Tracking the Time Course of Competition During Word Production: Evidence for a Post-Retrieval Mechanism of Conflict Resolution

Niels Janssen[1], Juan A. Hernández-Cabrera[1,2], Maartje van der Meij[1] and Horacio A. Barber[1,2]

[1]Universidad de La Laguna, 38205 La Laguna, Spain and [2]Basque Center on Cognition, Brain and Language, 20009 Donostia, Spain

Address correspondence to Dr Niels Janssen, Cognitive Neuroscience and Psycholinguistics lab, University of La Laguna, La Laguna 38205, Spain. Email: njanssen@ull.es

Producing a word is often complicated by the fact that there are other words that share meaning with the intended word. The competition between words that arises in such a situation is a well-known phenomenon in the word production literature. An ongoing debate in a number of research domains has concerned the question of how competition between words is resolved. Here, we contributed to the debate by presenting evidence that indicates that resolving competition during word production involves a postretrieval mechanism of conflict resolution. Specifically, we tracked the time course of competition during word production using electroencephalography. In the experiment, participants named pictures in contexts that varied in the strength of competition. The electrophysiological data show that competition is associated with a late, frontally distributed component that arises between 500 and 750 ms after picture presentation. These data are interpreted in terms of a model of word production that relies on a mechanism of cognitive control.

Keywords: competition, EEG, language production

Introduction

The production of words is one of the most remarkable aspects of the human ability to produce speech. Producing a word is often influenced by competition from other words that are similar in meaning. The occurrence of speech errors such as, for example, saying "dog" when the intended word was "cat," have been taken by researchers as evidence for the existence of competition between words (e.g., Dell 1986). (There exists different uses of the term "competition" in the language production literature. In one use of the term, competition reflects the general situation where more than a single word is considered during word production. In another more restricted use, competition refers to specific assumptions about a mechanism of word selection (e.g., Roelofs 1992). We will use the word competition in its general sense. See also Janssen (2013) for further discussion.) The presence of competition between words raises important questions about how speakers are able to maintain fluent and error-free speech. Researchers have proposed that the efficient production of speech depends in part on a cognitive mechanism whose job it is to resolve competition between words. However, although researchers in general agree about the existence of such a mechanism, the exact details about how competition is resolved during word production remains poorly understood. Here, we show that electrophysiological evidence clarifies how the resolution of competition during word production is achieved.

Current cognitive models of language production generally assume that the production of a word proceeds through at least three representational stages (e.g., Caramazza 1997). First, word production starts at a semantic stage where a to-be communicated message is constructed. This stage is followed by a word retrieval stage where the corresponding words are retrieved, and by a phonological encoding stage where the word's motor program is built (e.g., Schriefers et al. 1990). Within this framework, models assume that the competition between words that arises during word production is resolved at the word production stage. How such competition is resolved is currently a matter of intense debate (see, e.g., Spalek et al. 2013, for a recent review), where some models assume that competition is resolved by evaluating the relative activation levels of words (e.g., Roelofs 1992; Abdel Rahman et al. 2009), and other models assume that competition is resolved by adjusting connection weights between semantic and word representations (Oppenheim et al. 2010). Here, we did not focus specifically on the nature of the mechanism by which competition is resolved. Rather, our main focus was on the point in time "when" competition is resolved during word production. Given the assumption in current cognitive models of language production that competition is resolved during the retrieval of words from memory, one might infer that the mechanism of competition resolution is associated with a relatively 'early' stage of word production.

A different view on the time course of competition comes from models formulated within the tradition of research on cognitive control (e.g., Botvinick et al. 2001; Yeung et al. 2004; Badre 2008; Wiecki and Frank 2013). Cognitive control is a phenomenon that applies to a number of cognitive domains, among which word production. Models within this framework assume that the resolution of competition involves at least 2 components. First, there is a detection of competition and, second, the detection of competition may trigger a modulation of cognitive processing to prevent performance from deteriorating. In this framework, studies of word production have found that competition resolution is associated with networks of activity in the left frontal lobe (e.g., Thompson-Schill et al. 1997; Wagner et al. 2001; Schnur et al. 2006, 2009; Spalek and Thompson-Schill 2008; Snyder et al. 2010, 2011). For example, several studies have associated portions of the left inferior frontal lobe with word selection and controlled semantic retrieval (e.g., Wagner et al. 2001; Schnur et al. 2009; Snyder et al. 2011), and others have associated the anterior cingulate cortex with conflict detection (e.g., Botvinick et al. 2004). Given that frontal lobe activity arises late in the course of word production (e.g., Salmelin et al. 1994), one may assume that this frontal lobe activity related to competition does not reflect word retrieval directly, but reflects a postretrieval mechanism related to conflict resolution (e.g., Badre et al. 2005; Badre and Wagner 2007). In other words, according to these models, competition resolution should involve a "late" component that operates after the retrieval of words from memory.

In sum, there are two different views on the time course of competition during word production. One view assumes that

competition is resolved during the retrieval of words from memory and, hence, predicts that competition is associated with a relatively early stage of word production. In contrast, models proposed within the framework of cognitive control associate competition with a late component in word production. Here, we used the temporal precision of electroencephalography (EEG) to distinguish between these 2 proposals.

To track the time course of competition during word production, we exploited a pattern of results obtained in the semantic blocking task (e.g., Kroll and Stewart 1994). In our version of this task, participants named pictures presented in homogeneous sets where all the pictures shared meaning (e.g., dog, snake, fly), or in heterogeneous sets where pictures did not share meaning (e.g., dog, table, pen). In addition, participants did not name the pictures once, but repeatedly named the pictures in the 2 types of sets on consecutive "presentation cycles" (e.g., {dog,snake, fly}$_{cycle1}$, {snake,fly,dog}$_{cycle2}$, {fly,dog,snake}$_{cycle3}$, etc.). Crucially, previous studies have demonstrated that competition in this task arises on the second and subsequent presentation cycles. Specifically, on the later presentation cycles, naming latencies to pictures in the homogeneous sets are slower compared with the heterogeneous sets (Belke et al. 2005; Damian and Als 2005). On the first presentation cycle, latencies in the homogeneous sets are faster than in the heterogeneous sets (Abdel Rahman et al. 2007; Navarrete et al. 2012). This pattern of results in naming latencies has been interpreted as due to competition on later cycles, whereas no competition is present on the first cycle. Thus, the way the semantic context effect varies as a function of presentation cycle presents us with a unique opportunity to track the time course of competition during word production.

The effect of semantic context in the semantic blocking task has been used previously by 2 studies to track the time course of competition in word production (Maess et al. 2002; Aristei et al. 2011). These electrophysiological studies revealed that the effect of semantic context was associated with an early component (peaking after 150 and before 400 ms), suggesting that competition is resolved relatively early in the course of word production. However, these results are inconclusive for 2 reasons. First, the interpretation that this early effect reflects competition during word retrieval was challenged by a recent study of Janssen et al. (2011). In that study, the authors argued that the early effect of semantic context in the semantic blocking task reflects semantic processes, and not competition resolution. Second, the studies of Janssen et al. (2011), Aristei et al. (2011) and Maess et al. (2002) have not investigated the presence of late effects of semantic context. Specifically, these studies only investigated the EEG signal only at early time points, and simply did not examine the presence of semantic context effects late in the signal. Thus, the question of whether competition resolution involves a late component remains unknown.

Participants in our experiment named the pictures in the homogeneous and heterogeneous sets on 4 presentation cycles of the semantic blocking task. Given that naming latencies in the task are around 700 ms, we examined the EEG signal over a range of 1000 ms such that both early and late effects of semantic context could be detected. On the first presentation cycle where naming latencies have revealed facilitatory effects of semantic context, we expected a pattern of electrophysiological results that is usually associated with semantic facilitation effects: the N400 component. This component is thought to reflect the processing of semantic information in memory (e.g.,

Ganis et al. 1996; Janssen et al. 2011; Kutas and Federmeier 2011).

On subsequent presentation cycles where naming latencies have revealed inhibitory effects of semantic context (and hence competition), we expected a modulation of the N400 effect if the origin of this inhibitory effect of semantic context in naming latencies reflects the resolution of competition during an early word retrieval stage (e.g., Roelofs 1992; Abdel Rahman et al. 2009; Oppenheim et al. 2010). In contrast, we expected a late component on subsequent cycles if the locus of the inhibitory effect of semantic context reflects the resolution of competition by a late, postretrieval mechanism (e.g., Botvinick et al. 2001; Yeung et al. 2004; Badre 2008; Wiecki and Frank 2013).

Before reporting the experimental results, we highlight one important methodological aspect of the study reported here. This aspect concerns the statistical approach for analyzing the electrophysiological data. The standard approach in the EEG literature relies on experimental designs where amplitudes are averaged across participants or items. However, there are problems associated with such an approach. Most importantly, by averaging across items, one ignores item variance, and by averaging across participants, one ignores participant variance. Ignoring these sources of variance leads to an increased probability of type I errors (e.g., Bagiella et al. 2000; Baayen et al. 2008). In order to avoid this increase in type I errors, we relied on so-called mixed-effect models, where the random variance associated with participants and items is estimated simultaneously. Although the use of mixed models is relatively common in the behavioral literature (e.g., Baayen et al. 2008; Barr et al. 2013), we could find only one study in the EEG literature (Amsel 2011). Here, we further introduced mixed models for the analysis of electrophysiological data.

Materials and Methods

Participants
Thirty-one native speakers of Spanish took part in the experiment. Participants were students at the Universidad de La Laguna, and received course credit or were paid 10 Euro. Five participants with excessively high numbers of rejected trials (>40%) were discarded from further analyses.

Design
Twenty-five line drawings were selected from a standardized picture database (Snodgrass and Vanderwart 1980) There were 5 semantic categories (i.e., vehicles, tools, animals, furniture, clothing), with 5 depicted instances per category. Pictures were standardized such that each picture fit within a rectangle of 300×300 pixels. Each picture appeared in the center of the screen as white-line drawings on a black background.

The homogeneous and heterogeneous conditions were created by arranging the picture names in a 5×5 matrix such that each row contained items from a homogeneous semantic category, and each column contained items from a heterogeneous category. The resulting 5 sets of homogeneous and heterogeneous items were combined into 5 blocks of homogeneous and 5 blocks of heterogeneous trials. A block consisted of a sequence of 20 trials, on which a given set of 5 homogeneous or heterogeneous items was repeated 4 times (i.e., 4 presentation cycles).

In line with other studies using this task (e.g., Damian et al. 2001), homogeneous and heterogeneous blocks did not alternate in the experiment, but were organized into groups of 5 consecutive homogeneous (A) or heterogeneous (B) blocks. The order of the homogeneous and heterogeneous groups was counterbalanced within and across participants (ABBA; BAAB). Each participant therefore named 400 trials in 10 homogeneous and 10 heterogeneous blocks.

Procedure

The experiment started with a familiarization phase in which participants were presented with a booklet containing all the pictures and their corresponding names. Participants were instructed to carefully study the pictures and their names. Next, they practiced the naming task by producing the name of each of the 25 pictures in the experiment once. On each trial, a fixation cross appeared for 700 ms, followed by the presentation of the target picture for 500 ms, or until the voice-key triggered. The next trial started after the presentation of a blank screen for 1200 ms. After this practice session, the actual experiment started with a trial structure that was identical to that used in the practice session.

Electrophysiological Recordings

The continuous EEG signal was recorded with 27 Ag/AgCl electrodes embedded in an elastic cap using Easycap (http://www.easycap.de). On-line recordings were referenced to the left mastoid. The signal was amplified (BrainAmp amplifiers) and digitized at a sampling rate of 250 Hz, with a 0.01- to 100-Hz band-pass filter. The horizontal electro-oculogram (EOG) was measured by placing 2 electrodes at the outer canthi, and the vertical EOG was measured with two electrodes placed above and below the left eye. Electrode impedance was kept below 5 kΩ for all electrodes.

After acquisition, the data were offline re-referenced using the average of the left and right mastoids. In the next step, the data were high-pass filtered at 0.15 Hz. Next, we extracted independent components using the independent component analysis (ICA) function of EEGLAB (Delorme and Makeig 2004). The ICA analysis was performed on EEG data from which stretches of EEG that were unrelated to the experiment (e.g., recorded EEG prior to the start of the experiment, pauses) had been removed. This ensured optimal detection of the ICA components. For each participant, at maximum 2 independent components, identified as vertical or horizontal eye movements, were subtracted from the EEG signal.

The EEG signal was then segmented into epochs of 1100 ms, starting 100 ms before the onset of the picture and ending 1000 ms after the onset of the picture. Each epoch was low-pass filtered at 30 Hz. No baseline correction was performed. This is because in the experimental design used in the current study, the experimental condition was known prior to the picture stimulus onset and, hence, subtracting a prestimulus baseline may reduce signal detection. In addition, given that the statistical analyses reported below relied on mixed-effect methods, the EEG signal was not averaged across participants or items. Instead, all epochs from all participants and all items were considered for further analysis.

Electrophysiological Analysis

Artifact removal was performed at the level of individual epochs, and at the level of individual amplitudes. First, individual epochs were removed from the analyses if they contained an amplitude outside the interval [−100, 100] μV, or if they contained a voltage fluctuation that exceeded 100 μV within a 200-ms window that moved through the epoch in 100 ms steps (10.6% of all amplitudes). In addition, epochs were removed for which the participant did not produce the correct target stimulus, or for which the target naming latency was <300 ms or was >1500 ms (2.5% of all amplitudes). Finally, individual amplitudes within an epoch that were > or <3 standard deviations plus or minus a participant's mean amplitude were considered outliers and also excluded from further analysis (0.4% of all amplitudes). Of a total number of 77 220 000 amplitudes (i.e., 26 subjects × 400 trials × 275 time-points × 27 electrodes), the total number of amplitudes included in the analyses was 67 064 652 (86.8%). (Note the 275 time-points reflects the length of the epoch divided by the sampling rate [i.e., $\frac{1100\,\text{ms}}{4} = 275$ ms].)

These amplitudes were analyzed in 2 subsequent stages. The primary function of the first stage of the analysis was to identify the precise temporal boundaries of the early and late time windows discussed in the Introduction of the article. This first stage is common (and often necessary) in EEG studies where the temporal boundaries of time windows are often influenced by idiosyncratic factors, which

prevent a quantitative definition of such windows in an a priori fashion. A common method is to visually inspect the grand average waveforms of a single representative electrode, which then allows for the precise temporal boundaries of expected time windows to be determined for further targeted analyses. Here, we replaced this visual inspection method with a more informative, statistical method which we have termed an all-time-points analysis. (Note that this test is also suitable for the analysis of EEG data when no a-priori information about the temporal window of an effect is available [see Lage-Castellanos et al. 2010, for discussion].)

In the all-time-points analysis, the degree to which semantic context predicted amplitude was computed at every time-point, and for every electrode across the 1100 ms range. The degree to which amplitudes were predicted by the semantic context variable were estimated using mixed-effect modeling. As we explained above, mixed-effect models allow for the analysis of EEG data that is not averaged across participants or items, and is specified at the single-trial level. Our statistical model included the fixed-effect predictor semantic context, and included by-participant and by-item random intercepts. In addition, we included a control predictor trial. This predictor was assumed to control for the longitudinal variation in a participant's amplitude along the course of the experiment. However, as the focus in this study is on the effect of semantic context, the effect of trial is not further discussed. In order to reduce problems with multicollinearity, all fixed-effect predictors were centered on their sample means. We computed this all-time-points analysis separately for each individual cycle.

The effects of semantic context in these mixed-effect analyses are expressed using t-tests. We report the sign-preserving P-values of these tests, where positive P-values indicate that the heterogeneous condition is more negative than the homogeneous condition, and negative P-values indicate the opposite situation. The P-values were computed directly from the observed t-value, using as degrees of freedom the number of data points minus the number of fixed-effect predictors in the model. We repeated this procedure for all 27 electrodes. To protect against problems associated with multiple comparisons, we considered significant only those P-values that were <0.05 and extended consecutively over at least 48 ms (Guthrie and Buchwald 1991). As we mentioned above, the primary function of the all-time-points analysis was to determine the precise temporal boundaries of the early and late time windows identified on the basis of previous studies discussed in the Introduction of the article. In other words, the results obtained in the first stage of the analysis were not used for hypothesis testing, which was done in the second stage of the analysis.

The second stage of the analysis was a targeted statistical test for which inferences were warranted (i.e., a region of interest analysis). In this stage, we computed the effect of semantic context for the early and late time windows that were discussed in the Introduction section, and whose precise temporal boundaries were identified by the all-time-points test. The analyses also relied on the mixed-effect regression technique. However, instead of computing t-tests in these 2 time windows at every time-point, statistical effects were estimated from the amplitudes averaged across time-points in the 2 time windows. Thus, for every participant, electrode, and trial, we averaged across all time-points in each time window to obtain a single average amplitude for that time window. Note that this averaging procedure preserves the availability of the EEG data at the single-trial level because averaging took place across time-points within a trial, and no averaging across participants or items was performed. Also note that averaging across time-points drastically reduces the number of statistical tests that are performed, thereby reducing the problem of multiple comparisons.

In the analyses of the first and second windows, we examined how the semantic context effect depended on electrode location. To this end, we included a factor anteriority with 2 levels: anterior (F7, F3, F4, F8, FC5, FC1, FC2, FC6, Fp1, Fp2, Fz) versus posterior (P7, P3, P4, P8, CP5, CP1, CP2, CP6, O1, O2, Pz). The statistical model considered in the analyses included the fixed-effect predictors semantic context, anteriority, and their interaction, as well as the control predictor trial. As before, predictors were centered to reduce issues with collinearity. The random-effect structure included by-participant and by-item random intercepts and by-participant random slopes for all fixed-effect predictors. We did not include a parameter for the correlation between

random intercepts and slopes. The method for computing *P*-values relied on a technique called likelihood ratio testing, where the goodness of fit of a model containing a given fixed-effect predictor is compared with a model without the predictor but that is otherwise identical in random-effect structure (e.g., Barr et al. 2013). This likelihood ratio test follows a χ^2 distribution, from which *P*-values can be obtained. This method of obtaining *P*-values differs from that used in the all-time-points analysis. The likelihood ratio method is computationally intensive, and would have been problematic for the all-time-points analysis. As before, we computed these tests separately for each individual cycle.

Results

Behavioral Data

The reaction times were analyzed using a mixed-effect model structure that was comparable with that used in the analysis of the EEG data. We removed from the analyses all trials on which participants produced an incorrect name for the picture (2.3%), as well as trials on which the the target naming latency was <300 ms or was >1500 ms (0.1%). Of a total of 10 400 reaction times (i.e., 26 subjects × 400 trials), the total number of reaction times included in the analyses was 10 143 (97.5%).

The statistical model included fixed-effect predictors for trial and semantic context, and included by-participant and by-item random intercepts, as well as by-participant random slopes for the fixed-effect predictors. All fixed-effect predictors were centered to reduce potential problems with collinearity. As the focus in the current study is on the effect of semantic context, the effect of trial is not further discussed. *P*-values were obtained using likelihood ratio testing described above. We computed the effect of semantic context on reaction times separately for each presentation cycle.

The results confirmed those previously observed (Abdel Rahman et al. 2007; Navarrete et al. 2012). Specifically, on the first presentation cycle, participants named pictures in the homogeneous sets "faster" than pictures in the heterogeneous sets. In contrast, on the second, third, and fourth cycles, they named pictures in the homogeneous sets "slower" than pictures in the heterogeneous sets. (See Table 1 for details of the statistics, and Fig. 1 for a graphical presentation of the reaction time effects.)

Electrophysiological Data

Visual inspection of the results from the all-time-points analysis revealed the presence of the exact temporal boundaries of the early and late time windows discussed in the Introduction section (see Fig. 2). In an early time window between ~250 and 400 ms, positive *P*-values were found, suggesting that the heterogeneous condition was more negative than the homogeneous condition. In a later window between ~500 and 750 ms, negative *P*-values were found, suggesting that the homogeneous condition was more negative than the heterogeneous condition. These early and late effects appeared to depend on the cycle. Whereas the early effect seemed to be present only on cycles 1, 2, and 3, the later effect appeared on cycles 2, 3, and 4. Figure 3 presents the time course of the homogeneous and heterogeneous conditions for representative electrode Fz for each cycle, and Figure 4 presents a topographical representation of the amplitude differences between the homogeneous and heterogeneous conditions in the 2 time windows for each cycle.

Table 1
Overview of the behavioral results from the experiment

Variable	Cycle	β (standard error)	t	$\chi^2(1)$	P
Semantic context	1	−17.82 (7.20)	−2.47	5.68	0.0172
	2	20.22 (6.13)	3.30	9.41	0.0022
	3	35.37 (5.42)	6.52	25.81	0.0001
	4	27.77 (7.91)	3.51	10.41	0.0013

Listed are the beta-coefficient and its standard error, the corresponding *t*-value, as well as the χ^2 and *P*-value obtained from likelihood ratio tests (see text for details) for the variable semantic context as a function of presentation cycle. Note the polarity of the semantic context effect varied between cycle 1 vs. cycles 2–4.
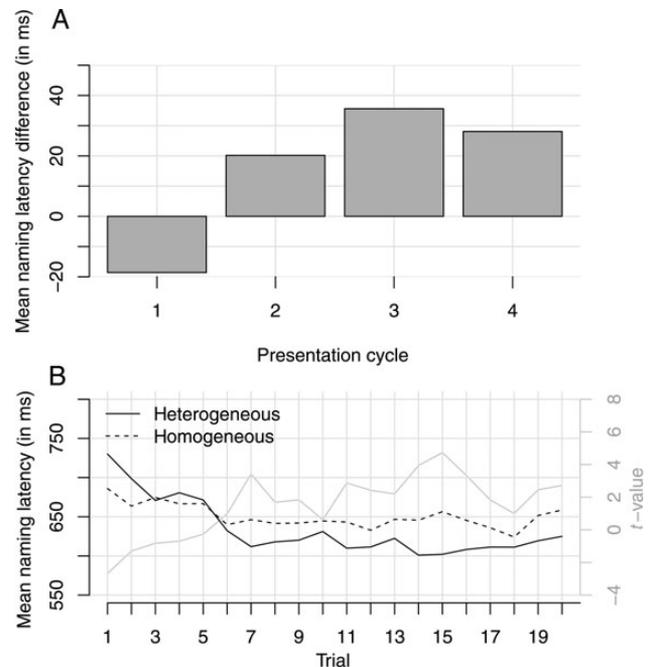

Figure 1. The progression of the behavioral effect of semantic context. Panel *A* shows how the effect varies as a function of presentation cycle. Panel *B* shows this effect in greater detail, and displays the effect across individual trials (each presentation cycle is averaged across 5 trials). Also plotted in panel *B* are the *t*-values of the context effect across trials (gray line). Note how the effect is facilitatory on the first cycle, and becomes inhibitory on subsequent cycles.

These observations were confirmed with targeted mixed-effect analyses of the semantic context effect in the early and late time windows whose temporal boundaries were determined by the all-time-points analysis. In the "early time window" (250–400 ms), on the first presentation cycle, there was an effect of semantic context, where amplitudes in the heterogeneous condition were more negative than in the homogeneous condition. This effect was also marginally present on cycle 2 and was robust on cycle 3. Finally, on cycle 4, no semantic context effect was found (see Table 2 for statistics). Further exploration of the interaction between anteriority and semantic context found for cycle 3 revealed that the effect of semantic context was found primarily on anterior ($t = 2.77$; $P = 0.0083$), and not posterior electrodes ($t = 1.25$; $P = 0.2076$).

In the "late time window" (500–750 ms), on the first presentation cycle, there was no effect of semantic context. In contrast, on cycles 2, 3, and 4, there were effects of semantic context (see Table 2 for statistics). On these cycles, amplitudes in the heterogeneous condition were more positive than in the
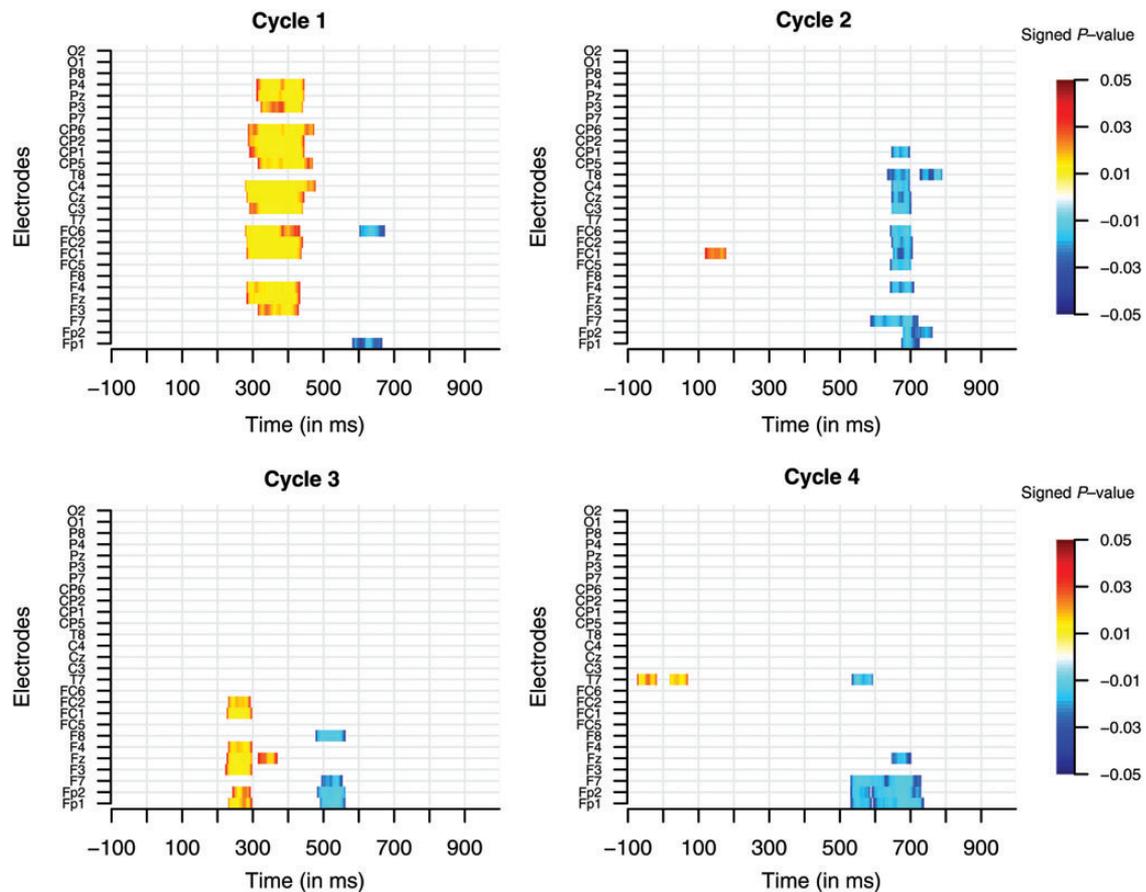
Figure 2. Exploratory all-time-points analysis of the electrophysiological effect of semantic context. Plotted are the *P*-values associated with the effect of semantic context at a particular time-point, electrode, and presentation cycle. *P*-values are signed, meaning that they indicate the direction of the semantic effect. To control for type I errors, *P*-values were plotted if were in the interval [−0.05, 0.05], and if they extended consecutively over a period of at least 48 ms (Guthrie and Buchwald 1991). Positive *P*-values (yellow) indicate that the heterogeneous condition is more negative than the homogeneous condition, whereas negative *P*-values (blue) indicate the opposite situation. Note how positive *P*-values are associated with an early time window, and how negative *P*-values are linked to a later time window and appear on cycles 2–4.

homogeneous condition. Further exploration of the interaction between anteriority and semantic context found for cycle 4 revealed that the effect of semantic context was found primarily on anterior ($t = −2.19$; $P = 0.0324$), and not posterior electrodes ($t = −1.29$; $P = 0.1967$). Figure 5 presents a graphical representation of the mean amplitude differences between the homogeneous and heterogeneous conditions across the four presentation cycles in the early and late time windows.

Discussion

In the current study, we tracked the time course of competition during word production. Participants named small sets of pictures in semantically homogeneous and heterogeneous contexts on 4 consecutive presentation cycles. On the first cycle, we found faster latencies in the homogeneous than in the heterogeneous sets (Abdel Rahman et al. 2007; Navarrete et al. 2012). On subsequent cycles, there were inhibitory effects of semantic context in the naming latencies, suggesting the presence of competition (Belke et al. 2005; Damian and Als 2005). The electrophysiological data revealed a semantic context effect in 2 time windows that further depended on the presentation cycle. Specifically, there was an early effect that appeared between 250 and 400 ms and was present on cycles 1–3. In addition, there was a late frontally distributed component that appeared between 500 and 750 ms and was present on cycles 2–4.

In the early window (250–400 ms), the polarity, scalp distribution and latency of the semantic effect are in line with those commonly observed in electrophysiological studies of semantic effects with pictorial stimuli (e.g., Ganis et al. 1996; Janssen et al. 2011). These N400 effects are commonly interpreted in terms of the processing of semantic information (e.g., Frenzel et al. 2011; Kutas and Federmeier 2011). One possible interpretation of this early effect would be in terms of the integration of semantic information (Lau et al. 2008). In this case, this early effect could reflect the differences in the ability of the system to integrate semantic information in the homogeneous and heterogeneous contexts. Alternatively, this early component could reflect the retrieval of semantic information from memory (Kutas and Federmeier 2011). This latter interpretation would be in line with other electrophysiological studies that have found that memory access during word production starts around 200 ms poststimulus onset (Costa et al. 2009; Strijkers et al. 2010; Blackford et al. 2012).

The critical finding reported here concerns the effect of semantic context in the late time window (500–750 ms). This late electrophysiological effect of semantic context was strongly correlated with the competition effect observed in the behavioral data. Specifically, the late electrophysiological effect was absent on cycle 1, and was present across cycles 2, 3, and 4, exactly like the competition effect in the naming latencies. In addition, the shift in polarity of the electrophysiological effect
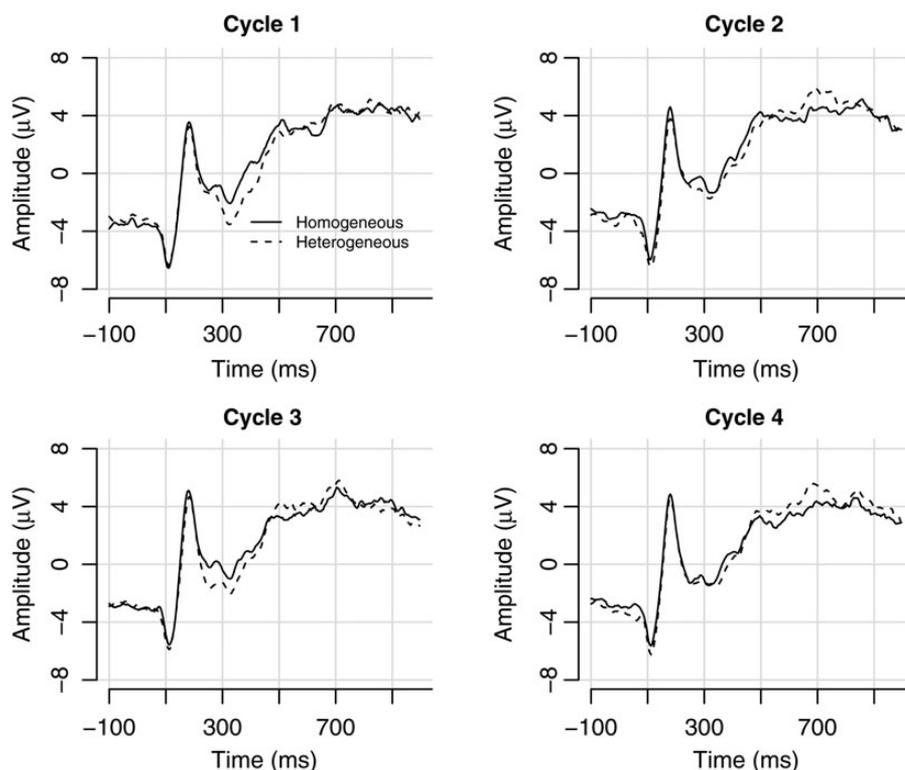
Figure 3. Average amplitudes of the heterogeneous and homogeneous conditions as a function of time and presentation cycle for electrode Fz. Note how the polarity of the semantic effect varies as a function of both time and cycle: Heterogeneous is more negative than homogeneous at an early time window, whereas heterogeneous is more positive than homogeneous at a later time window on cycles 2–4.
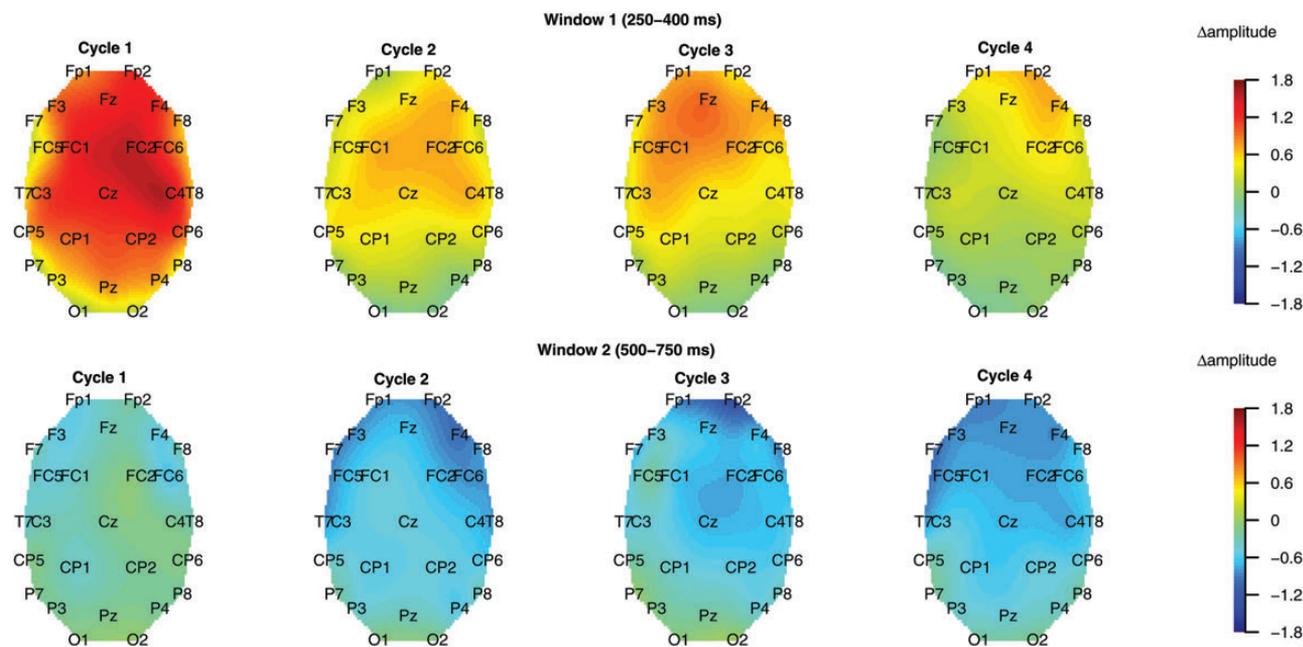

Figure 4. Topographic presentation of the amplitude differences between the homogeneous and the heterogeneous condition as a function of electrode location and presentation cycle. Note that, for window 1 (top row), there are mainly positive differences, whereas in window 2 (bottom row), there are mainly negative differences on cycles 2–4. Interpolation of the difference amplitudes based on thin plate splines (Perrin et al. 1987).

paralleled that of the naming latencies: The effect of semantic context in the naming latencies was facilitatory on cycle 1 and inhibitory on cycles 2–4. Similarly, whereas the heterogeneous condition was more negative than the homogeneous condition on cycle 1 (in the early time window), this polarity of the semantic effect reversed on cycles 2–4 (in the late time window). The tight correlation between the late electrophysiological component and the behavioral effect permits the identification of this late component with processing related to competition.

Table 2
Overview of the electrophysiological results from the experiment

Window	Variables	Cycle	β (standard error)	t	$\chi^2(1)$	P
Window 1	Semantic Context	1	0.61 (0.28)	2.15	4.39	0.0362
	Anteriority	1	4.78 (0.85)	5.59	21.09	0.0001
	Sem Cont:Anter	1	−0.12 (0.26)	−0.45	0.22	0.6399
	Semantic Context	2	0.35 (0.20)	1.79	3.22	0.0727
	Anteriority	2	5.08 (0.84)	6.05	23.40	0.0001
	Sem Cont:Anter	2	−0.30 (0.20)	−1.48	2.16	0.1413
	Semantic Context	3	0.52 (0.19)	2.69	6.62	0.0101
	Anteriority	3	5.46 (0.82)	6.64	26.39	0.0001
	Sem Cont:Anter	3	−0.63 (0.29)	−2.15	4.42	0.0356
	Semantic Context	4	0.09 (0.26)	0.33	0.11	0.7350
	Anteriority	4	5.56 (0.74)	7.45	30.38	0.0001
	Sem Cont:Anter	4	−0.19 (0.29)	−0.68	0.48	0.4905
Window 2	Semantic Context	1	−0.37 (0.36)	−1.02	1.07	0.3018
	Anteriority	1	−1.97 (0.63)	−3.01	8.47	0.0036
	Sem Cont:Anter	1	0.38 (0.24)	1.53	2.34	0.1263
	Semantic Context	2	−0.75 (0.31)	−2.43	5.53	0.0187
	Anteriority	2	−3.41 (0.65)	−5.22	19.15	0.0001
	Sem Cont:Anter	2	0.52 (0.33)	1.54	2.37	0.1240
	Semantic Context	3	−0.57 (0.31)	−1.84	3.29	0.0696
	Anteriority	3	−3.50 (0.62)	−5.63	21.27	0.0001
	Sem Cont:Anter	3	0.52 (0.42)	1.22	1.51	0.2187
	Semantic Context	4	−0.73 (0.35)	−2.08	4.14	0.0418
	Anteriority	4	−3.78 (0.60)	−6.26	24.54	0.0001
	Sem Cont:Anter	4	0.77 (0.40)	1.90	3.53	0.0603

Listed are the beta-coefficient and its standard error, the corresponding t-value, as well as the χ^2 and P-value obtained from likelihood ratio tests (see text for details) for the variables semantic context, anteriority and their interaction as a function of presentation cycle in time windows 1 and 2.
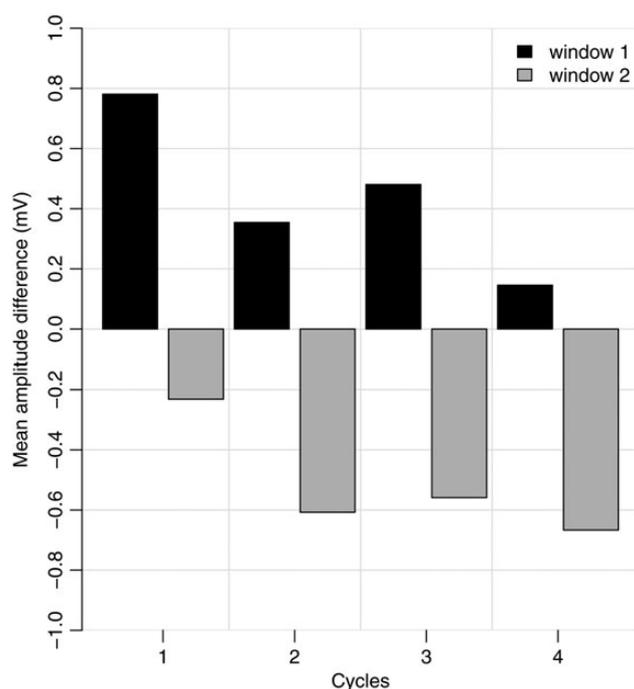

Figure 5. Graphical presentation of the differences between the mean amplitudes in the homogeneous and heterogeneous condition as a function of window and cycle. Note how in window 1 (black bars), differences are positive and large on the first cycle, whereas in window 2 (gray bars), differences are negative and large on cycles 2–4.

The presence of this late component is inconsistent with a class of models that associates the resolution of competition with the retrieval of words from memory (Roelofs 1992; Abdel Rahman et al. 2009; Oppenheim et al. 2010). Such models assume that the inhibitory effects of semantic context in the

naming latencies observed on cycles 2–4 reflect a mechanism of competition resolution that operates during the retrieval of words from memory. Such models therefore predict a modulation of the early electrophysiological effect of semantic context observed by the presentation cycles. At odds with this prediction, our results revealed that the inhibitory effect of semantic context in the naming latencies was associated with a late electrophysiological effect. Given that the average picture naming response in the experiment was around 650 ms (see Fig. 1), this late electrophysiological effect clearly reflects processes that arise after the retrieval of words from memory. Accordingly, the current results are at odds with models that associate the resolution of competition with the retrieval of words from memory.

Instead, the current results find a ready explanation in models of word production formulated within the context of cognitive control (e.g., Thompson-Schill et al. 1997; Wagner et al. 2001; Schnur et al. 2006; Snyder et al. 2010). These models have associated the resolution of competition with frontal lobe structures that play a role in word production after the retrieval of words from memory. These models therefore predict that competition resolution should be associated with a late component, which is precisely the pattern of results observed here.

What is the exact interpretation of this late component associated with competition? One possibility is that this late component reflects the detection of conflict during word production. This interpretation would be consistent with a number of studies in the language and cognitive control literatures that have found a late component associated with error detection (e.g., Botvinick et al. 2001; Kolk et al. 2003; Yeung et al. 2004; Ganushchak and Schiller 2008; Riès et al. 2011). The neural structure that is frequently associated with this late component is the anterior cingulate cortex (Botvinick et al. 2001). Within the context of our experiment, this late component could index the degree to which competition and associated performance problems arise during the retrieval of words from memory. Specifically, on cycle 1, when competition is weak and few word retrieval errors occur, no activity in the anterior cingulate cortex is expected, leading to no observed late component. In contrast, on cycles 2–4, when competition is strong and performance deteriorates, strong activity in the anterior cingulate cortex is expected, thereby explaining the appearance of the late component on these cycles. Thus, in this interpretation, the late component is thought to reflect a detection mechanism that is sensitive to the quality of performance in other neural areas.

Another possible interpretation of the late component is in terms of the selection of competing representations in frontal lobe structures (e.g., Thompson-Schill et al. 1997, 1998; Schnur et al. 2009; Snyder et al. 2010). One recent model has proposed that the left ventro-lateral prefrontal cortex is involved in resolving competition between words during language production (Snyder et al. 2010). This model assumes that the left ventro-lateral prefrontal cortex instantiates a mechanism of lateral inhibition which is used to resolve competition between words. The late component observed in the current experiment may index the activity of this mechanism of lateral inhibition. This framework also provides an explanation of the modulation of the late component by the presentation cycles observed here. Specifically, on cycle 1, no competition is present, meaning little lateral inhibition is required, leading to no activity in the left prefrontal cortex, and consequently, no

observed late component. In contrast, on cycles 2–4, competition is present, meaning that lateral inhibition is required, leading to activity in the left prefrontal cortex, and the observation of a late component. Thus, in this interpretation, the late component is assumed to reflect processes related to the selection of words. (Note that these 2 interpretations of the late component are not mutually exclusive. It may well be the case that the late component reflects the activity of a detection and a selection mechanism.)

One aspect of the data presented here can be used to distinguish between a detection and selection interpretation of the late component. Specifically, the detection and selection interpretations differ with respect to the role of the late component in determining the pattern of naming latencies in the semantic blocking task. If the late component reflects detection, the neural processes underlying this component play no causal role in determining naming latencies (i.e., they only reflect "detection" of competition). Accordingly, the facilitatory and inhibitory effects of semantic context observed in the naming latencies must be determined by the neural processes underlying the early component. This predicts that there should be a correlation between the electrophysiological effect of semantic context observed in the early time window and the reaction time effect. However, this correlation is not present: Whereas the reaction time effect changes polarity from facilitatory to inhibitory across presentation cycles, the electrophysiological effect in the early time window decreases across cycles and does not change polarity (see Fig. 5). Thus, a detection interpretation of the late component seems unlikely.

Alternatively, if the late component reflects selection processes, the neural processes underlying this component can directly influence naming latencies. In this scenario, the neural processes underlying the late component are responsible for the inhibitory effect in the naming latencies. Accordingly, the neural processes underlying the early component would be responsible for the facilitatory effect. This scenario fits well with the observed data: when the facilitatory effect is observed on presentation cycle 1, the early component is largest and the late component is absent. On the later presentation cycles where inhibitory effects are found, the early component is small in magnitude and the late component predominates (see Fig. 5). Thus, the overall correlation between the effects of semantic context across presentation cycles in the naming latencies and the electrophysiological effect of semantic context in the early and late time windows favors an interpretation of the late component in terms of selection processes.

In addition, the selection interpretation of the late component receives support from studies in the behavioral literature. These studies have used the picture-word interference paradigm to examine word production processes under conditions when producing a response no longer requires memory retrieval. For example, Janssen et al. (2008) and Dhooge and Hartsuiker (2011) have shown that specific properties of a prime word (i.e., semantic relationship with target, lexical frequency) can influence a naming response, even when that naming response no longer directly relies on memory retrieval. (The status of results obtained with this delayed naming version of the picture-word interference task is frequently debated (see, e.g., Spalek et al. 2013). Often not considered in this debate are the data of Spalek and Thompson-Schill (2008), who reported an fMRI study on semantic inference effects using a delayed naming task that resembles that of Janssen et al. (2008). In line with the observations of Janssen et al. (2008), Spalek and Thompson-Schill reported clear effects of semantic context in both temporal and frontal regions using this task, which they interpreted in terms of competition effects during lexical processing. This finding therefore further illustrates the presence of competition effects at a late time-point in the production of a word (Janssen et al. 2008).) The conclusion from these studies is that performance in the picture-word interference task relies on a mechanism of word selection that operates after the retrieval of words from memory (e.g., Miozzo and Caramazza 2003; Finkbeiner and Caramazza 2006; Mahon et al. 2007). These data therefore provide further independent support for the interpretation of the late component observed in the current experiment in terms of postretrieval selection processes.

A striking observation of the results presented here is the change in the polarity of the behavioral effect of semantic context across presentation cycles. Whereas the effect of semantic context is facilitative on cycle 1 (i.e., homogeneous faster than heterogeneous), the same effect is inhibitory on subsequent presentation cycles (i.e., homogeneous slower than heterogeneous). The cognitive control framework proposed here provides an explanation for this modulation of the semantic effect by presentation cycle. Specifically, on cycle 1, the presence of shared semantic features between words from the homogeneous condition lead to priming between words. Given that such shared semantic features are not present for words in the heterogeneous condition, one would expect semantic facilitation effects due to semantic priming on the first presentation cycle (e.g., Meyer and Schvaneveldt 1971). However, a negative consequence of the repeated naming of words with shared semantic features is that their activity levels will become increasingly similar, leading to strong competition and increased likelihoods for making errors (i.e., selecting the wrong word). This increased possibility for errors will lead to increased activity in the anterior cingulate cortex which may cue the onset of a winner-take-all selection mechanism in the form of lateral inhibition housed in left frontal regions (e.g., Snyder et al. 2010). The neural processes in this region are responsible for the observation of semantic interference effects on cycles 2–4. Thus, the pattern of behavioral results obtained here can be explained within a cognitive control framework.

Finally, the data reported here also have important implications for an ongoing debate in the behavioral literature on the nature of the mechanism of word selection (e.g., Mahon et al. 2007; Abdel Rahman et al. 2009; Spalek et al. 2013). This debate has centered on the question of the nature of the mechanism of word selection in language production. Specifically, 2 distinct hypotheses have been proposed: One where word selection is based on a winner-take-all principle using relative activations, and another where word selection is not based on a winner-take-all mechanism (e.g., Janssen and Caramazza 2011). The current results suggest that a third hypothesis should be entertained. Specifically, the current results suggest that word selection may rely on a winner-take-all mechanism, but only in circumstances where strong competition is present. When competition is absent, as is the case on the first presentation cycle in the semantic blocking task, this winner-take-all mechanism is dormant. When competition is present, as is the case on later presentation cycles, this mechanism becomes active. In other words, the dichotomy between a word selection mechanism that either employs a winner-take-all mechanism or not may be insufficient to characterize the available data.

To conclude, the current results reveal that the resolution of competition during word production is associated with a late component that arises around 500- to 750-ms poststimulus onset. Given that the average naming latency in the experiment was about 650 ms, this late component reflects processes that arise after the retrieval of words from memory. These data present a challenge to a class of models that associate the inhibitory effects of semantic context with the resolution of competition during the retrieval of words from memory. In contrast, these data are supported by models formulated within the context of cognitive control, which assume that late acting frontal lobe structures play a crucial role in competition resolution. Within the cognitive control framework, the late component may reflect the detection of conflict among competing representations in temporal regions by the anterior cingulate cortex, and/or the selection of competing representations via a winner-take-all mechanism present in the left inferior frontal gyrus. Future studies that can reveal the dynamics of localized activities in the temporal and frontal lobe regions are necessary to further clarify these possible interpretations of the late component.

Funding

Notes

Conflict of Interest: None declared.

References

Abdel Rahman RA, Melinger A. 2009. Semantic context effects in language production: a swinging lexical network proposal and a review. Lang Cognitive Proc. 24(5):713–734.

Abdel Rahman RA, Melinger A. 2007. When bees hamper the production of honey: lexical interference from associates in speech production. J Exp Psychol Learn. 33(3):604–614.

Amsel BD. 2011. Tracking real-time neural activation of conceptual knowledge using single-trial event-related potentials. Neuropsychologia. 49(5):970–983.

Aristei S, Melinger A, Rahman RA. 2011. Electrophysiological chronometry of semantic context effects in language production. J Cogn Neurosci. 23(7):1567–1586.

Baayen RH, Davidson DJ, Bates DM. 2008. Mixed-effects modeling with crossed random effects for subjects and items. J Mem Lang. 59:390–412.

Badre D. 2008. Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. Trends Cogn Sci. 12(5):193–200.

Badre D, Poldrack RA, Paré-Blagoev EJ, Insler RZ, Wagner AD. 2005. Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. Neuron. 47(6):907–918.

Badre D, Wagner AD. 2007. Left ventrolateral prefrontal cortex and the cognitive control of memory. Neuropsychologia. 45(13):2883–2901.

Bagiella E, Sloan RP, Heitjan DF. 2000. Mixed-effects models in psychophysiology. Psychophysiology. 37(1):13–20.

Barr DJ, Levy R, Scheepers C, Tily HJ. 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. J Mem Lang. 68(3):255–278.

Belke E, Meyer AS, Damian MF. 2005. Refractory effects in picture naming as assessed in a semantic blocking paradigm. Q J Exp Psychol. 58:667–692.

Blackford T, Holcomb P, Grainger J, Kuperberg G. 2012. A funny thing happened on the way to articulation: N400 attenuation despite behavioral interference in picture naming. Cognition. 123:84–99.

Botvinick M, Braver T, Barch D, Carter C, Cohen J. 2001. Conflict monitoring and cognitive control. Psychol Rev. 108(3):624–652.

Botvinick M, Cohen J, Carter C. 2004. Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn Sci. 8(12):539–546.

Caramazza A. 1997. How many levels of processing are there in lexical access? Cogn Neurop. 14:177–208.

Costa A, Strijkers K, Martin C, Thierry G. 2009. The time course of word retrieval revealed by event-related brain potentials during overt speech. Proc Natl Acad Sci USA. 106:21442–21446.

Damian MF, Als LC. 2005. Long-lasting semantic context effects in the spoken production of object names. J Exp Psychol Learn. 31 (6):1372–1384.

Damian MF, Vigliocco G, Levelt WJ. 2001. Effects of semantic context in the naming of pictures and words. Cognition. 81(3):B77–B86.

Dell GS. 1986. A spreading-activation theory of retrieval in sentence production. Psychol Rev. 93(3):283–321.

Delorme A, Makeig S. 2004. EEGlab: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Meth. 134(1):9–21.

Dhooge E, Hartsuiker RJ. 2011. The distractor frequency effect in a delayed picture-word interference task: further evidence for a late locus of distractor exclusion. Psychon B Rev. 18(1):116–122.

Finkbeiner M, Caramazza A. 2006. Now you see it, now you don't: on turning semantic interference into facilitation in a stroop-like task. Cortex. 42(6):790–796.

Frenzel S, Schlesewsky M, Bornkessel-Schlesewsky I. 2011. Conflicts in language processing: a new perspective on the n400–p600 distinction. Neuropsychologia. 49(3):574–579.

Ganis G, Kutas M, Sereno MI. 1996. The search for "common sense": an electrophysiological study of the comprehension of words and pictures in reading. J Cogn Neurosci. 8:89–106.

Ganushchak L, Schiller N. 2008. Brain error-monitoring activity is affected by semantic relatedness: an event-related brain potentials study. J Cogn Neurosci. 20(5):927–940.

Guthrie D, Buchwald JS. 1991. Significance testing of difference potentials. Psychophysiology. 28(2):240–244.

Janssen N. 2013. Response exclusion in word–word tasks: a comment on roelofs, piai and schriefers. Lang Cognitive Proc. 28(5):672–678.

Janssen N, Caramazza A. 2011. Lexical selection in multi-word production. Front Psychol. 2:81.

Janssen N, Carreiras M, Barber H. 2011. Electrophysiological effects of semantic context in picture and word naming. Neuroimage. 57:1243–1250.

Janssen N, Schirm W, Mahon BZ, Caramazza A. 2008. Semantic interference in a delayed naming task: evidence for the response exclusion hypothesis. J Exp Psychol Learn. 34(1):249–256.

Kolk HH, Chwilla DJ, van Herten M, Oor PJ. 2003. Structure and limited capacity in verbal working memory: a study with event-related potentials. Brain Lang. 85(1):1–36.

Kroll JF, Stewart E. 1994. Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. J Mem Lang. 33:149–174.

Kutas M, Federmeier KD. 2011. Thirty years and counting: finding meaning in the n400 component of the event related brain potential (ERP). Annu Rev Psychol. 62:621–647.

Lage-Castellanos A, Martínez-Montes E, Hernández-Cabrera J, Galán L. 2010. False discovery rate and permutation test: an evaluation in ERP data analysis. Stat Med. 29(1):63–74.

Lau EF, Phillips C, Poeppel D. 2008. A cortical network for semantics: (de) constructing the n400. Nat Rev Neurosci. 9(12):920–933.

Maess B, Friederici AD, Damian M, Meyer AS, Levelt WJ. 2002. Semantic category interference in overt picture naming: sharpening current density localization by PCA. J Cogn Neurosci. 14 (3):455–462.

Mahon BZ, Costa A, Peterson R, Vargas KA, Caramazza A. 2007. Lexical selection is not by competition: a reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. J Exp Psychol Learn. 33(3):503–535.

Meyer DE, Schvaneveldt RW. 1971. Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. J Exp Psychol. 90(2):227–234.

Miozzo M, Caramazza A. 2003. When more is less: a counterintuitive effect of distractor frequency in the picture–word interference paradigm. J Exp Psychol Gen. 132(2):228–252.

Navarrete E, Del Prato P, Mahon BZ. 2012. Factors determining semantic faciliation and interference in the cyclic naming paradigm. Front. Psychol. 38:1–15.

Oppenheim G, Dell G, Schwartz M. 2010. The dark side of incremental learning: a model of cumulative semantic interference during lexical access in speech production. Cognition. 114(2):227–252.

Perrin F, Pernier J, Bertnard O, Giard M, Echallier J. 1987. Mapping of scalp potentials by surface spline interpolation. Electroencephalogr Clin Neurophysiol. 66(1):75–81.

Riès S, Janssen N, Dufau S, Alario F, Burle B. 2011. General-purpose monitoring during speech production. J Cogn Neurosci. 23 (6):1419–1436.

Roelofs A. 1992. A spreading-activation theory of lemma retrieval in speaking. Cognition. 42:107–142.

Salmelin R, Hari R, Lounasmaa O, Sams M. 1994. Dynamics of brain activation during picture naming. Nature. 368(6470):463–465.

Schnur TT, Schwartz MF, Brecher A, Hodgson C. 2006. Semantic interference during blocked-cyclic naming: evidence from aphasia. J Mem Lang. 54(2):199–227.

Schnur TT, Schwartz MF, Kimberg DY, Hirshorn E, Coslett HB, Thompson-Schill SL. 2009. Localizing interference during naming: convergent neuroimaging and neuropsychological evidence for the function of Broca's area. Proc Natl Acad Sci USA. 106(1):322–327.

Schriefers H, Meyer AS, Levelt WJ. 1990. Exploring the time course of lexical access in language production: picture-word interference studies. J Mem Lang. 29(1):86–102.

Snodgrass J, Vanderwart M. 1980. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. J Exp Psychol Human. 6(2):174–215.

Snyder H, Banich MT, Munakata Y. 2011. Choosing our words: retrieval and selection processes recruit shared neural substrates in left ventrolateral prefrontal cortex. J Cogn Neurosci. 23 (11):3470–3482.

Snyder HR, Hutchison N, Nyhus E, Curran T, Banich MT, O'Reilly RC, Munakata Y. 2010. Neural inhibition enables selection during language processing. Proc Natl Acad Sci USA. 107(38):16483–16488.

Spalek K, Damian MF, Bölte J. 2013. Is lexical selection in spoken word production competitive? Introduction to the special issue on lexical competition in language production. Lang Cognitive Proc. 28(5):597–614.

Spalek K, Thompson-Schill SL. 2008. Task-dependent semantic interference in language production: an fMRI study. Brain Lang. 107 (3):220–228.

Strijkers K, Costa A, Thierry G. 2010. Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. Cereb Cortex. 20(4):912–928.

Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ. 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. Proc Natl Acad Sci USA. 94(26):14792–14797.

Thompson-Schill SL, Swick D, Farah MJ, D'Esposito M, Kan IP, Knight RT. 1998. Verb generation in patients with focal frontal lesions: a neuropsychological test of neuroimaging findings. Proc Natl Acad Sci USA. 95(26):15855–15860.

Wagner A, Paré-Blagoev E, Clark J, Poldrack R. 2001. Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. Neuron. 31(2):329–338.

Wiecki TV, Frank MJ. 2013. A computational model of inhibitory control in frontal cortex and basal ganglia. Psychol Rev. 120(2):329.

Yeung N, Botvinick MM, Cohen JD. 2004. The neural basis of error detection: conflict monitoring and the error-related negativity. Psychol Rev. 111(4):931.