



ERP evidence for ultra-fast semantic processing in the picture–word interference paradigm

Roberto Dell'Acqua^{1,2*}, Paola Sessa¹, Francesca Peressotti¹, Claudio Mulatti¹, Eduardo Navarrete¹ and Jonathan Grainger³

¹ Cognition and Language Laboratory, Department of Developmental Psychology, University of Padova, Padova, Italy

² Centre for Cognitive and Brain Sciences, University of Padova, Padova, Italy

³ Laboratoire de Psychologie Cognitive, Aix-Marseille Université and CNRS, Marseille, France

Edited by:

Thomas F. Münte, University of Magdeburg, Germany

Reviewed by:

Thomas F. Münte, University of Magdeburg, Germany
Nicole Wicha, University of Texas at San Antonio, USA

*Correspondence:

Roberto Dell'Acqua, Centre for Cognitive and Brain Sciences, University of Padova, Via Venezia 8, 35131 Padova, Italy.
e-mail: dar@unipd.it

We used the event-related potential (ERP) approach combined with a subtraction technique to explore the timecourse of activation of semantic and phonological representations in the picture–word interference paradigm. Subjects were exposed to to-be-named pictures superimposed on to-be-ignored semantically related, phonologically related, or unrelated words, and distinct ERP waveforms were generated time-locked to these different classes of stimuli. Difference ERP waveforms were generated in the semantic condition and in the phonological condition by subtracting ERP activity associated with unrelated picture–word stimuli from ERP activity associated with related picture–word stimuli. We measured both latency and amplitude of these difference ERP waveforms in a pre-articulatory time-window. The behavioral results showed standard interference effects in the semantic condition, and facilitatory effects in the phonological condition. The ERP results indicated a bimodal distribution of semantic effects, characterized by the extremely rapid onset (at about 100 ms) of a primary component followed by a later, distinct, component. Phonological effects in ERPs were characterized by components with later onsets and distinct scalp topography of ERP sources relative to semantic ERP components. Regression analyses revealed a covariation between semantic and phonological behavioral effect sizes and ERP component amplitudes, and no covariation between the behavioral effects and ERP component latency. The early effect of semantic distractors is thought to reflect very fast access to semantic representations from picture stimuli modulating on-going orthographic processing of distractor words.

Keywords: event-related potentials, semantic processing, phonological processing

INTRODUCTION

There is a general consensus today that picture naming is semantically mediated, in that the phonological representation of the object name is recovered from a structural representation of the object via semantics (e.g., Indefrey and Levelt, 2004; Bi et al., 2009; Chauncey et al., 2009; Mulatti et al., 2010). In line with this view, the few studies to have used event-related potentials (ERPs) to study language production have provided evidence for semantic influences intervening before phonological influences (van Turenhout et al., 1997; Schmitt et al., 2000; Rodriguez-Fornells et al., 2002). Current estimates of the timing of component processes in picture naming converge on a time-window of 150–200 ms for semantic activation, followed about 150 ms later by phonological word-form activation (see Indefrey and Levelt, 2004, for review). In line with these estimates, there is evidence from recent ERP studies that activation of whole-word phonological representations can be initiated as early as 180–200 ms post-picture onset (Costa et al., 2009; Strijkers et al., 2010). This estimated timing of semantic access from picture stimuli dovetails nicely with estimates derived from studies where participants have to detect the presence of objects belonging to a particular semantic category in a complex scene (e.g., Thorpe et al., 1996; Kirchner and Thorpe, 2006).

A large part of the insights into the nature and timecourse of the different processing stages underpinning language production have been provided through chronometric studies employing the

picture–word interference paradigm (see Mahon et al., 2007, for a comprehensive review). Two findings are pervasively associated with the use of such paradigm, semantic interference (i.e., longer picture naming time when a semantically related distractor word is superimposed on the picture relative to when the word is semantically unrelated), and phonological facilitation (i.e., shorter picture naming time when a phonologically related distractor word is superimposed on the picture relative to when the word is phonologically unrelated). The dominant interpretation of semantic interference in the picture–word interference paradigm is that the distractor word provides bottom-up support for a competing lexical representation that is activated top-down from semantic representations that are activated by the picture stimulus. For example, on seeing a picture of a truck, not only is the lexical representation for “truck” activated, but also semantically compatible words such as “car.” Presenting the word “car” as a distractor increases the activity of its lexical representation, hence increasing its ability to interfere with the production of “truck.” According to this account, the earliest possible influence of a semantically related distractor word during picture naming is during the process of lexical selection. Current estimates locate this process at around 200 ms post-picture onset (e.g., Costa et al., 2009). The dominant interpretation of phonological facilitation effects, on the other hand, locates them at the level of phonological segments that are activated subsequently to

phonological word-forms. There is, however, evidence that both semantic interference and phonological facilitation operate at the level of whole-word phonological representations (Starreveld and La Heij, 1995, 1996). In this case, the phonological word-form corresponding to the picture name would be activated by the phonologically related distractor word.

To further explore the timecourse of semantic and phonological effects in the picture–word interference paradigm, in the present study we manipulated the semantic and phonological relationship between picture names and distractor words while concurrently recording ERPs time-locked to the onset of picture–word stimuli. An overview of prior attempts in the same direction and a critical revisitation of past results from a subset of particularly recent studies hinging on an analogous rationale are reported in Section “Discussion.” Distractor words could be from the same semantic category as the pictures, could share their initial phonemes with the picture name, or had no relationship with the pictures. Subjects were exposed, as is typical in the most classical variant of the picture–word interference paradigm, to the picture of an object that had one word placed more or less at its center, under the requirement to name the picture and disregard the word. The pictures used in the semantic and phonological conditions were the same, and the different sets of words paired with the pictures were carefully matched for a number of physical and lexical attributes. This allowed us to isolate an unequivocal ERP reflection of the picture/word relationship manipulation by subtracting the raw ERP response to one picture when presented with a given word from the raw ERP response to the same picture when presented with another (carefully balanced) word. We expected to replicate the standard behavioral finding of semantic interference and phonological facilitation relative to unrelated distractors. Most important, however, the ERP methodology devised in the present context was optimized to explore the timing at which semantic and phonological information becomes available and modulates the process of picture naming. Furthermore, correlations between these ERP indexes of semantic and phonological activation and behavioral measures of picture naming speed were carried out to infer differences and similarities of the cerebral sources of semantic and phonological information modulating picture naming.

MATERIALS AND METHODS

PARTICIPANTS

Twenty-seven students at the University of Padova, 14 males, with a mean age of 24 years, participated in the experiment for course credit. All had normal or corrected-to-normal vision, and gave their informed consent prior to participation. Thirteen of them participated in the semantic condition, the others in the phonological condition.

APPARATUS AND STIMULI

Forty-eight line drawings of real-world concepts with a high name agreement ($H = 0.01$) were selected from Dell'Acqua et al.'s (2000) database. In the semantic condition, each picture was paired with a same-category related word and an unrelated word. In the phonological condition, each picture was paired with a phonologically related word and an unrelated word. The phonologically related words shared with the picture names the initial two

to three phonemes. The phonologically unrelated words shared a maximum of 1 phoneme with the picture names, which was never in initial or final position. No distractor word was part of the response (picture names) set. Examples of the stimuli used in the present experiment are illustrated in **Figure 1**. The four word sets were matched for number of letters, number of syllables, lexical frequency, concept familiarity, concept typicality, and age of acquisition ($ts < 1$). Equiluminant pictures and words were displayed in white (45 cd/m²) against the black background (8 cd/m²) of a 19" cathode-ray tube monitor, controlled by an IBM-clone and MEL software. The same apparatus was used to control vocal response recordings, using a high-impedance microphone placed before the participant's mouth and connected to the IBM-clone via a response-box provided by Psychology Software Tools®. Words were displayed at the center of the monitor in Romantri 32 font, surrounded by the pictures that could all be inscribed in a 6° × 6° (visual angle) area of the monitor.

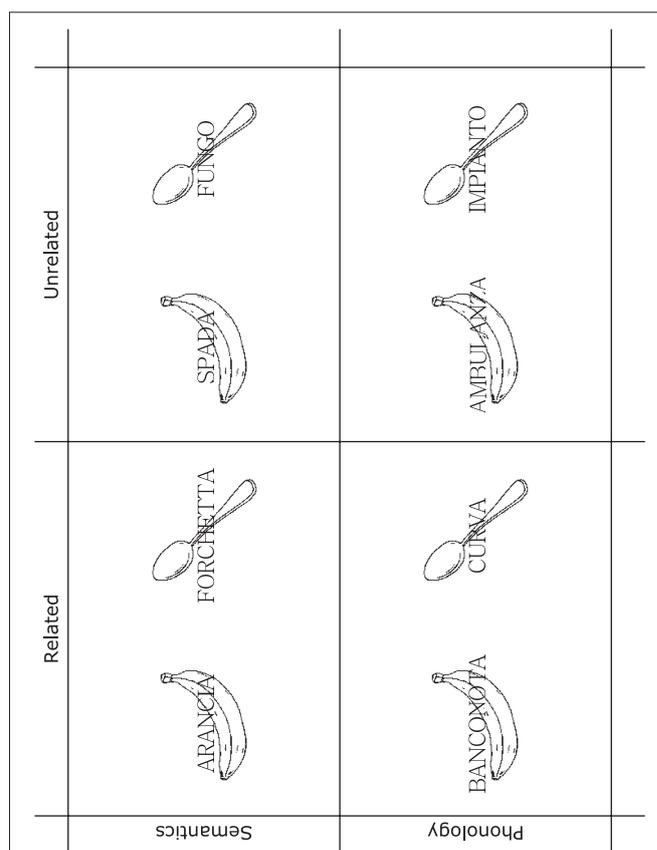


FIGURE 1 | Examples of the stimuli used in the experiment.

Semantics, from left to right (Italian phonology in parentheses): the picture of a banana (/ba'nana/) and a spoon (/kuk'kjaio/) paired with the distractor words orange (/a'ranʃa/), fork (/for'ketta/), sword (/ʃpada/), and mushroom (/ʃungo/). Phonology, from left to right: the same pictures paired with the distractor words banknote (/banʒo'nota/), curve (/ʃkurva/), ambulance (/ambu'lantsa/), and installation (/im'pjanto/). Picture and words in the figure are approximately to scale with those displayed on the monitor during the experiment. Words are reported in the figure in Roman font (rather than Romantri font used in the experiment, not available for our graphical software), which is just slightly less serifed than Romantri.

DESIGN AND PROCEDURE

Each trial began with the presentation of a fixation point at the center of the monitor for 1000 ms. The offset of the fixation point was followed by a blank interval of 800 ms, and by the presentation of a picture–word stimulus. Participants were instructed to name the picture as fast and accurately as possible while ignoring the word. The acoustic onset of the vocal responses was detected through a microphone placed before the participant's mouth. The experimental list of 96 stimuli (i.e., 48 pictures \times 2 words) was repeated 3 times for each participant, and organized at run-time in 6 blocks of 48 trials, that were preceded by a block of 24 practice trials with stimuli that were not included in the experimental list. In each block, picture–word relatedness levels were randomly intermixed and equiprobable. Pictures were not repeated in the same block.

ERP RECORDING AND ANALYSIS

Electroencephalography (EEG) activity was recorded continuously using Electro-Cap[®] head-cap with tin electrodes from 19 standard sites and the right earlobe, and referenced to the left earlobe. Horizontal EOG (HEOG) was recorded bipolarly from electrodes placed at the outer canthi of both eyes. Vertical EOG (VEOG) was recorded bipolarly from two electrodes, above/below the left eye. EEG, HEOG, and VEOG signals were amplified using a Brain-Amp apparatus (Brain Products[®]), filtered using a bandpass of 0.01–80 Hz, and digitized at a sampling rate of 250 Hz. Recording was controlled using Brain-Vision Recorder 1.4 (Brain Products[®]). Impedance at each site was maintained below 5 k Ω . The EEG was re-referenced offline to the average of the left and right earlobes, and segmented into 600 ms (–100 to 500 ms) epochs time-locked to the onset of the picture–word stimuli. Data processing and analysis were carried out using Brain-Vision Analyzer 2.0 (Brain Products[®]). Epochs at each electrode site were baseline corrected using the mean activity during the –100 to 0 ms pre-stimulus period. Trials associated with a HEOG exceeding ± 30 μ V, eye blinks exceeding ± 60 μ V, or any other artifact exceeding ± 80 μ V during the epoch were discarded from analyses (18% of total, evenly distributed among recording channels). Only trials associated with a correct response were considered for ERPs generation. In each phonological/semantic condition, separate ERPs were computed for pictures paired with related and unrelated words, and difference ERPs were generated by subtracting ERPs locked to unrelated picture–word stimuli from ERPs locked to related picture–word stimuli. Mean amplitude and latency of the difference ERPs were explored in two time windows, i.e., 50–200 and 250–450 ms post-stimulus after pooling (i.e., averaging of unweighted values) proximal electrodes to isolate nine regions of interest (ROIs) along the sagittal and the coronal cerebral axes, i.e., left-anterior (pooled Fp1, F3, F7), mid-anterior (Fz), right-anterior (pooled Fp2, F4, F8), left-central (pooled T3, C3), mid-central (Cz), right-central (pooled T4, C4), left-posterior (pooled P3, P7, O1), mid-posterior (Pz), and right-posterior (pooled P4, P8, O2) regions. ERP latencies in these regions were estimated using the jackknife approach (Ulrich and Miller, 2001), by finding the time at which a jackknife waveform reached 30% of the amplitude in each time-window. The Greenhouse–Geisser correction for non-sphericity was applied when appropriate. For excess of artifacts (less than 30% of trials available for ERP generation), data from three subjects were discarded from analyses,

1 in the semantic condition, and 2 in the phonological condition. Following subjects' rejection, data from an equal number of subjects (i.e., 12) were retained in each of the two conditions.

RESULTS

BEHAVIOR

Correct naming times (RTs) were screened for outliers using the procedure described by Van Selst and Jolicoeur (1994) that led to the rejection of 2.2% of the data. The resulting mean RTs and the proportion of correct naming responses were submitted to separate analyses of variance (ANOVAs) considering condition (phonology vs. semantics) as a between-subject factor and relatedness (related vs. unrelated picture–word stimuli) as a within-subject factor. A summary of the results is reported in **Table 1**.

The RT analysis revealed a significant interaction between condition and relatedness ($F(1, 22) = 35.9$, $\eta_p^2 = 0.620$, $p < .001$; other $F_s < 1$). RTs to semantically related stimuli were 39 ms longer and RTs to phonologically related stimuli were 31 ms shorter than RTs to the corresponding unrelated stimuli, which did not differ significantly ($t < 1$). Both these effects (i.e., semantic interference and phonological facilitation) were significant in ANOVAs carried out separately on each condition ($F(1, 11) = 19.9$, $\eta_p^2 = 0.644$, $p < 0.001$; $F(1, 11) = 15.9$, $\eta_p^2 = 0.592$, $p < 0.01$; respectively). The accuracy analysis revealed significant main effects of condition ($F(1, 22) = 11.6$, $\eta_p^2 = 0.346$, $p < 0.001$), relatedness ($F(1, 22) = 28.4$, $\eta_p^2 = 0.563$, $p < 0.001$), and a significant interaction between these factors ($F(1, 22) = 13.6$, $\eta_p^2 = 0.382$, $p < 0.01$). Subjects were equally accurate in related/unrelated phonological and unrelated semantic conditions ($F < 1$), and less accurate in the related semantic condition.

ERP: 50–200 MS TIME-WINDOW

Separate illustrations of the unsubtracted ERPs in the phonological and semantic conditions in the nine ROI (see “ERP Recording and Analysis”) are reported in **Figures 2 and 3**.

The general structure of the ERPs shown in **Figures 2 and 3** appears to be characterized at the anterior and central ROIs by the presence of the visual N1 anterior subcomponent. The large deflection toward positivity was likely a P2 component which was followed by a N400-like negative-going wave. At posterior ROIs the P1 wave was followed by a posterior N1 subcomponent (which typically peaks later than the anterior N1 subcomponent). The large deflection toward positivity was likely a linear combination of overlapping P2 and P3 components. To ascertain that ERPs elicited by unrelated picture–word stimuli did not differ significantly between the semantic and phonological conditions in the present time-window, an ANOVA was conducted on the amplitude of the unrelated ERPs considering ROI as a within-subject factor and condition (phonology vs. semantics) as a between-subject factor. The analysis revealed a significant main effect of ROI ($F(1,$

Table 1 | Mean naming latencies in ms (and percentage of errors) in the phonological and semantic conditions, as a function of the picture–word relationship.

	Phonology	Semantics
Related	843 (2.6)	910 (6.6)
Unrelated	874 (2.9)	871 (3.0)

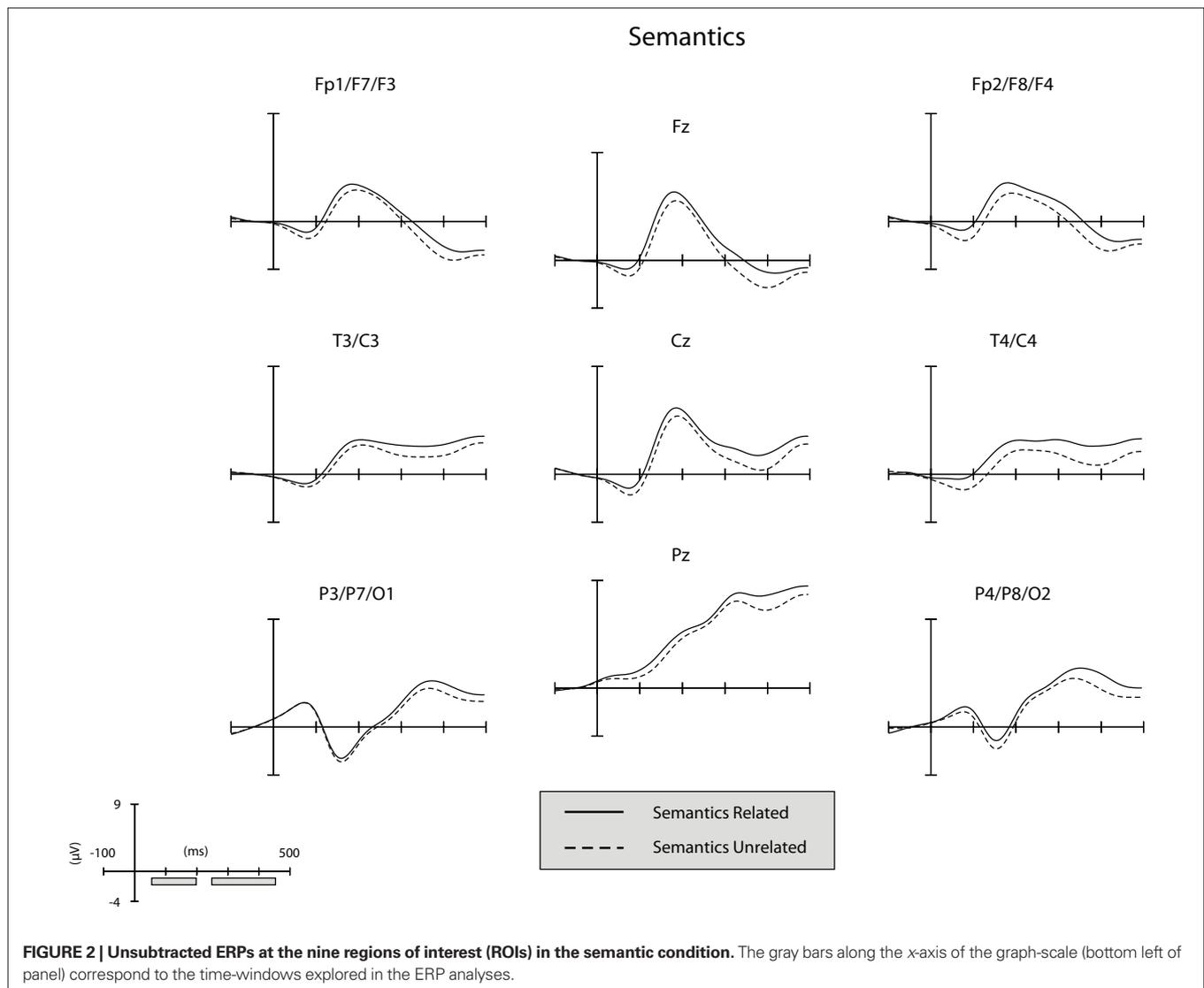


FIGURE 2 | Unsubtracted ERPs at the nine regions of interest (ROIs) in the semantic condition. The gray bars along the x-axis of the graph-scale (bottom left of panel) correspond to the time-windows explored in the ERP analyses.

8) = 26.5, $\eta_p^2 = 0.946$, $p < .001$). Neither the main effect of condition nor the interaction between ROI and condition produced significant modulatory effects on the amplitude of the ERPs elicited by unrelated picture–word stimuli (both $F_s < 1$; $\eta_p^2 < 0.045$; $p > 0.8$).

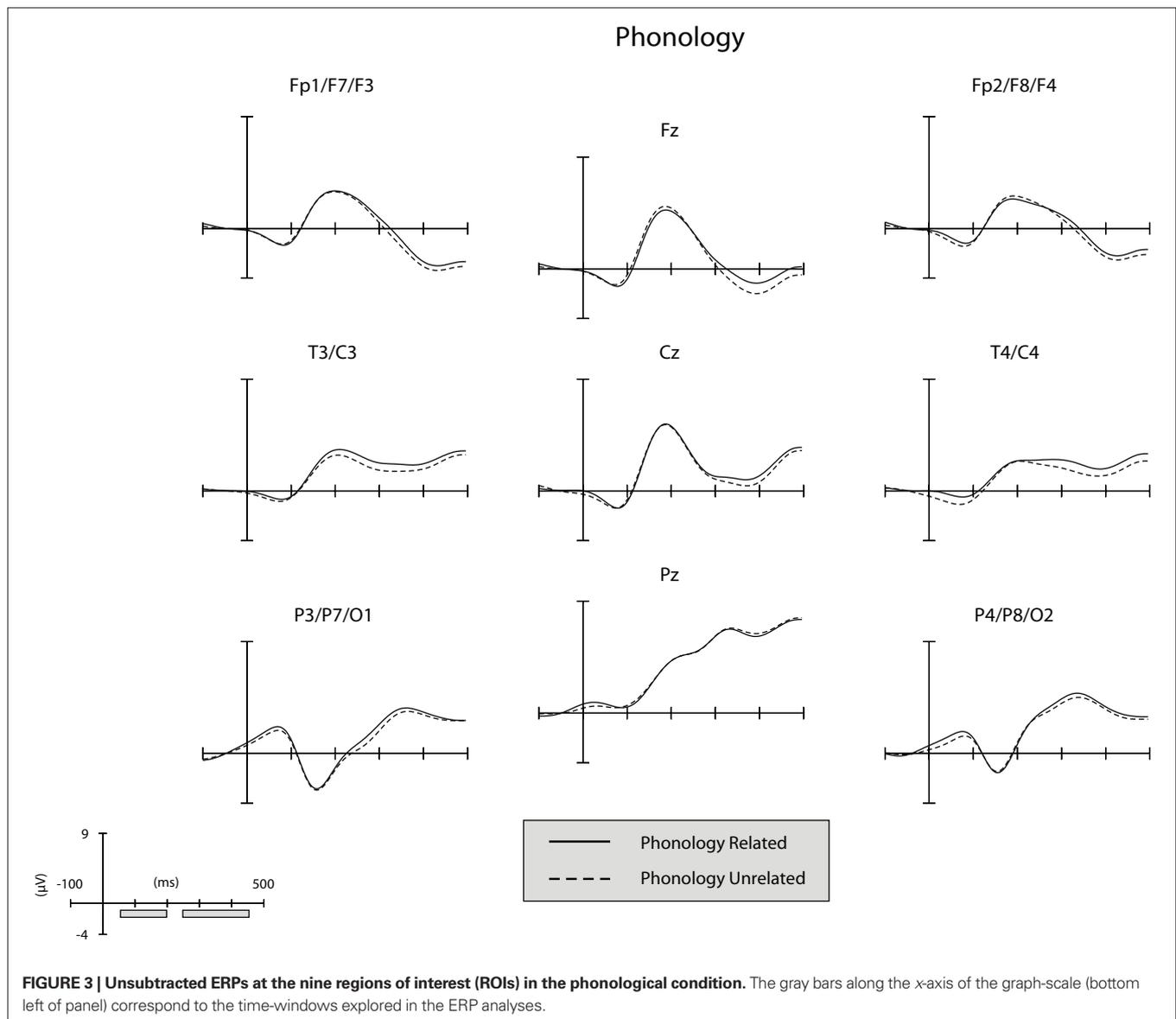
An illustration of difference (related minus unrelated) ERPs for the nine ROIs and in each condition of the present experiment is reported in **Figure 4**. **Figure 4** includes also a graphical representation of baseline/control ERP activity generated by subtracting ERPs elicited by unrelated picture–word stimuli in the phonological and semantic conditions.

An ANOVA on the amplitude values of the difference ERPs, considering condition as a between-subject factor and ROI as a within-subject factor, showed a significant effect of ROI ($F(1, 8) = 6.2$, $\eta_p^2 = 0.767$, $p < 0.001$), and a significant interaction between condition and ROI ($F(1, 8) = 2.8$, $\eta_p^2 = 0.596$, $p < .05$). A series of Bonferroni-corrected t -tests indicated a difference ERP amplitude in the semantic condition that was significantly non-nil at all regions (min $t(1, 11) = 2.6$; max $p = 0.027$; min $\eta_p^2 = 0.381$) except at the left-posterior region (P3/P7/O1; $t(1, 11) = 1.4$, $p = 0.2$; $\eta_p^2 = 0.151$), and a difference

ERP amplitude in the phonological condition that was statistically nil at all regions with the marginally significant exception of the right-central region (T4/C4; $t(1, 11) = 2.0$, $p = 0.065$; $\eta_p^2 = 0.267$).

A step-wise regression analysis considering the RT size of the semantic effect (RT to related minus RT to unrelated picture–word stimuli) as dependent variable and non-nil difference ERP amplitudes in the semantic condition as predictors revealed that 72% (corrected R^2) of the RT variance could be accounted for by the linear combination of the ERP amplitudes recorded at the left-frontal (Fp1/F7/F3) and left-central (T3/C3) sites ($F(1, 11) = 15.2$, $p < 0.02$; $\eta_p^2 = 0.580$)¹. There was no correlation between

¹The present ROI-as-regressor topographical approach was adopted to pinpoint recording channels characterized by an inter-individual distribution of values that correlates with inter-individual distributions of semantic interference and phonological facilitation effects observed on behavior. It must be noted that this approach is far more informative *vis-à-vis* a standard topographical analysis aimed at establishing an electrophysiological (topographical) correlate of behavioral effects. The present approach allows to explore and isolate *parametric* links between the distribution of electrical values recorded at scalp and behavioral effects' estimates, independently on their respective absolute values. Only rarely is such solution attained with more standard approaches.



the difference ERP amplitude recorded at T4/C4 in the phonological condition and the RT size of the phonological effect (RT to unrelated minus RT to related picture–word stimuli; $r = 0.02$, $t < 1$). The jackknife estimate of the latency of the difference ERP in the semantic condition was 106 ms. A regression analysis analogous to that carried out on amplitude values was conducted using individual (not jackknife) latency estimates of the semantic ERP difference as predictors of the RT size of the semantic effect. The regression indicated no significant covariation between these variables ($F_s < 1$).

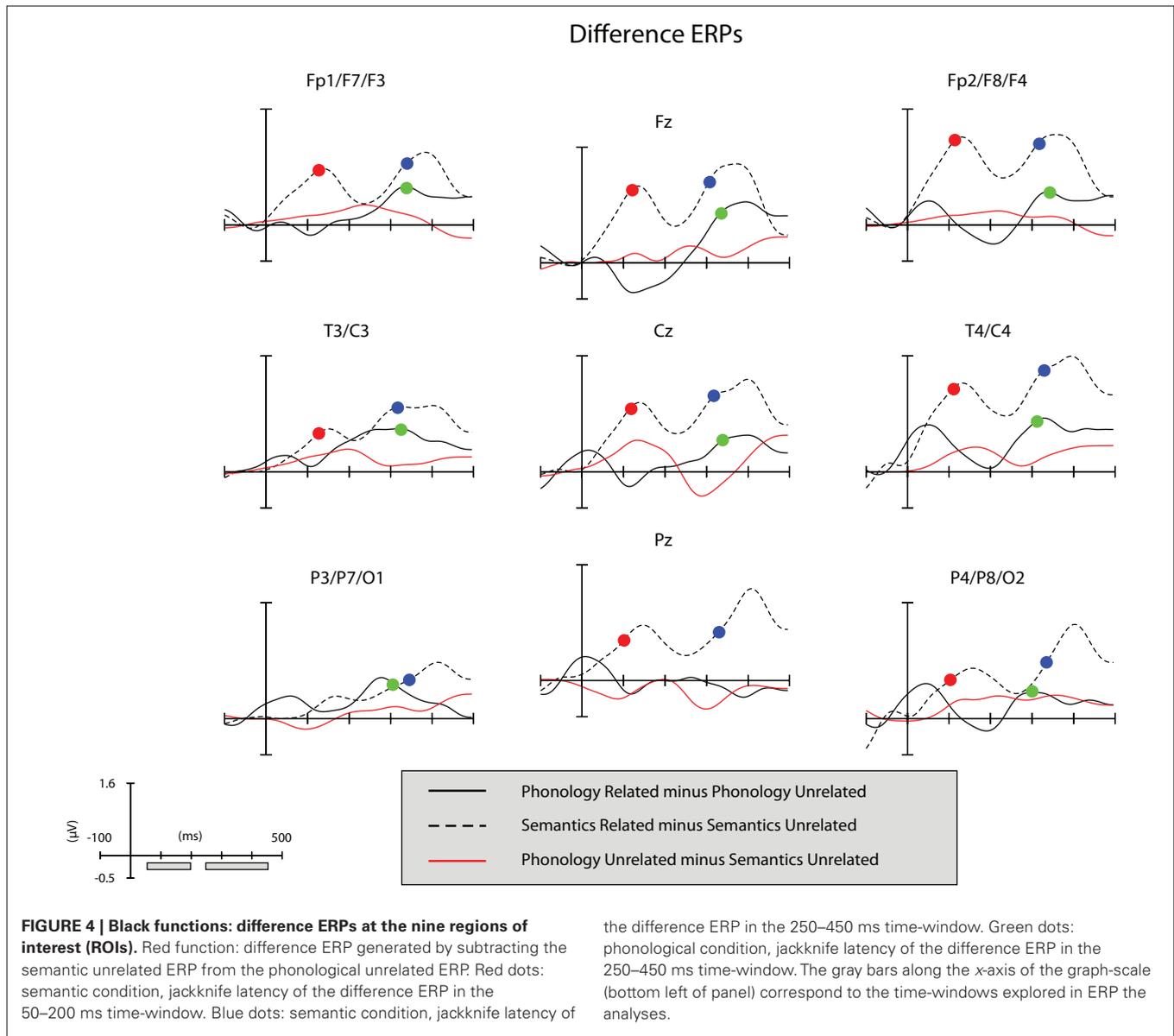
ERP: 250–450 MS TIME-WINDOW

An ANOVA was conducted on the amplitude of the unrelated ERPs considering ROI as a within-subject factor and condition (phonology vs. semantics) as a between-subject factor. The analysis revealed a significant main effect of ROI ($F(1, 8) = 52.2$, $\eta_p^2 = 0.972$, $p < 0.001$). Neither the main effect of condition nor the interac-

tion between ROI and condition produced significant modulatory effects on the amplitude of the ERPs elicited by unrelated picture–word stimuli (both $F_s < 1$; $\eta_p^2 < 0.040$; $p > 0.7$).

An ANOVA on the amplitude values of the difference ERPs, considering condition as a between-subject factor and ROI as a within-subject factor, showed no significant factor effects ($F_s < 1$). Bonferroni-corrected t -tests indicated a difference ERP amplitude significantly different from 0 at all regions in both the semantic condition (min $t(1, 11) = 2.4$; max $p = 0.041$; min $\eta_p^2 = 0.344$) and in the phonological condition (min $t(1, 11) = 2.7$; max $p = 0.033$; min $\eta_p^2 = 0.399$) with the exception, in this latter condition, of mid-posterior region (Pz; $t(1, 11) = 0.35$, $p = 0.78$; $\eta_p^2 = 0.011$).

A regression analysis considering the RT size of the semantic effect as dependent variable and non-nil difference ERP amplitudes in the semantic condition as predictors revealed that 82% of the RT variance could be accounted for by the linear combination of the difference ERP amplitudes recorded at the left-frontal (Fp1/F7/F3) and left-central



(T3/C3) sites ($F(1, 11) = 25.2, p < 0.001; \eta_p^2 = 0.700$). A distinct regression analysis considering the RT size of the phonological effect and the non-nil difference ERP amplitudes in the phonological condition as predictors revealed that 69% of RT variance could be accounted for by the linear combination of the difference ERP amplitudes recorded at the left-frontal (Fp1/F7/F3) and right-posterior (P4/P8/O2) sites ($F(1, 11) = 12.1, p < 0.005; \eta_p^2 = 0.524$). The jackknife estimates of the latency of the difference ERP in the semantic and phonological conditions were 320 and 321 ms, respectively. An ANOVA on the latency values considering condition as a between-subject variable and ROI as a within-subject variable showed no significant effects ($F_{s_{corrected}} < 1$; Ulrich and Miller, 2001). Separate regression analyses using individual latency estimates of difference ERPs as predictors of the RT size of the semantic and phonological effects showed no significant covariation between these variables ($F_s < 1$).

DISCUSSION

The behavioral results showed that the different classes of distractor words used in the present context exerted modulatory effects on picture naming in the expected direction. Relative to unrelated words, semantically related words prolonged picture naming times and phonologically related words shortened picture naming times. The electrophysiological results in the semantic condition showed a bimodal distribution of scalp activity, with a surge in the first 50 ms post-stimulus, a decrease after 200 ms, and a full recovery after 250 ms. The two semantic components had latencies of 106 and 320 ms, respectively. A later and gradual increase of scalp activity was observed in the phonological condition, whose latency was comparable to the latency of the later semantic component, i.e., 321 ms. Overall, ERP activity was widely distributed on the scalp in both the semantic and phonological

conditions. However, regression analyses on ERP amplitude differences revealed a different topography of the regions where ERP activity was most tightly interconnected with facilitatory and inhibitory effects on naming responses in the two conditions. Consistently with prior evidence, the scalp distribution of semantic effects was entirely confined to temporal and frontal regions of the left hemisphere (e.g., Indefrey and Levelt, 2004). The scalp distribution of phonological effects overlapped at left-frontal regions with the semantic effects, but extended in addition to an occipito-temporo-parietal region of the right hemisphere (e.g., Rugg, 1984).

The most important findings of the present study are those related to the timecourse of the observed effects at the electrophysiological level. One key finding is the extremely rapid onset of the early semantic component, which deserves some attention before discussing the later semantic and phonological components. Here we entertain four Scenarios for interpreting the early semantic effect (see Figure 5).

SCENARIO 1

Form only account. Effects of semantic relatedness are generated by the distractor word providing bottom-up support for a whole-word orthographic representation that is partially activated by an approximate visual analysis of the picture stimulus (e.g., the picture of a “truck” activates the lexical representation for “car”

as well as “truck;” cfr. Bar et al.’s, 2006, model of object identification). Therefore, the presence of a relatedness effect implies that whole-word orthographic representations are activated by both the picture and the distractor word, without necessarily involving prior activation of semantics.

SCENARIO 2

Semantics only. Effects of semantic relatedness are generated by interactions between semantic representations activated by the picture and distractor word. The presence of a relatedness effect implies that both the picture and the distractor word have already activated semantic representations.

SCENARIO 3

Semantics-to-phonology. Effects of semantic relatedness are generated by the distractor word providing bottom-up support for a lexical representation (other than the picture name) that is partially activated during the process of lexical selection (activation of whole-word phonological forms from semantics). Therefore, the presence of a relatedness effect implies that the process of lexical selection has already been initiated, and a phonological word form has been activated by the distractor word. This also implies that semantic representations have been activated by the picture, but does not necessarily imply that the distractor word has already activated semantic representations.

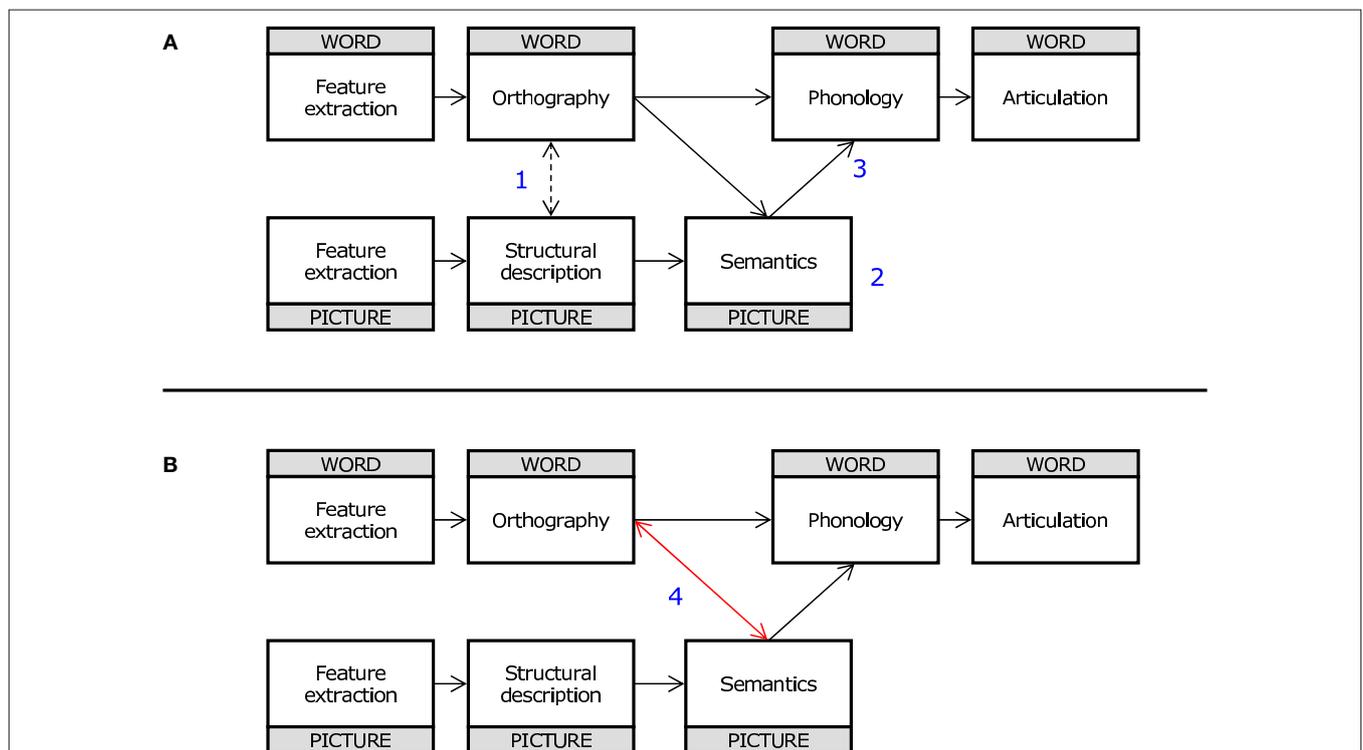


FIGURE 5 | (A) Functional architecture subtending picture and word naming proposed by Chauncey et al. (2009; Figure 4). Blue numbers in the panel refer to the distinct interpretative Scenarios examined in Section “Discussion” (see text for details). **(B)** The same functional architecture reported in **(A)** modified with the

addition of the red bidirectional arrow connecting semantics and whole-word orthographic processing (Scenario 4) providing the best conceptual fit with the present set of ERP results. The blue number in the panel refers to one of the Scenarios examined in Section “Discussion” (see text for details).

SCENARIO 4

Semantics-to-orthography. Effects of semantic relatedness are generated by the distractor word providing bottom-up support for a whole-word orthographic representation that also receives top-down support from semantic representations activated by the picture.

Given that there is no independent evidence that object representations can directly activate whole-word representations (dashed connection in **Figure 5**) without passing via semantics, we can tentatively rule out Scenario 1. Furthermore, current estimates of the timing of semantic access from words are against Scenario 2. According to this proposal, distractor words should also have accessed semantic representations before 100 ms. Although current estimates of the time it takes to access semantic representations from a printed word vary considerably (see Grainger and Holcomb, 2009, for discussion), the fastest estimates are still beyond 100 ms. One recent result situates the earliest semantic effects in the N2 range (Dell'Acqua et al., 2007), onsetting around 180–200 ms post-stimulus onset (see also Pulvermüller, 2001; Hauk et al., 2006). The timing of the early semantic effect allows us to rule out Scenario 3, according to which the process of lexical selection must be already initiated in order to observe the effect. As noted in the introduction, current estimates of the initiation of lexical selection in picture naming point to 180 ms post-picture onset as the earliest moment in time. This implies that the early semantic effect must be located either at the level of semantic representations or orthographic representations. Since we have already ruled out Scenario 2 (semantics), Scenario 4 (orthography) therefore remains the most viable option.

According to Scenario 4, ultra-fast access to semantic representations from picture stimuli initiates feedback processes from semantics to orthographic representations that are involved in on-going processing of the distractor word. This proposal is in line with the evidence showing faster access to semantics from pictures compared with words. According to this proposal, the early semantic effect has little to do with picture naming, and should in principle be observed in a task that does not require a naming response. So why does it correlate with naming latencies? The early semantic effect would simply be the orthographic equivalent of the standard semantic interference effect that is located at the level of whole-word phonological representations (lexical selection). It is therefore not surprising that its amplitude correlates with the size of the semantic interference seen in picture naming latencies.

Reconciling the present finding supporting ultra-fast semantic access – novel *per se* and never shown before – with past analogous attempts using EEG where such effects were not found in ERPs is arduous, in light of the numerous differences between the present design and those used in prior work. Words and pictures in the present context were displayed unimodally (visually), at a stimulus onset asynchrony (SOA) equal to 0 ms (synchronously), under the requirement of naming pictures, and using to-be-ignored distractors that were not part of the response (pictures' name) set (on this particular point, see Caramazza and Costa, 2000). Each one or any combination of these features may have had a non-neutral influence on the capability of the present design to detect very early semantic effects. The seminal work of

Greenham et al. (2000) apart, in which semantic associate and not coordinate words were used as distractors, we note that a subset of prior studies used a cross-modal presentation of the stimuli, generally displaying words auditorily and pictures visually, sometimes at SOAs different from 0 ms, and including distractor words in the response set (e.g., Jescheniak et al., 2002; Aristei et al., 2010). Behavioral evidence suggesting that semantic interference effects observed under uni-modal and cross-modal conditions may originate from distinct sources has been reported in a seminal study by Damian and Martin (1999). Furthermore, displaying close-to-concurrent stimuli cross-modally may be problematic considering the likely engagement of multi-purpose, amodal control mechanisms that are not engaged when stimuli are presented unimodally (Jolicœur and Dell'Acqua, 1998), even when one of the stimuli displayed in either modality can be ignored (Eimer and Van Velzen, 2002). On the other hand, the use of non-0 SOAs when displaying two (or more) stimuli may trigger processing dynamics which are potentially inflated by attentional perturbations known as attentional blink phenomena. When stimuli are not synchronous and the first stimulus is particularly salient (e.g., as is the case when the onset of a uni-modal or cross-modal distractor is associated with an abrupt onset preceding a to-be-responded-to stimulus), part of the processing of the second stimulus is normally postponed (Dell'Acqua et al., 2009), even when the first stimulus is a word that must be ignored (Stein et al., 2010). Under these circumstances, the possibility that subjects attend to and actively process that first stimulus cannot be excluded *a priori*, with detrimental consequences on the bandwidth and strength of semantically modulated signals (Vachon and Jolicœur, 2010). Finally, the task requirement of vocalizing picture names when examining processing in the picture–word interference paradigm using ERPs is fundamental in order to maximize the chances of engaging lexicalization subroutines. Differently, as when manual responses are used to classify pictures in some form (e.g., Xiao et al., 2010), hypotheses concerning interactions between lexical codes (or worse, the timecourse of their expected interaction) may be just lame arguments. In fact, when all these problematic aspects are overcome, semantically mediated ERP evidence with similar temporal characteristics to that described herein emerges (Hirschfeld et al., 2008), albeit of a nature that is perhaps only partially consistent with that proposed to account for the present findings.

This second peak of effects of semantically related distractors coincided temporally with the effects of phonologically related distractors, at about 320 ms post-stimulus onset. Evidence compatible with concomitant activation of semantic and phonological codes during lexicalization has been provided by Starreveld and La Heij (1995, 1996) in a series of experiments manipulating semantic and orthographic relatedness in a picture–word naming design analogous to that used in the present context. In these experiments, semantically related words (e.g., “mare”) hampered picture naming (e.g., “cat”), while orthographically related words (e.g., “cap”) sped up picture naming. Notably, semantic interference was substantially reduced when distractor words were related both orthographically and semantically to picture names (e.g., “calf”; see also, Rayner and Springer, 1986). According to the authors, the interaction between semantic and orthographic

relatedness would be compatible with concurrent generation and interplay of semantic interference and phonological facilitation effects (e.g., Schriefers et al., 1990). Given the timing of our effects at about 320 ms, we propose that both of them, the semantic effect and the phonological effect, reflect processing at the level of lexical selection prior to phonological encoding (see for similar timing effects, Costa et al., 2009). Semantically related distractors would affect processing at this level via the combined activation input from the distractor word and from semantic representations activated by the picture that are compatible with the distractor word. Phonologically related distractors would influence the same level of processing at about the same time but in a different way, by providing extra bottom-up support for the picture name via whole-word phonological representations activated by the distractor word. This proposal dovetails nicely with Bi's et al. (2009) idea that phonological facilitation in the picture–word interference paradigm may well have multiple sources, one of which likely involves phonological segments and/or articulatory codes. This would involve direct connections between a pre-lexical orthographic code and phonological segments or articulatory units (not shown in **Figure 5**), and would explain why the regression analyses revealed an additional source for the phonological effect over and above the source shared with the late semantic effect.

One may find it puzzling that effects of opposite direction at the behavioral level of analysis (i.e., interference in the semantic condition and facilitation in the phonological condition) were both associated with ERP components of similar, negative, polarity. Although an in-depth discussion about the complexity of mapping ERP absolute polarities and direction of behavioral effects is beyond the scope of the present work, we appeal here to two examples that are illustrative of the non-linear relationship between the functional/behavioral and ERP domains. Consider first the N2pc component, namely, the enhanced negativity usually recorded at posterior occipito-parietal regions contralateral to the side of presentation of a lateralized to-be-attended-to target stimulus (with the concomitant presentation of a lateralized symmetrical to-be-ignored distractor stimulus). In absolute terms, both contralateral and ipsilateral ERPs are negative shortly after the presentation of the lateralized stimuli. Nonetheless, some (e.g., Hickey et al., 2008) have provided evidence that increased contralateral negativity reflects target activation, whereas the reduced ipsilateral negativity reflects distractor inhibition. Consider in turn a second example, which comes from studies

using multi-stimulus designs like rapid serial visual presentation or psychological refractory period. Generally, ERPs elicited by stimuli trailing to-be-attended-to stimuli unfold as flows of progressively increasing negativity up to 500/600 ms post-stimulus, perhaps owing to the contingent negative variation (CNV) effects often observed when subjects are instructed to expect more than one stimulus to process. Under these circumstances, a systematic manipulation of response frequency sought to generate a “surprise effect” (Donchin, 1981) tends to be reflected in a ERP negativity decrease in a 350–600 ms time-window locked to trailing stimuli. Nonetheless, such decrease in negativity is held to belong to the P3b class of frequency-related ERPs (e.g., Dell'Acqua et al., 2005), showing that a component of established positive polarity may be camouflaged by latent ERP components and/or spurious modulatory factors that may reverse the sign of ERP component voltage. Our inclination is therefore to interpret the outcome of the ROI-as-regressor analysis as a valuable information concerning the different patterns of cerebral circuitry at work when processing phonological and semantic information, and to abstain from making any commitments concerning the specific processing algorithms implemented in these circuitries that were ultimately reflected in opposite effects – interference in the semantic condition and facilitation in the phonological condition – on overt behavior.

In sum, we recorded ERPs in the picture–word interference paradigm in order to plot the timecourse of effects of semantically related and phonologically related distractors. We found evidence for a very early influence of semantically related distractors at around 100 ms post-stimulus onset. This early effect is interpreted as reflecting ultra-fast access to semantic representations from picture stimuli that then provides feedback to on-going orthographic processing of word stimuli. We also found evidence for later semantic and phonological effects arising at around 320 ms post-stimulus onset, and having partially overlapping sources. These effects are taken to reflect activation of whole-word phonological representations that is modulated (albeit in different ways) by semantically related and phonologically related distractors, with the latter having a possible additional influence at the level of phonological segments and/or articulatory codes.

ACKNOWLEDGMENTS

Jonathan Grainger was supported by ERC Advanced Grant #230313. We thank Dr. Patrik Pluchino for coordinating the data collection phase.

REFERENCES

- Aristei, S., Melinger, A., and Abdel Rahman, R. (2010). Electrophysiological chronometry of semantic context effects in language production. *J. Cogn. Neurosci.* (in press).
- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmidt, A. M., Dale, A. M., Hamalainen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., and Halgren, E. (2006). Top-down facilitation of visual recognition. *Proc. Natl. Acad. Sci. U.S.A.* 103, 449–454.
- Bi, Y., Xu, Y., and Caramazza, A. (2009). Orthographic and phonological effects in the picture–word interference paradigm: evidence from a logographic language. *Appl. Psycholinguist.* 30, 637–658.
- Caramazza, A., and Costa, A. (2000). The semantic interference effect in the picture–word interference paradigm: does the response set matter? *Cognition* 75, 51–64.
- Chauncey, K., Holcomb, P. J., and Grainger, J. (2009). Primed picture naming within and across languages: an ERP investigation. *Cogn. Affect. Behav. Neurosci.* 9, 286–303.
- Costa, A., Strijkers, K., Martin, C., and Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21442–21446.
- Damian, M. F., and Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 345–361.
- Dell'Acqua, R., Jolicœur, P., Luria, R., and Pluchino, P. (2009). Reevaluating encoding-capacity limitations as a cause of the attentional blink. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 338–351.
- Dell'Acqua, R., Jolicœur, P., Vespignani, F., and Toffanin, P. (2005). Central processing overlap modulates P3 latency. *Exp. Brain Res.* 165, 54–68.

- Dell'Acqua, R., Lotto, L., and Job, R. (2000). Naming times and standardized norms for the Italian PD/DPSS set of pictures: direct comparisons with American, English, French, and Spanish published databases. *Behav. Res. Methods Instrum. Comput.* 32, 588–615.
- Dell'Acqua, R., Pesciarelli, F., Jolicœur, P., Eimer, M., and Peressotti, F. (2007). The interdependence of spatial attention and lexical access as revealed by early asymmetries in occipit-parietal ERP activity. *Psychophysiology* 44, 436–443.
- Donchin, E. (1981). Surprise!... surprise? *Psychophysiology* 18, 493–513.
- Eimer, M., and Van Velzen, J. (2002). Crossmodal links in spatial attention are mediated by supramodal control processes: evidence from event-related potentials. *Psychophysiology* 39, 437–449.
- Grainger, J., and Holcomb, P. J. (2009). Watching the word go by: on the time-course of component processes in visual word recognition. *Lang. Linguist. Compass* 3, 128–156.
- Greenham, S. L., Stelmack, R. M., and Campbell, K. B. (2000). Effects of attention and semantic relation on event-related potentials in a picture–word naming task. *Biol. Psychol.* 50, 79–104.
- Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., and Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage* 30, 1383–1400.
- Hickey, C., Di Lollo, V., and McDonald, J. J. (2008). Electrophysiological indices of target and distractor processing in visual search. *J. Cogn. Neurosci.* 21, 760–775.
- Hirschfeld, G., Jansma, B., Bölte, J., and Zwitserlood, P. (2008). Interference and facilitation in overt speech production investigated with event-related potentials. *Neuroreport* 19, 1227–1230.
- Indefrey, P., and Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144.
- Jescheniak, J. D., Schriefers, H., Garrett, M. F., and Friederici, A. D. (2002). Exploring the activation of semantic and phonological codes during speech planning with event-related brain potentials. *J. Cogn. Neurosci.* 14, 951–964.
- Jolicœur, P., and Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cogn. Psychol.* 36, 138–202.
- Kirchner, H., and Thorpe, S. (2006). Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. *Vision Res.* 46, 1762–1776.
- Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., and 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. Mem. Cogn.* 33, 503–535.
- Mulatti, C., Lotto, L., Peressotti, F., and Job, R. (2010). Speed of processing explains the picture–word asymmetry in conditional naming. *Psychol. Res.* 74, 71–81.
- Pulvermüller, F. (2001). Brain reflections of words and their meaning. *Trends Cogn. Sci.* 5, 517–524.
- Rayner, K., and Springer, C. J. (1986). Graphemic and semantic similarity effects in the picture–word interference task. *Br. J. Psychol.* 77, 207–222.
- Rodriguez-Fornells, A., Schmitt, B. M., Kutas, M., and Münte, T. F. (2002). Electrophysiological estimates of the time course of semantic and phonological encoding during listening and naming. *Neuropsychologia* 40, 778–787.
- Rugg, M. D. (1984). Event-related potentials and the phonological processing of words and non-words. *Neuropsychologia* 22, 435–443.
- Schmitt, B. M., Münte, T. F., and Kutas, M. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology* 37, 473–484.
- Schriefers, H., Meyer, A. S., and Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: picture–word interference studies. *J. Mem. Lang.* 29, 86–102.
- Starreveld, P. A., and La Heij, W. (1995). Semantic interference, orthographic facilitation and their interaction in naming tasks. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 686–698.
- Starreveld, P. A., and La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 896–918.
- Stein, T., Zwickel, J., Kitzmantel, M., Ritter, J., and Schneider, W. X. (2010). Irrelevant words trigger an attentional blink. *Exp. Psychol.* 57, 301–307.
- Strijkers, K., Costa, A., and Thierry, G. (2010). Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. *Cereb. Cortex* 20, 912–928.
- Thorpe, S., Fize, D., and Marlot, C. (1996). Speed of processing in the human visual system. *Nature* 381, 520–522.
- Ulrich, R., and Miller, J. (2001). Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. *Psychophysiology* 38, 816–827.
- Vachon, F., and Jolicœur, P. (2010). Impaired semantic processing during task-set switching: evidence from the N400 in rapid serial visual presentation. *Psychophysiology* (in press).
- Van Selst, M., and Jolicœur, P. (1994). A solution to the effect of sample size on outlier elimination. *Q. J. Exp. Psychol.* 47A, 631–650.
- van Turennot, M. I., Hagoort, P., and Brown, C. M. (1997). Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *J. Exp. Psychol. Learn. Mem. Cogn.* 23, 787–806.
- Xiao, X., Zhang, Q., Jia, L., Zhang, Y., and Luo, J. (2010). Temporal course of cognitive control in a picture–word interference task. *Neuroreport* 21, 104–107.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

Received: 07 May 2010; paper pending published: 26 May 2010; accepted: 05 October 2010; published online: 27 October 2010.
 Citation: Dell'Acqua R, Sessa P, Peressotti F, Mulatti C, Navarrete E and Grainger J (2010) ERP evidence for ultra-fast semantic processing in the picture–word interference paradigm. *Front. Psychology* 1:177. doi: 10.3389/fpsyg.2010.00177
 This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.
 Copyright: © 2010 Dell'Acqua, Sessa, Peressotti, Mulatti, Navarrete and Grainger. This is an open-access article subject to an exclusive license agreement between the authors and the *Frontiers Research Foundation*, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.