Desroches, Newman, R. L., & Joanisse, 2009), but also as reflecting the ease with which a target word is selected from the set of activated lexical candidates (Desroches et al., 2009; Dufour et al., 2013; O'Rourke & Holcomb, 2002).

We were particularly interested in examining whether dichotic repetition priming effects vary as a function of a talker change between the primes and the targets, and if so, the precise moment during word processing at which this interaction occurs. Because it has been previously observed that talker-specific effects are more likely to emerge with low than with high frequency words (Luce et al., 1999; Luce and Lyons, 1988; Dufour & Nguyen, 2014), word frequency was manipulated. Given this prior evidence, we predicted that the influence of talker change on repetition priming effects should be more marked for low than for high frequency words. Moreover, we predicted that if talker specific representations interact with abstract representations during all the stages of lexical processing going from activation to the selection of the target word, this interaction could be observed both on the P350 and N400 components. To minimize strategic anticipations of the target words from the prime words, we used a dichotic presentation of the primes and the targets, with primes presented in the left ear at a lower intensity than the targets, and the targets presented in the right ear (see Grainger & Holcomb, 2015). To avoid contamination of ERPs by motor responses on the critical words, participants engaged in a go/no-go lexical decision task in which they were instructed to respond only when they heard a non-word in their right ear (i.e. attended ear). Note that several studies have shown that talker-specific information is stored in the right hemisphere

and abstract representations in the left hemisphere (e.g. González & McLennan, 2007; see also the right ear advantage in dichotic listening with linguistic stimuli: e.g. Kimura, 1961; Tervaniemi & Hugdahl, 2003). Hence, by presenting target words in the right ear, we promoted processing in the left hemisphere, and therefore the use of abstract representations during target word processing, in line with the central aim of our study to specify the moment at which access to abstract representations is influenced by talker-related variation. Note also that all the target words in the attended right ear were all pronounced by the same talker, and thus talkers - in one ear – were not mixed. Thus, the effects under investigation are those arising during target processing due to the presentation of a related prime word in the unattended left ear, either pronounced by the same talker as the target word or by a different talker.

Method

Participants: Forty right-handed French speakers (30 women) between 19 and 38 years old from the University of Aix-Marseille participated in the experiment. All participants reported having no neurological or hearing impairments. They were paid 15 euros for their participation and each of them gave written informed consent prior to experimentation.

Materials: One hundred and forty monosyllabic words, three to four phonemes in length, were selected from Vocolex, a lexical database for the French language (Dufour, Peereman, Pallier & Radeau, 2002). All words had their uniqueness point – the phonemic position at which a word can be reliably identified - after their last phoneme. Seventy were of

low frequency, and the remaining 70 words were of high frequency. Both sets of target words were matched on number of phonemes, uniqueness point and phonological neighborhood density. Note that phonological neighbors were calculated by summing the number of words that can be formed by adding, substituting, or deleting one phoneme in the words (Luce & Pisoni, 1998). The 140 words were used both as targets and repeated primes. For each of the target words, a control prime having no phonemes in common with the corresponding target was selected. The control primes and the target words were matched on number of phonemes, uniqueness point and phonological neighborhood density. The stimuli were recorded by a male and a female native speaker of French, in a sound-attenuated room, and digitized at a sampling rate of 44 kHz with 16-bit analog to digital recording. Each talker produced each words three times, and we exploited these different productions so that within each talker, the high and low frequency target words were matched in average duration, and the control primes were matched in average duration with the repeated primes[2]. The characteristics of the target words and the control primes are summarized in Table 1. The target words and their control primes are provided in the Appendix.

<Insert Table 1 about here>

Because each target was paired with two different primes (repeated and control), and no participant was presented with the same target twice, two experimental lists were created. Each list included the 140 target words – 70 of high frequency and 70 of low frequency. Half of the high frequency words were preceded by the repeated primes and the other half by the control primes. Also, half of the low frequency words were preceded by the repeated primes and the other half by the control primes. Within each list, the target words were always heard as produced by the male speaker. The two lists were then divided in two sub-lists so that the

prime words were heard as produced by the female speaker in one sub-list and by the male speaker in the other sub-list. Note that our experimental design differs from that used in prior behavioral studies examining talker-specificity effects and in which participants saw one experimental list composed of three types of prime, namely repeated primes by the same talker, repeated primes by different talker, and control primes. Because on the one hand, such a design requires to split our target words in three groups and because on the other hand, ERP averaging imposes a number of trials greater than 30 per condition for one participant, we used an experimental design in which the type of prime was a within-participant factor with two levels (repeated, control), and talker was a between-participant factor also with two levels (same, different). Hence, within each talker condition (same, different) each list included the 140 target words (70 of high frequency and 70 of low frequency), and within each frequency category, half of the targets words were preceded by a repeated prime and the other half by a control prime. As a result, each participant received 35 trials per condition, which is generally considered to be the minimum necessary for ERP averaging in this kind of experiment. Note also that in the lists in which the primes and the targets were pronounced by the male talker, the same tokens were used both as repeated primes and targets. For the purpose of the go/no-go lexical decision task, 18 non-words were created by changing only the last phoneme of real words and were added to the stimulus lists. Nine of the non-words were used as “go” target trials, and each of them was paired with a prime word that had no phoneme in common with the target. The other 9 non-words were used as primes and were paired with target words having no phoneme in common. These non-words used as prime stimuli in filler trials allowed us to ensure that participants only paid attention to stimuli presented in the right attended ear. According to the sub-lists, the non-word primes were pronounced either by the male or the female talker. Each sub-list thus included 158 trials in total and 9 (5.7%) were “go trials”.

Procedure: The participants were tested in a sound-attenuated booth and the primes and targets were presented via headphones. The targets were presented in the right ear at a comfortable listening level (60 dB SPL). To enhance participant's attention on the target items, the primes were presented in the left ear at a lower intensity level (54 dB). Each trial began with a fixation point. 500 ms later, the primes were presented and rapidly followed, 50 ms later, by the onset of the target word (i.e. SOA = 50 ms). The fixation point remained on the screen until the offset of the target, and participants were asked to refrain from blinking and from moving their eyes during the display of the fixation point. An inter-trial interval of 1500 ms separated the offset of targets and the beginning of the following trial. Participants were instructed to attend to the right ear and to quickly press a button only when they detected a non-word. They were not informed of the presence of non-words in the unattended left ear. The participants were tested on only one experimental sub-list and began with a block of 10 practice trials. The experiment lasted around 20 minutes and participants performed the experiment without a break.

Electroencephalogram (EEG) recording and preprocessing: The electroencephalogram was recorded from the scalp with a 64-channel BioSemi Active Two AD-box at a sampling of 2048Hz. The data were filtered online at 0.16-100Hz and referenced to the left mastoid for visualization purposes. In addition to the 64 scalp electrodes, two additional electrodes were attached below the left eye and to the right of the right eyes to monitor for vertical and horizontal eye movements, respectively. Two other electrodes were also attached over the left and the right mastoid bones. Individual electrodes were adjusted to a stable offset lower than 20 kΩ. Offline the EEG signals were down-sampled to 512 Hz, re-referenced to the average of left and right mastoids and a bandpass filter was applied (0.4 – 30Hz). Independent Component Analysis (ICA) decomposition was applied, using the EEGLAB implementation

of the Infomax algorithm, to remove artifacts generated by eye blinks and saccades (Makeig, Bell, Jung & Seijnowski, 1996). Noisy electrodes were rejected before ICA and were interpolated using cubic spline interpolation after correction by ICA. For all participants, the percentage of trials retained after data preprocessing exceeded 80% and the number of trials was kept uniform across conditions.

Data analysis: The epochs, starting 50 ms before target word onset (baseline) and ending 1000 ms after, were averaged for each condition and for each participant. To select appropriate time-windows for analyses, a two-tailed cluster-based permutation test was performed to analyse the repetition priming effect. This test is based on the clustering of adjacent spatio-temporal samples and simplifies the resolution of the Multiple Comparison Problem (Maris & Oostenveld, 2007). The permutation distribution is calculated by carrying out 2000 random partitions and then selecting those samples with a permutation *p*-value below the critical cluster alpha-level ($p \leq 0.05$). From these samples, clusters are formed based on spatio-temporal adjacency and we defined neighbouring electrodes using the triangulation algorithm implemented in the Matlab toolbox, FieldTrip (Oostenveld, Fries & Maris, 2011). Those clusters with a Monte Carlo *p*-value less than 0.025 (two-tailed test, one for positive and one for negative) were retained. By plotting the electrode clusters presenting a significant priming effect against the difference between the mean topography of repeated primes and control primes over 50ms time intervals, we were able to determine the temporal extent of this priming effect. This analysis revealed electrodes clusters presenting a significant priming effect over frontal regions with a left hemisphere bias in the 250 ms to 400 ms interval. A significant priming effect was also observed from 450 ms to 800 ms. Within this time window, electrode clusters presenting a significant effect were concentrated over central and

parietal regions bilaterally from 450 ms to 600 ms, while they were more widely distributed from 600 ms to 800 ms with an increasingly right hemisphere bias. Furthermore, a visual examination of the grand average waveforms and the global field power (GFP) of the control prime vs. repeated prime comparisons for the talker (same, different) and the frequency (low, high) conditions allowed us to identify an early time window from 100 to 200 ms. Based on theses analyses, we selected our time-windows: 100-200 ms (P200), 250-400 ms (P350), 450-600 ms (N400). Interestingly, the first two time-windows coincide with the processing of the initial phoneme/cluster and vowel of the target words, and the last time-window coincides with the processing of the last phoneme of the target words. Based on this observation, we defined a fourth time-window from 650 to 800 ms beginning just after target words offset – target word duration was around 622 ms on average–, which allowed us to analyze later effects. Data from these four selected time-windows were subjected to further statistical analysis to explore possible modulations of the repetition priming as a function of both talker change and word frequency.

An ANOVA was performed on these four time windows with frequency (low, high), and prime (repeated, control) as within-participant factors, and talker (same, different) as between-participant factor. The analysis also included the factors site (frontocentral, centroparietal, posterior) and laterality (right, left) to examine in greater detail the topographical distribution of the effects. The scalp surface was divided into six regions of six electrodes each: left frontal (AF7, AF3, F7, F5, F3, F1), right frontal (AF8, AF4, F8, F6, F4, F2), left fronto-central (FC5, FC3, FC1, C5, C3, C1), right fronto-central (FC6, FC4, FC2, C6, C4, C2,), left centro-parietal (CP5, CP3, CP1, P5, P3, P1), right centro-parietal (CP6, CP4, CP2, P6, P4, P2). The Greenhouse-Geisser correction was applied (Greenhouse & Geisser, 1959) and the corrected p values are reported below. A Bonferroni correction was

used in post-hoc comparisons. Grand-average waveforms for high frequency words at various electrodes are displayed in Figure 1, and those for low frequency words in Figure 2.

Results

Behavioral results: The percentage of correct non-word detection in the attended right ear reached 84%. The presence of non-words in the unattended left ear was detected in only 2.5 % of cases. Participants incorrectly classified word targets as nonwords on 1.27% of the word trials. The mean RT for the "go" non-word trials was 1045 ms (range: 900-1300 ms).

ERP results:

100-200 ms time-window: The first repetition priming effect was seen in this first time window. Targets following control primes showed more positive values than targets following repeated primes [$F(1,38)=13.30$, $p<.001$]. The factors site and laterality were also significant, showing more positive values on fronto-central and centro-parietal sites than on frontal sites [$F(2,76)=12.09$, $p<.001$], and more positive values on the left hemisphere in comparison to the right hemisphere [$F(1,38)= 8.16$, $p<.01$]. The main effect of frequency [$F(1,38)=0.80$, $p>.20$] and the critical interaction Prime × Frequency × Talker [$F(1,38)=1.09$, $p>.20$] were not significant.

250-400 ms time-window: In this epoch, we found a significant main effect of prime [$F(1,38)=32.94$, $p<.0001$], with targets following control primes showing more positive

values than targets following repeated primes. A main effect of word frequency [$F(1,38)=9.61$, $p<.01$] was also observed with low frequency words showing more positive values than high frequency words. The interaction between prime and laterality [$F(1,38)=5.53$, $p<.05$] as well as the interaction between prime and region [$F(2,76)=4.36$, $p<.05$] were also significant. These interactions were due to stronger priming effects on the left hemisphere in comparison to the right hemisphere, and to stronger priming effects on frontal and fronto-central sites in comparison to centro-parietal sites. The critical interaction Prime × Frequency × Talker [$F(1,38)=1.25$, $p>.20$] was again not significant.

450-600 ms time-window: In this epoch, there was a main effect of prime [$F(1,38)=35.76$, $p<.0001$], with targets following control primes showing more negative values than targets following repeated primes. Crucially, there was a three-way interaction between prime, talker and frequency [$F(1,38)=7.49$, $p<.01$]. As expected, this interaction was due to a smaller repetition priming effect when different talkers were used for the primes and the targets for low frequency words [$F(1,38)=3.95$, $p=.05$] but not for high frequency words [$F(1,38)=0.20$, $p>.20$] [3]. Subsequent comparisons with a bonferroni-correction showed that the repetition priming effect for low frequency words reached significance both in the case of a talker match ($p<.0001$) and in the case of a talker mismatch ($p<.05$) between the primes and the targets. The interaction between prime and region [$F(2,76)=48.40$, $p<.0001$] was significant, and was due to a stronger effect of prime on centro-parietal sites in comparison to fronto-central sites [$F(1,38)=13.01$, $p<.001$], as well as a stronger effect of prime on fronto-central sites in comparison to frontal sites [$F(1,38)=85.49$, $p<.0001$]. The interaction between word frequency and site was significant [$F(2,76)=4.12$, $p<.05$], and was due to a stronger effect of word frequency (high frequency words generating more negative values than low frequency words) on centro-parietal sites in comparison to frontal sites [$F(1,38)=4.49$, $p<.05$],

and on fronto-central site in comparison to frontal sites [$F(1,38)=4.79$, $p<.05$]. Finally, there was a significant main effect of site [$F(2,76)=12.63$, $p<.001$] with more negative values on centro-parietal sites in comparison to frontal and fronto-central sites. This main effect of site was also found to be greater on the left hemisphere than on the right hemisphere [$F(2,76)=8.15$, $p<.01$]. Figures 3 and 4 provide a zoom on the Cz electrode, to illustrate the modulation of the repetition priming as a function of a talker change that occurred in this time-window for low and high frequency words, respectively. Figure 5 provides an illustration at Cz of the difference in the magnitude of the priming effect (control minus repeated primes) as a function of a talker change for both low and high frequency words.

650-800 ms time-window: In this last time-window, the main effect of prime [$F(1,38)=93.70$, $p<.0001$] was again significant, with targets following control primes showing more negative values than targets following repeated primes. The main effect of frequency was no longer significant [$F(1,38)=1.04$, $p>.20$]. There was a three-way interaction between prime, talker, and frequency [$F(1,38)=4.19$, $p<.05$]. Unexpectedly however, this interaction was due to a smaller repetition priming effect when the same talker was used for the primes and the targets for high frequency words [$F(1,38)=4.37$, $p<.05$]. No modulation of the priming effect as a function of a talker change was observed for low frequency words in this last time-window [$F(1,38)=0.11$, $p>.20$]. Subsequent comparisons with a bonferroni-correction showed that the repetition priming effect for high frequency words reached significance both in the case of a talker match ($p<.0001$) and in the case of a talker mismatch ($p<.01$) between the primes and the targets. The interaction between prime and laterality [$F(1,38)=11.51$, $p<.01$] as well as the interaction between prime and region [$F(2,76)=14.51$, $p<.001$] were also significant. These interactions were due to stronger priming effect on the right hemisphere in comparison to the left hemisphere, and to a stronger effect of prime on

both centro-parietal sites and fronto-central sites in comparison to frontal sites. Finally, there was a significant main effect of site [$F(2,76)=80.51$, $p<.001$] with more negative values on centro-parietal sites in comparison to frontal and fronto-central sites, and more negative values on fronto-central sites compared to frontal sites.

<Insert Figures 1, 2, 3, 4 & 5 about here>

General Discussion

This ERP dichotic repetition priming study was conducted with three main objectives. First, we were interested in short-term repetition priming effects, and asked whether these effects are influenced by talker-specific information. Second, and in the case of a positive answer to this first question, we were interested in determining the exact moment in the processing of spoken words at which talker-specific influences occur, and affect the magnitude of the repetition priming effect. Third, because behavioral studies using long-term repetition priming paradigm have reported stronger talker specificity effects for low than for high frequency words (Luce et al., 1999; Luce and Lyons, 1988; Dufour & Nguyen, 2014), we examined in more details the interaction between word frequency and talker specificity effects, and more specifically the precise moment during spoken word recognition that this interaction takes place.

The earliest ERP differences between control and repeated primes were found between 100 and 200 ms from target word onset. More precisely, the first repetition priming effect was seen on a positive-going waveform, a P200-like component, with repeated primes eliciting less positive-going waveforms than control primes. Because this first repetition effect roughly occurs during the processing of the first phoneme of the targets, it likely reflects a pre-lexical stage of processing (see Grainger & Holcomb, 2015, for an early repetition priming effect in a similar time-window[4]). Repetition priming effects presumably occur because the processing of the target words involves the use of pre-lexical representations that were already activated during prime processing. Importantly, this first repetition priming effect was not sensitive to a talker change between the primes and the targets. Hence, it appears that talker-specific

information does not affect the pre-lexical stage of spoken words processing. This key finding is in line with Grainger and Holcomb's (2015) proposal that the earliest repetition priming effects obtained in dichotic repetition priming are subtended by source-invariant sublexical phonological representations.

A second repetition priming effect was observed between 250-400 ms from word onset, and more particularly on the P350 component, with repeated primes again eliciting less positive-going waveforms than control primes. Crucially, this repetition effect was accompanied by a word frequency effect, with high frequency words eliciting less positive-going waveforms than low frequency words. Because the word frequency effect is a clear signature of lexical access, we believe that this repetition priming effect reflects a lexical stage of processing. The effect presumably occurs because the processing of target words involves the use of lexical representations that were already activated during prime processing. Again, at this level of processing no modulation of the repetition priming effect as a function of a talker change between the primes and the targets was observed. Interestingly, the P350 has been interpreted as reflecting activation of phonological word forms (Dufour et al., 2013; Friedrich et al., 2004a; 2004b; Friedrich et al., 2009), and both the repetition priming effect and the word frequency effect have an important characteristic that corroborates such a claim. They roughly emerge during the processing of the first two phonemes of the target words, that is the point in time at which lexical activation is assumed to start (Marslen-Wilson & Welsh, 1978). Considering that we have tapped into activation of lexical representations, the absence of a talker specificity effect within this second time-window suggests that talker-specific information does not affect the earliest moments of lexical access.

A third repetition priming effect was observed between 450 and 600 ms from word onset, and more particularly on the N400 component, with repeated primes again eliciting lower amplitudes than control primes. Again, this repetition effect occurs accompanied by a word frequency effect, and again, it likely reflects the engagement of lexical representations. Crucially, talker specific information was found to have its first and only influence in this time-window. In accordance with behavioral studies (Luce et al., 1999; Luce and Lyons, 1988; Dufour & Nguyen, 2014), we reported a modulation of the repetition priming effect as a function of a talker change between the primes and the targets only for low frequency words. Because, the effects reported in this time-window approximately occurred during processing of the last phoneme and extended to the end of the target words, they likely reflect the moment at which the target word is selected as the best candidate and is recognized. As a consequence, a match in talker identity between the prime and the target would have facilitated the selection of low frequency target words, and a greater repetition priming effect is observed in case of same talker lexical items.

A repetition priming effect was also observed just after target word offset in a time-window between 650 and 800ms, and was significant for both low and high frequency words, and for both a talker match and a talker mismatch. Curiously, however, the effect was found to be stronger in the case of a talker mismatch for high frequency words. Because this effect occurs after target word offset and likely after lexical access is completed, we think that it reflects processes other than those involved in lexical identification. Although response biases in terms of strategic anticipations of target words based on the repeated phonemes of the primes and targets are unlikely in dichotic priming with primes and targets overlapping temporally, a possibility however is that once words are identified, participants engage in a comparison process between the prime words presented in one ear and the target words

presented in the other ear. Although RT measurements combined with ERPs are necessary to test for this claim, this repetition priming effect could reflect a congruency checking process, a process known – in classic sequential priming – to facilitate "word" decisions in a standard lexical decision task when there is a phonological congruency between the primes and the targets (see Norris et al., 2002). The observation that the repetition effect is greater when two different talkers are used, at least for high frequency words, could also be explained by the fact that this congruency checking mechanism is facilitated when the primes and the targets are pronounced by different talkers and are therefore more distinct. This advantage was not observed for low frequency words probably because in the preceding N400 time-window, the repetition priming effect was found to be smaller for low frequency words when there was a talker mismatch.

The fact that the N400 component reflects selection of the target word during spoken word recognition has been proposed in other studies (Desroches et al., 2009; Dufour et al., 2013; Dumay, Benraïss, Barriol, Colin, Radeau & Besson, 2001; O'Rourke & Holcomb, 2002). For example, testing the proposal of the Cohort model (Marslen-Wilson & Welsch, 1978) that words are recognized when the information in the speech signal is compatible with no other words than the target, O'Rourke & Holcomb (2002) have reported that the N400 component peaked sooner for words with early than late recognition points when ERPs were time-locked to word onset. Another demonstration comes from Desroches et al. (2009)'s study that examined the influence of lexical candidates that match either the initial or the final phonemes of the target word on the recognition of the target word. To this end, they used a picture-word matching task in which participants saw a picture followed by an auditory probe word, and had to judge whether the picture and the probe were the same. The probe could be identical to the picture label (e.g., CONE–cone); could share the rime with the picture label

(e.g., CONE–bone); could share the initial phonemes with the picture label (e.g., CONE–comb); or was unrelated to the picture label (e.g., CONE -fox). Desroches et al. (2009) observed a reduction in the amplitude of the N400 component in the rime overlap condition (CONE–bone). For the initial overlap condition (CONE–comb), the magnitude of the N400 component increased significantly in comparison to the other conditions at a slightly later time point (late N400; 410-600 ms). The increase in the late N400 amplitude in the initial overlap condition has been taken as evidence in favor of a competition process between words overlapping in their initial phonemes during spoken word recognition. In particular, the presence of a close competitor makes it harder to select the target word from among the set of activated candidates. Hence, our interpretation of the N400 component fits well with that of other studies in which the N400 has been interpreted as reflecting lexical selection.

In the present study, prime and target words produced by the male talker had a mean duration of 622 ms, whereas the prime words produced by the female talker had a mean duration of 539 ms. This difference in prime duration was therefore confounded with our same vs. different talker factor, and could therefore be the cause of the observed talker-specificity effect. We think that this is unlikely for two reasons. First, because these timing differences were the same for both high and low frequency words, any talker specificity effect resulting only from such differences should not modulate as function of target word frequency. Second, one might expect effects driven by differences in stimulus duration to affect relatively early ERP components, whereas the talker-specificity effect we observed occurred only on the N400 component and not before. Nevertheless, although it remains unclear how prime duration in itself could have caused the talker specificity effect reported in the present study, we certainly agree that future research should examine how prime duration might impact on the magnitude of repetition priming effects in the dichotic priming paradigm.

Taken together, we believe that our results add to the growing body of evidence showing that under some circumstances talker-specific information influences spoken word recognition (e.g. Creel, Aslin and Tanenhaus, 2008; Dufour & Nguyen, 2014; McLennan & Luce, 2005; McLennan & González, 2012). In accordance with behavioral studies, we reported that talker-specificity effect is more likely to occur during lexical identification for low frequency than for high frequency words. The observation that talker-specificity effect is limited to low frequency words and that repetition priming effect still occurs – although diminished – in the case of a talker mismatch on low frequency words, suggest that abstract representations are primarily activated, and that this is this type of representation that dominates spoken word recognition. Note that exemplar-based model would also predict smaller specificity effects for high frequency than for low frequency words. In the case of high frequency words, the large number of exemplars that are activated would cancel out individual differences, and this would make it less likely to observe talker-specificity effect for these words. Such an explanation has been envisaged by Goldinger (1998) to account for the influence of word frequency in the access to detailed information in a production task, and in particular the observation that low frequency words are imitated to a greater extent than high frequency ones. Nonetheless, as noted by Goldinger (1998) and Luce and Lyons (1998), the simultaneous activation of many traces for a high frequency word may both obscure the acoustic details associated with each particular trace, and cause the generation of a more abstract, generic, representation for the target word. Thus, even within strict episodic model, abstract representations nonetheless exist and come into play in spoken word recognition. This kind of model thus approaches hybrid models in which representations are initially episodic, and become more abstract with repeated exposure to the words (Grossberg, 1986; McLennan & Luce, 2005).

Our results could be in accordance with a view of a mental lexicon incorporating both abstract and talker-specific representations (McLennan & Luce, 2005). The difficulty however with such an approach resides in our observation that talker specific information influences only a late stage of spoken word recognition, namely lexical selection. No evidence for an influence of talker specific information was observed during lexical activation, and it is rather hard to envisage, within a hybrid view of the mental lexicon, that talker-specific representations does not come into play at the stage of lexical activation. Recent studies (Hanique, Aalders & Ernestus, 2013; Nijveld, Mulder, Bosch & Ernestus, 2016) question hybrid models of spoken word recognition and argue that talker-specificity effects have little to do with the mental lexicon, instead they would engage detailed representations stored in episodic memory. Our observation that the talker-specific effect is modulated by word frequency, an effect that indexes some form of lexical activation, seems to indicate that our effects are driven at least in part by the lexical representations. A possibility to account for our findings is to envisage some interactions between the abstract representations stored in the mental lexicon and the detailed representations of words stored in episodic memory. This interaction would take place during the selection process, and "episodic" representations would act in boosting the selection of target words that require either a strong accumulation of perceptual evidence to reach the recognition threshold, as would be the case for low frequency words[5], or when the speech input is degraded and does not fully match abstract representations, as in the case with foreign-accented words (McLennan & González, 2012) or with dysarthric speech (Mattys & Liss, 2008).

In conclusion, our study is in accordance with prior behavioral research showing that word frequency modulates effects of talker-specificity on spoken word recognition. In particular, talker-specific information influences the identification of low frequency words but

not that of high frequency words. Crucially, our study shed light on the exact moment at which talker-specific influences arise during spoken word recognition. We showed that talker-specific influences occur late during spoken word recognition, and most likely only start to have an impact during lexical selection. This late influence of talker specific information is all the more interesting in that it allows us to reject an alternative explanation in terms of an increase in "processing effort" in order to deal with the variability engendered by a talker change between the primes and the targets (e.g. Creelman, 1957; Mullennix et al., 1989). Indeed, if our results were mainly due to differences in processing effort, then the talker-specificity effect should have been observed relatively early during the time-course of spoken word recognition, and well before the N400 component. Also, this late influence of talker-specific information suggests that our effects were not merely due to a greater acoustic distance between the primes and the targets in the different talker condition, since in such a case a modulation of the repetition priming effect as a function of a talker change should have also occurred during pre-lexical processing, and therefore should have been visible in earlier ERP components. Nevertheless, future research could examine the possible role of additional factors on talker-specificity effects by comparing, for example, lists blocked by talker, as in the present study, with lists in which the same and different talker conditions are intermixed.

On the basis of the present findings, we argue that spoken word processing relies primarily on abstract representations. The challenge now for future work is to better understand the origin of the influences related to talker-specific information. Taking into account recent studies (Hanique et al., 2013; Nijveld et al., 2016), we suggest that talker-specific effects might result from the interactions between two memory systems, namely episodic memory and the mental lexicon. Given that hemispheric differences in the processing of talker-specific information have been revealed in prior research (e.g. González &

McLennan, 2007), it would be also interesting to continue the present work by comparing the time-course of the talker-specific effect as a function of the ear of presentation of the target words. This and other future work should further reveal the utility of combining the dichotic priming paradigm with ERP recordings as a tool for uncovering basic mechanisms underlying speech perception and spoken language comprehension.

Footnote

1: Activation of abstract and talker-specific representations is susceptible to follow a different time course in case of allophonic variability, which results from articulatory and acoustic differences among speech sounds that belong to the same phonemic category. In a series of experiments focusing on flaps – neutralized allophones of intervocalic /t/ and /d/ phonemes in American English –, McLennan and Luce (2003) have shown that the specific allophonic representation dominates processing when responses were rapid, while evidence for an access to the abstract underlying representation (/t/and /d/) was obtained when responses were slowed. McLennan and Luce (2003, 2005) have proposed an account of the time-courses of allophonic and talker variability within the ART (Adaptive Resonance Theory) framework (Grossberg, 1986) and in which the speed with which access to the one or the other type of representations is dependent on their frequency. The flap being more frequent in intervocalic contexts than the underlying /t/ and /d/ phonemes, it is primarily activated. In contrast, in the case of talker variability, and in particular in the case of unfamiliar talkers (Maibauer, Markis, Newell & McLennan, 2014), the specific features related to each of the talker being less frequent than the abstract features found in the words, it is the abstract representation that is first activated and that dominates processing.

2: Talker gender was selected as the talker-specific information to be manipulated in the present study since it has been widely used in previous studies examining talker specificity effects in spoken word recognition (e.g. Creel et al., 2008; Maibauer et al., 2014; McLennan & Luce, 2005). Note that our talkers also differed in their speaking rate since on average the acoustic duration of words recorded by the male talker was longer than that of the words recorded by the female talker. Hence any talker specificity effects could be driven by either the gender of the talkers, the speaking rate of the talkers or both. This study specifically

addresses when in processing of words talker-specificity effect takes place, but was not designed to examine the specific variables that drive talker-specificity effect. This difference in the speaking rate of the male and female talker necessarily caused a difference in prime duration as a function of the talker used. This point will be discussed in the General Discussion.

3: Note that the three-way interaction between prime type (control, repeated), frequency (low, high) and talker (same, different) indicates how the magnitude of the priming effect (the difference between control and repeated primes) is modulated by both word frequency and talker change. Obviously, additional analyses performed directly on the magnitude of the difference between control and repeated primes with frequency (low, high), as within-participant factor, and talker (same, different) as between-participant factor led to exactly the same patterns of results, and with exactly the same statistical values. This is of course true for the four time-windows of interest. As an illustration, in this particular time-window (*450-600 ms),* there was a significant two-way interaction between talker and frequency [$F(1,38)=7.49$, $p<.01$]. Again, this interaction is due to a significant difference in the magnitude of the difference between the same and different talker condition for low [$F(1,38)=3.95$, $p=.05$] but not for high frequency [$F(1,38)=0.20$, $p>.20$] words.

4: Note that in Grainger and Holcomb (2015), the first repetition effect was seen on a negative-going component, with targets following control primes producing greater negativities than targets following repeated primes. Because in their ERPs analyses, Grainger and Holcomb (2015) used only the left mastoid bone as reference, we re-ran our analyses with this reference only, and found exactly the same pattern of results. The difference in the polarity of the first repetition priming effect between the two studies could be due to the use of different tasks (go/no-go semantic categorization task in Grainger and Holcomb (2015)).

5: In models with discrete lexical representations, such as in interactive-activation models (e.g., McClelland & Elman, 1986) or in the Cohort Model (Marslen-Wilson, 1990), frequency is generally coded in the resting activation level of lexical units and thus determines the baseline activation level of each word. As high frequency words have higher resting activation levels than low frequency ones, they reach the recognition threshold earlier and thus are selected faster as the best candidate.

References

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. Perception & Psychophysics, 61, 206 –219.

Creel, S. C., Aslin, R. N., Tanenhaus, M. K. (2008). Heeding the voice of experience: The role of talker variation in lexical access. *Cognition* 106, 633–664.

Creelman, C. D. (1957). The case of the unknown talker. Journal of the Acoustical Society of America, 29, 655.

Desroches, A. S., Newman, R. L., & Joanisse, M. F. (2009). Investigating the time course of spoken word recognition: Electrophysiological evidence for the influences of phonologicalsimilarity. *Journal of Cognitive Neuroscience, 21,* 1893–1906.

Dufour, S., Brunellière, A. & Frauenfelder, U.H. (2013): Tracking the time course of word frequency effects in auditory word recognition with event-related potentials. *Cognitive Science, 34,* 489-507.

Dufour, S., Nguyen, N. (2014). Access to talker-specific representations is dependent on word frequency. *Journal of Cognitive Psychology,* 26, 256-262.

Dufour, S., Peereman, R., Pallier, C., & Radeau, M. (2002). VoCoLex: A lexical database on phonological similarity between French words. *L'Année Psychologique*, *102*, 725–746.

Dumay, N., Benraïss, A., Barriol, B., Colin, C., Radeau, M., & Besson, M. (2001). Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *Journal of Cognitive Neuroscience*, *13*, 121-143.

Dupoux, E., Kouider, S. & Mehler, J. (2003). Lexical access without attention? Explorations using dichotic priming *Journal of Experimental Psychology-human Perception and Performance*, 29, 172-184.

Friedrich, C. K., Kotz, S. A., Friederici, A. D., & Alter, K. (2004a). Pitch modulates lexical identification in spoken word recognition: ERP and behavioral evidence. *Journal of Cognitive Brain Research*, *20*, 300–308.

Friedrich, C. K., Kotz, S. A., Friederici, A. D., & Gunter, T. C. (2004b). ERPs reflect lexical identification in word fragment priming. *Journal of Cognitive Neuroscience, 16,* 541–552.

Friedrich, C.K., Schild, U., & Röder, B. (2009). Electrophysiological indices of word fragment priming allow characterizing neural stages of speech recognition. *Biological Psychology, 80,* 105-113.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105, 251–279.

González, J., & McLennan, C. T. (2007). Hemispheric differences in indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *33*,410–424.

Grainger, J. & Holcomb, P. J. (2015). An ERP investigation of dichotic repetition priming with temporally overlapping stimuli. *Psychonomic Bulletin and Review, 22*, 289-296.

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

Grossberg, S. (1986). The adaptive self-organization of serial order in behavior: Speech, language, and motor control. In E. C. Schwab & H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines: Vol. 1. Speech perception* (pp. 187–294). New York: Academic Press.

Grossberg, S., & Myers, C. W. (2000). The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review 107, 4,* 735–767.

Grossberg, S., & Stone, G. (1986). Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review 93,* 46–74.

Hanique, I., Aalders, E., & Ernestus, M. (2013). How robust are exemplar effects?. *The Mental Lexicon* 8, 269-294.

Hintzman, D. L. (1986). “Schema abstraction” in a multiple-trace memory model. *Psychological Review*, *93*, 411 – 428.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*, 528 – 551.

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, *15*, 166 –171.

Luce, P. A., Charles-Luce, J., & McLennan, C. T. (1999, August). Representational specificity of lexical form in the production and perception of spoken words. *Paper presented at the proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, CA.

Luce, P. A., & Lyons, E. A. (1998). Specificity of memory representations for spoken words. *Memory & Cognition, 26*, 708–715.

Luce, P. A. and Pisoni, D. B. (1998). Recognizing spoken words: The Neighborhood Activation Model. *Ear and Hearing*, *19*, 1–36.

Maibauer, A. M., Markis, T. A., Newell, J. & McLennan, C. T. (2014). Famous talker effects in spoken word recognition. *Attention, Perception, and Psychophysics*, *76*, 11-18.

Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. (1996). Independent component analysis of electroencephalographic data. In: D. S. Touretzky, M. C. Mozer, & M.E. Hasselmo (Eds.), *Advances in neural information processing systems* (Cambridge: MA: MIT Press, pp. 145-151.

Maris, E., Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*, 177–190.

Marslen-Wilson, W.D. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (PP. 148-172). Cambridge, MA: MIT Press.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interaction and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10*, 29-63.

Mattys, S. L., & Liss, J. M. (2008). On building models of spoken word recognition: When there is as much to learn from natural “oddities” as artificial normality. *Perception & Psychophysics, 70*, 1235–1242.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1– 86.

McLennan, C.T., & González, J. (2012). Examining talker effects in the perception of native- and foreign-accented speech. *Attention, Perception & Psychophysics, 74*, 824-830.

McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. Journal of Experimental Psychology: Learning, Memory, and Cognition, 29, 539 – 553.

McLennan, C. T., Luce, P. A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 31, 306–321.

McQueen, J. M. & Sereno, J. (2005). Cleaving automatic processes from strategic biases in phonological priming. *Memory and Cognition, 33*, 1185 - 1209.

Mullennix, J. W., Pisoni,D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. Journal of the Acoustical Society of America, 85, 365-378.

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition*, *52*, 189– 234.

Norris, D., McQueen, J.M, & Cutler, A. (2002). Bias effects in facilitatory phonological priming. *Memory and Cognition*, *30*, 399-411.

Nijveld, A., Mulder, K., Bosch, L. T. & Ernestus, M. (2016). ERPs reveal that exemplar effects are driven by episodic memory instead of the mental lexicon. *Paper presented at LabPhon 15*, Cornell University, New-York.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics, 60*, 355–376.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). Field-Trip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Computational Intelligence and Neuroscience, 156869.

O'Rourke, T.B., Holcomb, P.J. (2002). Electrophysiological evidence for the efficiency of spoken word processing. *Biological Psychology, 60*, 121–150.

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 309-328.

Pufahl, A., & Samuel, A. G. (2014). How lexical is the lexicon? Evidence for integrated auditory memory representations. *Cognitive Psychology, 70*, 1–30.

Tervaniemi, M. & Hugdahl, K. (2003). Lateralization of auditory-cortex functions. Brain Research Reviews, 43, 231–246.

Table 1: Characteristics of the stimulus sets (mean values)

	Low Frequency		**High Frequency**	
	Target Words	**Control Primes**	**Target Words**	**Control Primes**
Frequency (in occurrence per million)	5.67	5.24	139.84	145.43
Number of phonemes	3.11	3.11	3.11	3.11
Uniqueness point	4.11	4.09	4.10	4.06
Neighborhood density	24.70	25.67	24.67	26.47
Female speaker duration (in ms)	539	539	539	539
Male speaker duration (in ms)	622	622	622	622

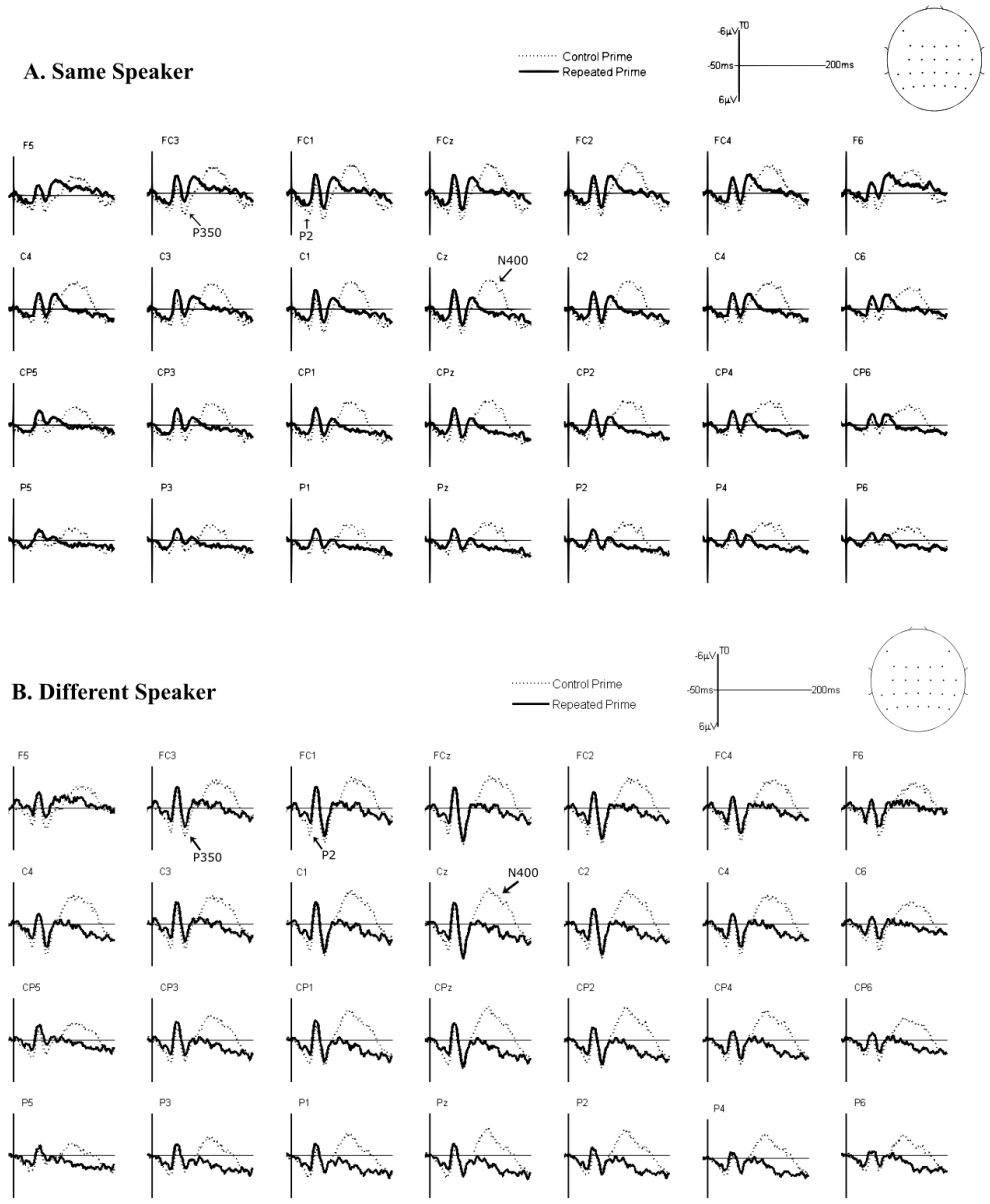


Figure 1 : Grand-average waveforms at various recordings sites for high frequency words in a time window between -50 and 1000 ms after the onset of target words. In this and following figures, amplitude (**µ**V) is represented on the Y-axis with negative voltage up, and time (ms) on the X-axis.

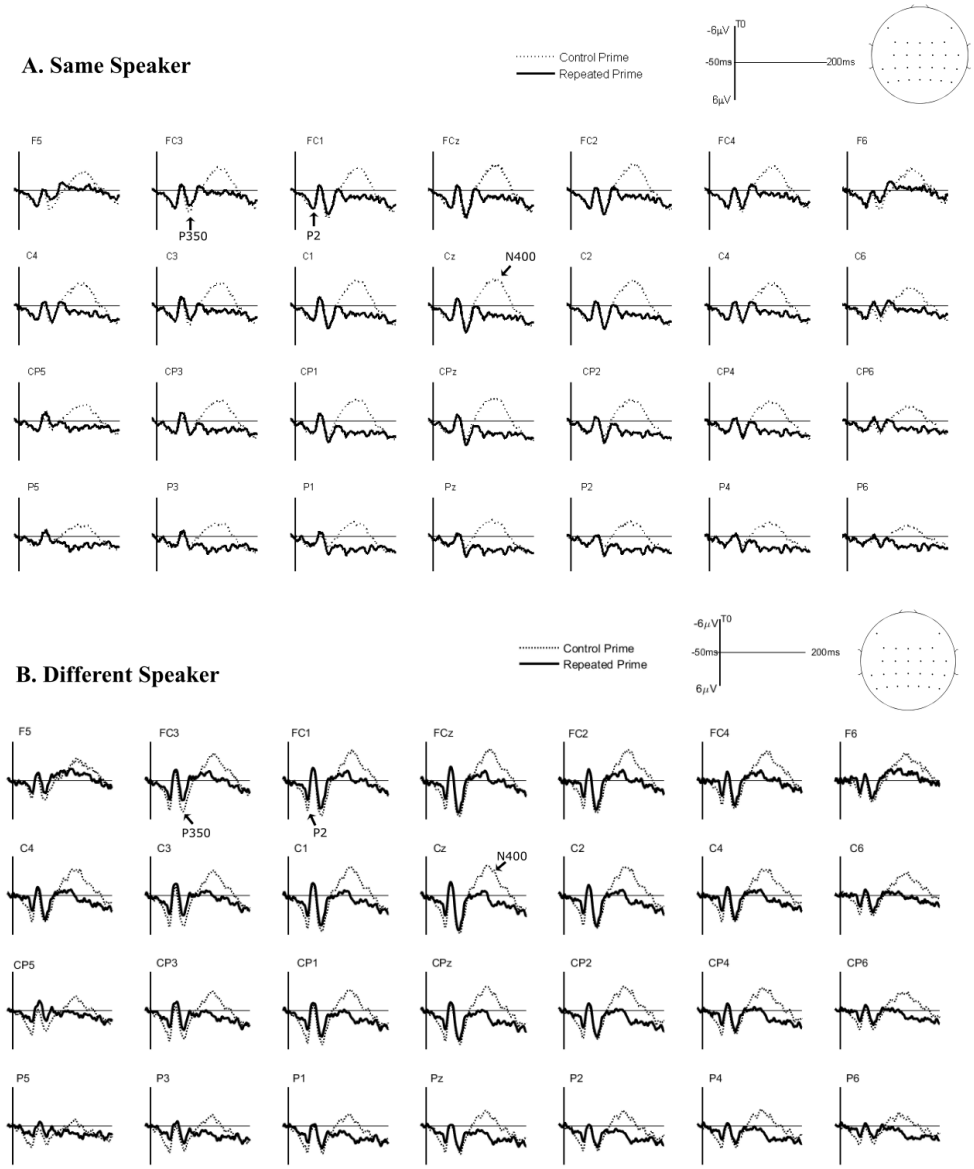
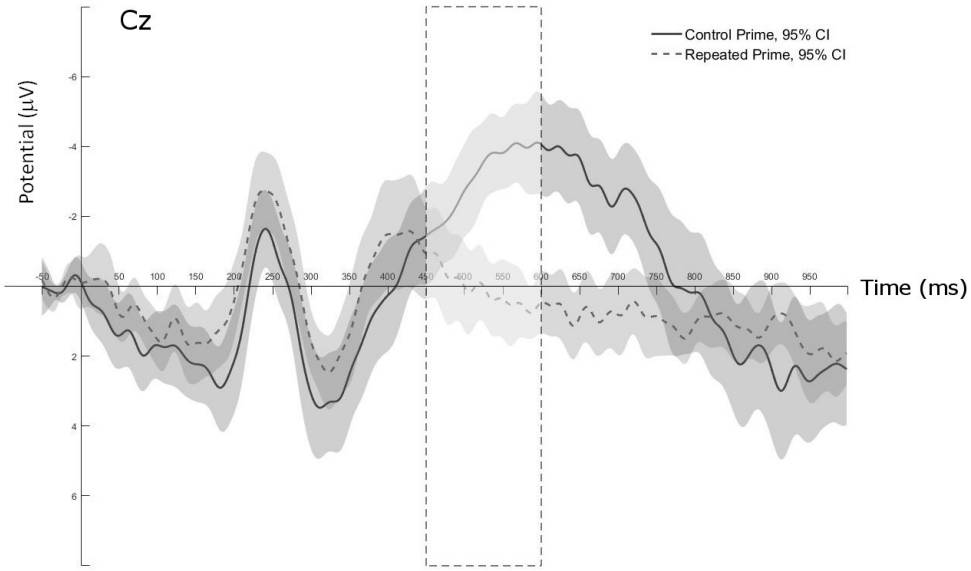


Figure 2 : Grand-average waveforms at various recordings sites for low frequency words in a time window between -50 and 1000 ms after the onset of target words.

A. Same Speaker



B. Different Speaker

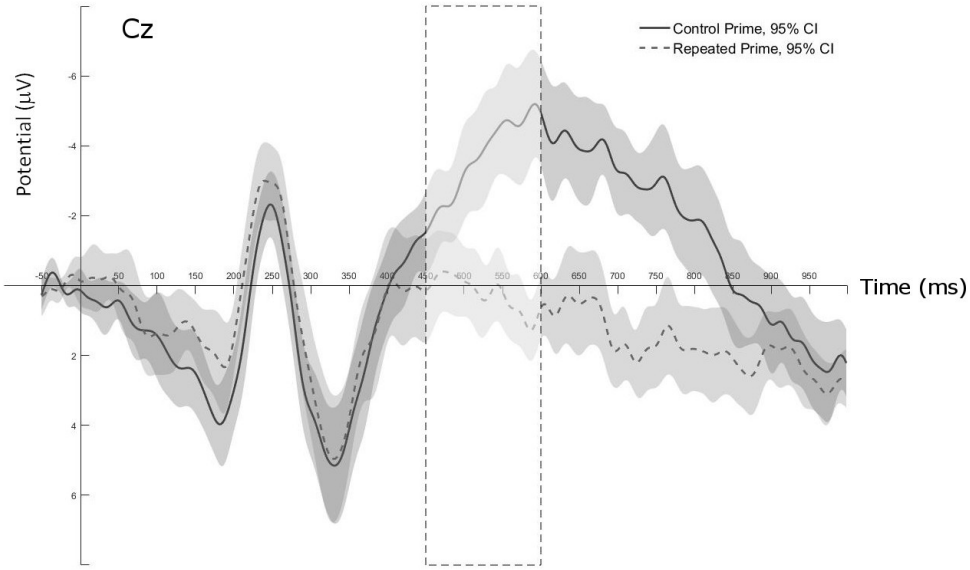
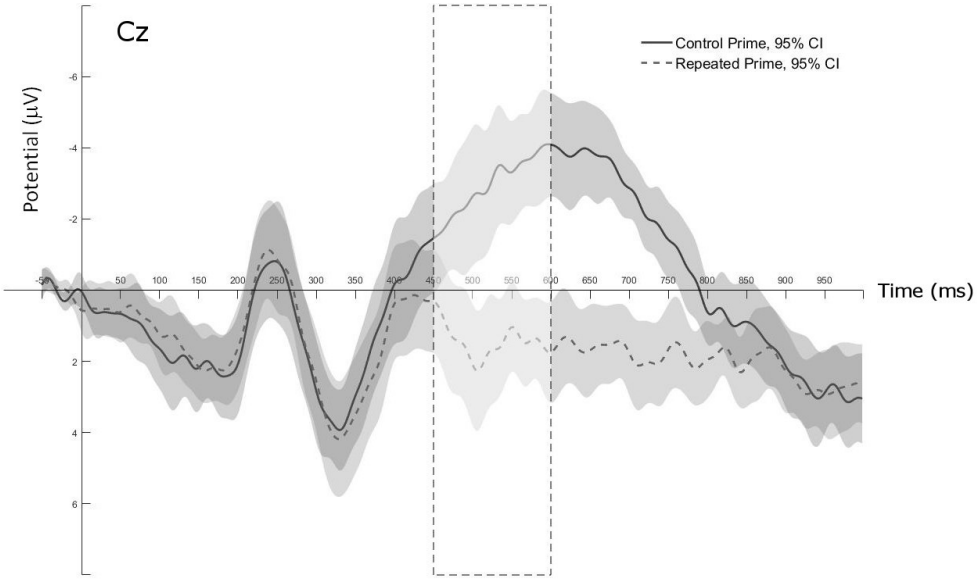


Figure 3 : Grand-average waveforms at Cz for high frequency words in a time window between -50 and 1000 ms after the onset of target words. The dotted rectangular box illustrates the repetition priming effect found in the 450-600 ms time window, during processing of the last phoneme of the target words.

A. Same Speaker



B. Different Speaker

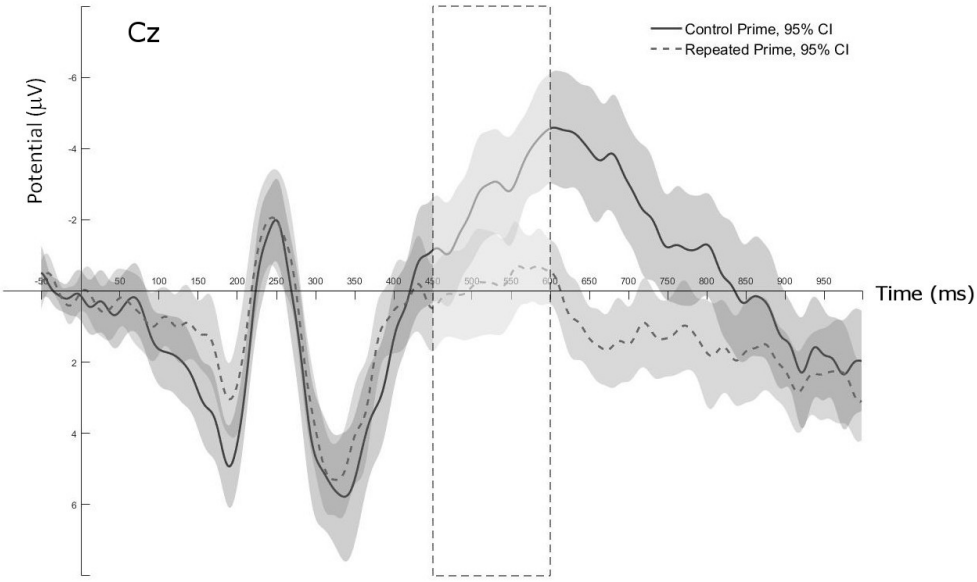
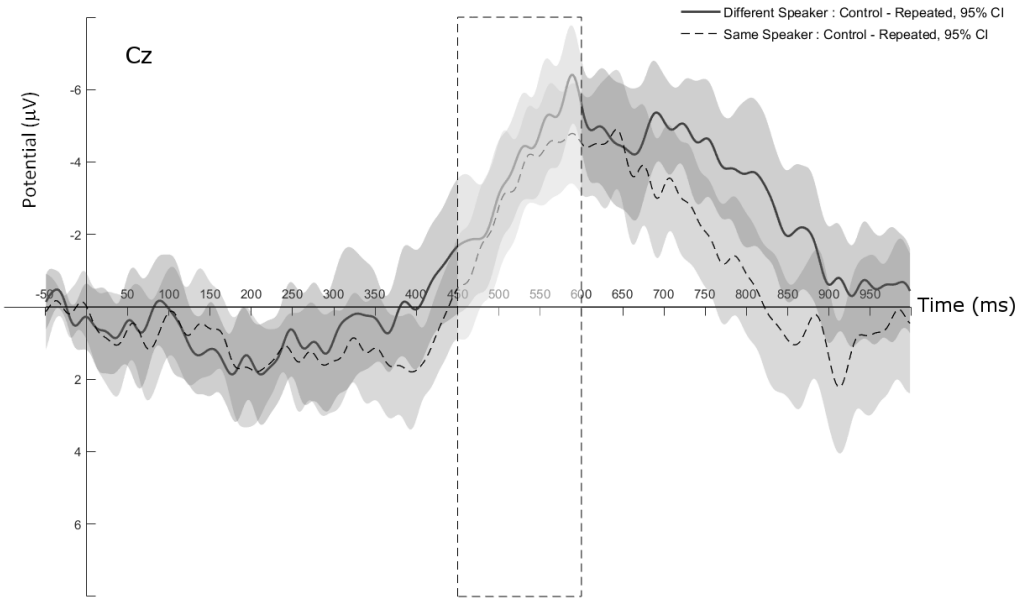


Figure 4 : Grand-average waveforms at Cz for low frequency words in a time window between -50 and 1000 ms after the onset of target words. The dotted rectangular box illustrates the repetition priming effect found in the 450-600 ms time window, during processing of the last phoneme of the target words.

A. High Frequency Words



B. Low Frequency Words

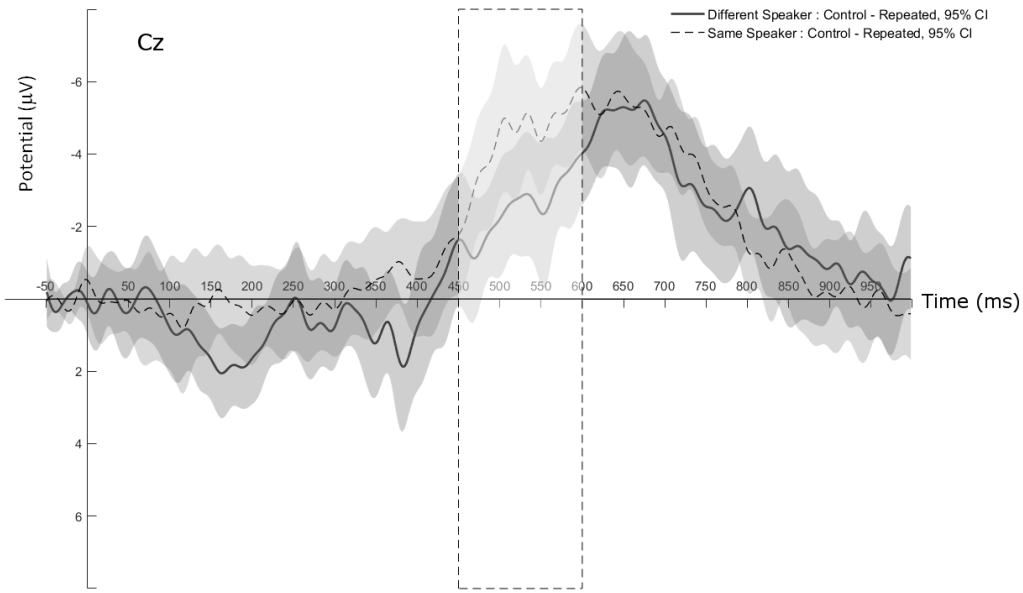


Figure 5: Grand average difference waveforms (Control minus repeated primes) at Cz as a function of a talker change for both high frequency words and low frequency words. The dotted rectangular box illustrates the repetition priming effect found in the 450-600 ms time window, during processing of the last phoneme of the target words.

Appendix: Target words and control primes used in the Experiment

Low Frequency		High Frequency	
Target words	**Control Primes**	**Target words**	**Control Primes**
mangue	baffe	ligne	genre
quiche	lange	signe	bord
cèpe	buse	soeur	vide
ronce	bûche	langue	mise
dinde	bagne	nord	laisse
galle	phoque	rouge	mille
luge	bave	rire	sol
naine	dard	bouche	mer
nonne	soute	route	type
ruche	veuf	chef	mur
guêpe	rime	peur	face
cil	dune	fils	tour
figue	souche	bonne	mal
lime	quinte	fille	terre
rhume	sonde	ville	jeune
louve	gaine	sens	cœur
quille	gente	doute	guerre
natte	soude	femme	verre
puce	nage	seul	corps
ride	gamme	vite	père
fugue	ponte	monde	cours
rate	fève	tête	passe
cuve	latte	jour	peine
chèque	ruse	dire	salle/sale
beige	douille	chaque	fort
loupe	gage	page	fête
niche	châle	lune	sainte
bulle	chatte	vague	lutte
bise	mule	bête	nul
botte	phare	dur	caisse
gomme	bouc	poche	tombe
fade	pull	mode	conte
cale	digue	robe	manque
dupe	gel	date	juge
pelle	lard	foule	veille
peigne	louche	dame	quinze
cube	rame	somme	lire
char	site	sac	file
bus	gîte	sud	chaire
laine	biche	rêve	neuf

panne	bosse	chance	port
case	touffe	taille	douze
moule	rampe	lampe	masse
pince	rude	sec	base
veine	mite	riche	danse
bac	fente	chasse	pire
bague	tempe	note	russe
gaffe	dette	gaz	lourd
cure	bol	neige	chaîne
douche	noce	gare	coupe
mèche	dalle	chaise	douce
singe	fouille	bande	pur
poule	nerf	tâche	banque
tige	cape	code	vif
vigne	sotte	choc	rare
cire	coq	vol	mine
fiche	loge	feuille	change
mouche	pique	jambe	pêche
nappe	voûte	seuil	couche
roche	pompe	chute	balle/bal
mousse	rail	touche	reine
tasse	bille	folle	mince
crampe	blatte	classe	trente
graine	flaque	groupe	blanche
crasse	blouse	place	grande
globe	truffe	branche	style
trousse	planque	fleur	trace
grappe	cloque	frère	stade
griffe	blague	plage	grise
tranche	flash	glace	preuve