

The Influence of Language on Perception: Listening to Sentences about Faces Affects the Perception of Faces

Ayelet N. Landau,¹ Lisa Aziz-Zadeh,² and Richard B. Ivry¹

¹Psychology Department, University of California, Berkeley, Berkeley, California 94720, and ²Brain and Creativity Institute and Department of Occupational Sciences, University of Southern California, Los Angeles, California 90089

We examined the effect of linguistic comprehension on early perceptual encoding in a series of electrophysiological and behavioral studies on humans. Using the fact that pictures of faces elicit a robust and reliable evoked response that peaks at ~ 170 ms after stimulus onset (N170), we measured the N170 to faces that were preceded by primes that referred to either faces or scenes. When the primes were auditory sentences, the magnitude of the N170 was larger when the face stimuli were preceded by sentences describing faces compared to sentences describing scenes. In contrast, when the primes were visual, the N170 was smaller after visual primes of faces compared to visual primes of scenes. Similar opposing effects of linguistic and visual primes were also observed in a reaction time experiment in which participants judged the gender of faces. These results provide novel evidence of the influence of language on early perceptual processes and suggest a surprising mechanistic description of this interaction: linguistic primes produce content-specific interference on subsequent visual processing. This interference may be a consequence of the natural statistics of language and vision given that linguistic content is generally uncorrelated with the contents of perception.

Introduction

Vision and language provide two primary means to access conceptual knowledge. We can appreciate the completion of a track competition by watching the winner cross the finish line or by listening to the description of a radio announcer. The relationship between nonlinguistic and linguistic forms of representation has long stood as a foundational problem in epistemology and cognitive psychology. One core issue centers on characterizing interactions between these forms and, in particular, how the operation of a phylogenetically younger linguistic system might be scaffolded on more ancient motor and perceptual processes (Gallese and Lakoff, 2005; Feldman, 2006). In the current studies, we use physiological and behavioral measures to investigate interactions between language and the visual perception of faces.

The motivation for this study comes, in part, from investigations of action understanding. This literature has emphasized that the comprehension of observed actions entails reference to our own ability to produce those actions. More recently, neuroimaging studies have shown that processing linguistic phrases about actions also activates premotor regions in an effector-specific manner (Hauk and Pulvermüller, 2004; Hauk et al., 2004; Tettamanti et al., 2005; Aziz-Zadeh et al., 2006, 2008), suggesting shared neural representation for the conceptual knowledge of action.

A similar question can be asked concerning the relationship between language and perception. Numerous studies demonstrate that performance in one task domain is influenced by the conceptual relationship of task-irrelevant information in a different domain (Glenberg and Kaschak, 2002). For example, people are slower to judge that two words are unrelated if the words share perceptual properties (e.g., racket and banjo) compared to when they were perceptually distinct (Zwaan and Yaxley, 2004). Other studies have taken the opposite approach, showing that incidental verbal information modulates perceptual acuity (Meteyard et al., 2007). Such results provide compelling demonstrations of interactions between representations engaged by linguistic and nonlinguistic input. However, the manner and stage of processing at which these interactions occur is unclear. It remains unknown whether linguistic representations exert an influence on processes traditionally associated with early perceptual stages or whether the interactions occur at later processing stages.

In the present work, we look at the influence of language on perception, focusing on face perception. Aziz-Zadeh et al. (2008) reported that activation in the fusiform face area (FFA) was differentially modulated when people listened to sentences about faces compared to places. Surprisingly, FFA activation was weaker after face sentences. The temporal resolution of functional magnetic resonance imaging (fMRI), however, is not sufficient to examine the stage of processing responsible for this modulation, nor were any behavioral measures obtained to assess the functional consequences of this content-specific linguistic modulation. We address these issues in the present study. We use electroencephalography (EEG) to examine the temporal dynamics of the interaction of language and perception. We exploit the fact that an early evoked response, the N170, is elicited after the visual presentation of faces (Bentin et al., 1996; see also, Bötzel

Received April 20, 2010; revised July 19, 2010; accepted Sept. 13, 2010.

We thank Nola Kleffuss, Meriel Owen, and Sarah Holtz for their assistance in data collection.

Correspondence should be addressed to Ayelet N. Landau at her present address: Ernst Strüngmann Institute (ESI) in Cooperation with Max Planck Society, Deutschordenstr. 46, D-60528 Frankfurt, Germany. E-mail: ayelet.landau@gmail.com.

DOI:10.1523/JNEUROSCI.2046-10.2010

Copyright © 2010 the authors 0270-6474/10/3015254-08\$15.00/0

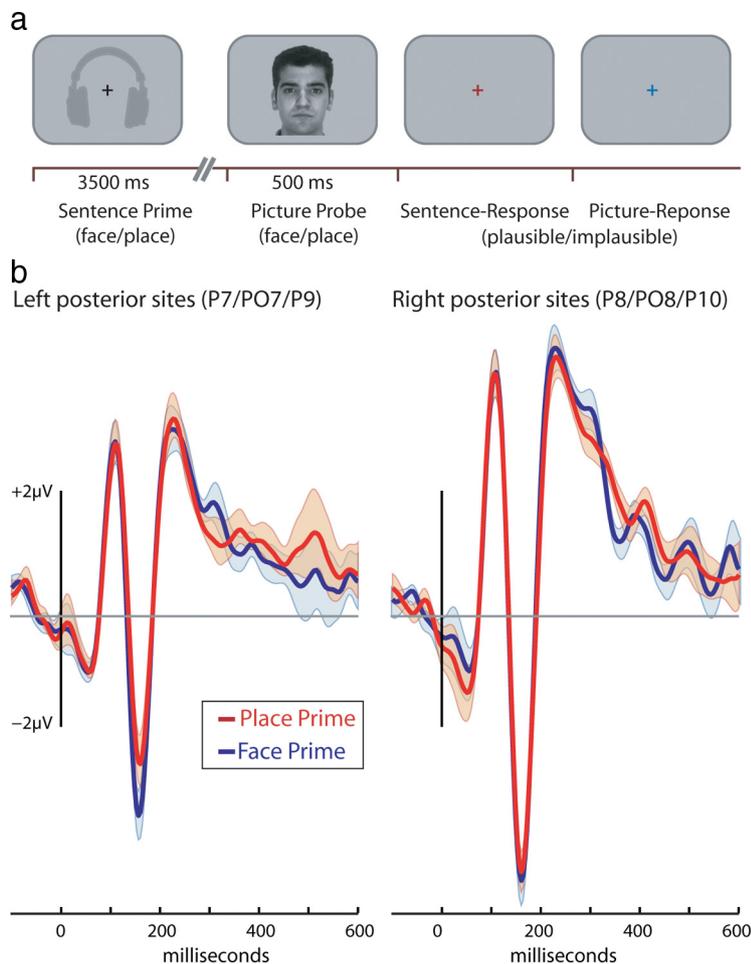


Figure 1. *a*, Timeline of the experimental procedure for sentence priming in experiment 1. Priming sentences, presented through headphones, either referred to faces or scenes. The picture probes were either of a face or a place. Colored fixation crosses (red and blue) prompted each plausibility response. *b*, Event-related response to face picture probes appearing after face sentence primes (red) and place sentence primes (blue). Shaded areas represent the SE of the difference between the two waveforms. Consequently, the same error term is superimposed on both waveforms.

and Grüsser, 1989; Jeffreys, 1989). In the first experiment, we ask whether this visually based response is modulated by the semantic content of linguistic primes. We predicted that the N170 response to faces would be differentially modulated by sentence primes that described faces compared to sentence primes that described scenes. In a second experiment, we replaced the sentence primes with picture primes, providing a contrast between the efficacy of linguistic and nonlinguistic primes on the N170 response. Finally, we conducted a reaction time (RT) experiment to compare performance changes associated with these two types of primes.

Materials and Methods

Participants

A total of 59 participants were recruited from the Berkeley Research Participation Pool. Thirteen participated in experiment 1 (6 females), 19 in experiment 2 (12 females), and 27 in experiment 3 (15 females). All participants had normal or corrected-to-normal vision. All participants were right-handed as determined by the Edinburgh handedness questionnaire (Oldfield, 1971) and were native English speakers. The protocol was approved by the University of California Berkeley Institutional Review Board.

Picture stimuli

The pictures consisted of 88 grayscale images of unfamiliar faces, viewed from the front (used previously by Landau and Bentin, 2008) and 88

grayscale images of scenes (used previously by Aziz-Zadeh et al., 2008). An additional 20 implausible pictures were created from a subset of these images. Of these, 10 were faces in which the internal components (e.g., the nose) were misplaced, and 10 were scenes that included oversized and out-of-context placed objects. Each picture subtended a visual angle of $\sim 5^\circ$.

Sentence primes

Sentence primes were used in experiments 1 and 3. The sentences were used in a previous study looking at hemodynamic changes in the FFA to linguistic stimuli (Aziz-Zadeh et al., 2008). These consisted of 176 auditory sentences, recorded by a native English speaker. Half of the sentences described a facial feature, about either a famous face (e.g., “George Bush has wrinkles around his eyes.”) or a generic face (e.g., “The farmer has freckles on his cheeks.”). The other half of the sentences pertained to scenes, describing either a famous place (e.g., “The Golden Gate Bridge’s towers rise up from the water”) or a generic place (e.g., “The house has a couch near the fireplace.”). In addition to plausible sentences, there were eight sentences in which the semantics constituted an inaccurate or unreasonable statement (e.g., “The farmer’s lashes are ten feet long,” “The park has a lake made of lead.”).

The face and place sentences were matched in terms of number of words and syllables (average number of words, 7.9; average number of syllables, 11; average sentence duration, 2.5 s). Aziz-Zadeh et al. (2008) also did extensive pretesting to match the sentences for understandability and difficulty. During pretesting for that experiment, participants were required to make speeded responses, indicating if the sentence described a plausible or implausible concept. Participants were accurate on $>95\%$ of the trials with similar reaction times across the different conditions (average of 3 s from sentence onset).

Picture primes

Picture primes were used in experiments 2 and 3. In these experiments, 44 of the 88 face stimuli were designated primes, and the other 44 were designated probes. Similarly, 44 of the place stimuli were designated primes, and the other 44 were designated probes. The assignment to these two roles was counterbalanced across participants.

Procedure

The general trial structure was identical in all three experiments (Fig. 1). Each trial started with the presentation of a black cross at the center of the screen. Participants were instructed to maintain fixation on the cross. After 500 ms, a prime was presented. Three seconds after prime onset, a picture appeared at the center of the screen (average interstimulus interval of 500 ms). The picture was presented for 500 ms and was then replaced by a colored cross that served as a response prompt.

Experiment 1. The primes were sentences, played over headphones. The probes were pictures of either a face or scene. Participants were required to make two responses on each trial, the first indicating the plausibility of the prime and the second the plausibility of the probe. For each response, they pressed one of two response keys with their right hand. There were a total of 352 experimental and 64 implausible trials (32 implausible prime and 32 implausible probes). Each sentence prime was presented twice during the course of the experiment, once preceding a picture of a scene and once preceding a picture of a face. Each picture probe appeared twice, once after a prime describing a face and once after a prime describing a place. Repetitions were always separated by a min-

imum of 132 trials. Pairing of the primes and probes was counterbalanced between subjects. Although our primary analysis is focused on the event-related potential (ERP) data, we note here that accuracy on plausibility judgments was very high, with a mean \pm SD prime hit rate of $93 \pm 0.02\%$ and a mean \pm SD probe hit rate of $99 \pm 0.01\%$.

Experiment 2. The same design was used in experiment 2, with the exception that the auditory sentence primes were now replaced by visual picture primes (Fig. 2A). Again, performance on the plausibility judgments was very high, with a mean \pm SD prime hit rate of $93 \pm 0.03\%$ and a mean \pm SD probe hit rate of $96 \pm 0.02\%$.

Experiment 3. Only faces were presented as probes in experiment 3, and, for these stimuli, participants were required to make a gender-discrimination judgment. The 88 face stimuli used in experiments 1 and 2 were coded for gender, and a subset of 80 of these were selected such that half were of male faces and half were of female faces. Participants completed two blocks of 80 trials each. In one block, the face probes were preceded by auditory sentence primes, with half of the primes describing faces and the other half describing scenes. In the other block, the primes were pictures, half of faces and half of scenes. The order of the two blocks was counterbalanced across participants. Within each priming modality (sentences or pictures), a face probe was repeated twice, once after a face-related prime and once after a place-related prime. Separate sets of face probes (40 per set) were used for each prime modality, with the assignment of sets to modality counterbalanced between participants. A given face could not be a prime and probe in the same block. Implausible stimuli were not used in experiment 3.

After the onset of the probe, participants were required to make a speeded response, indicating whether the picture was of a man or woman. Responses were made with the index and middle fingers of the right hand, with the assignment of gender and finger counterbalanced between participants. Participants were not required to respond to the prime stimuli during the RT blocks. However, to encourage processing of the primes, participants were informed that a memory test for the primes would be administered after each block. The memory test consisted of eight trials, four with novel stimuli and four with stimuli that had been used previously during the block as primes. Participants made a button press to indicate whether the stimulus was new or old. Mean recognition performance was $>75\%$.

EEG data acquisition and preprocessing. EEG was recorded in the first two experiments with a Biosemi Active Two system at a sampling rate of 512 Hz. Recordings were obtained from 64 electrode sites using a modified 10–20 system montage. Horizontal electro-oculographic (EOG) signals were recorded at the left and right external canthi, and vertical EOG signals were recorded below the right eye. All scalp electrodes, as well as the EOG electrodes, were referenced offline to the tip of the nose. Preprocessing of the data was done in Brain Vision Analyzer. Trials with eye movement or blinks were removed from the data using an amplitude criterion of $\pm 75 \mu\text{V}$ or lower (16.6% of trials eliminated). The EEG signal was segmented into epochs starting 100 ms before probe onset and continuing until 1000 ms after probe onset.

ERP analysis

Segmented data were averaged separately for each condition. Averaged waveforms were bandpass filtered (0.8–17 Hz) and baseline corrected, using the 100 ms preprobe epoch. For each participant, the peak of the N170 component (Bentin et al., 1996) was determined as a local minimum occurring between 130 and 220 ms. The primary dependent variable was the amplitude of this component, recorded from posterior

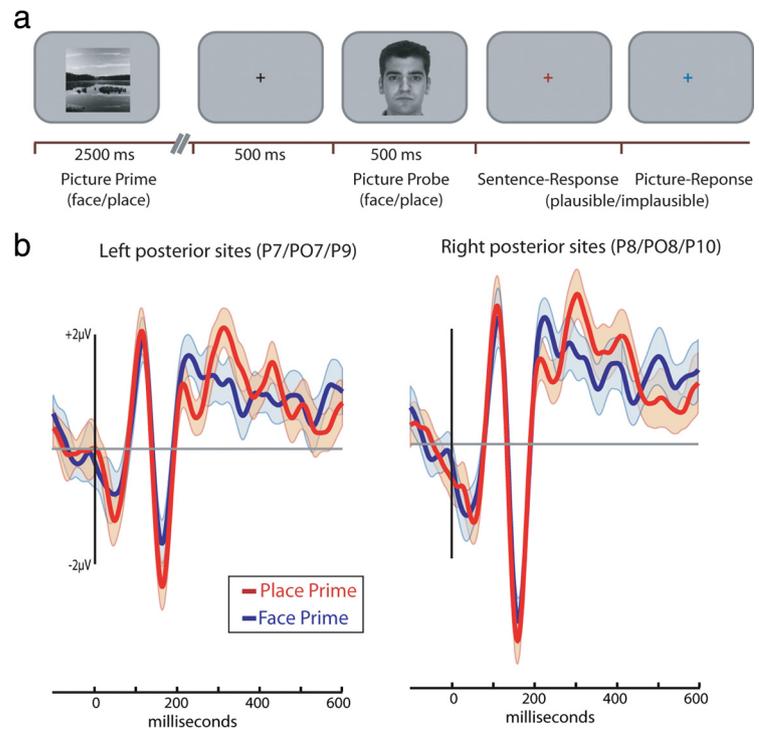


Figure 2. *a*, Timeline of the experimental procedure for picture priming in experiment 2. Primes were either pictures of faces or scenes and were followed by face or place picture probes. *b*, Event-related response to face picture probes appearing after a face picture primes (red) and place picture primes (blue). Shaded areas represent the SE of the difference between the two waveforms as in Figure 1.

electrodes over both hemispheres (P7/8, PO7/8, and P9/10). Because the response to faces was similar over all three electrodes ($F < 1$), the three sites were collapsed for all analyses. The ERP responses to probes following famous and unfamiliar face primes were similar, as were the ERP responses to probes following famous and unfamiliar scene primes ($F < 1$). The data were collapsed over the famous and unfamiliar items for each category in the analyses reported below. Trials with implausible sentences were excluded from the analysis.

Given that we did not include a baseline condition in which the sentence primes were absent, the analyses here focus on the relative effect of the place and face priming sentences on the N170 to the face stimuli. Note that a moderate N170 is also observed, albeit in attenuated form, after the presentation of scenes. However, in the current study, we did not observe a significant peak between 150 and 200 ms to the place probes and do not report these analyses below.

Results

Experiment 1

To examine how language comprehension interacts with perceptual processing, we examined the ERP response to picture probes. If language comprehension engages perceptual substrates in a category-specific manner, then the N170 ERP component elicited by a picture of a face should be differently modulated when preceded by face sentence primes compared to when they are preceded by place sentence primes. Consistent with this prediction, the face-sensitive N170 was modulated by the priming manipulation: the N170 was larger to face probes that followed face-related sentences compared to face probes that followed place-related sentences (Figs. 1B, 3, left). Moreover, this linguistic modulation of the N170 was only observed in electrodes positioned over the posterior aspect of the left hemisphere.

We performed a three-way ANOVA with prime type (face or place sentence), probe type (face or place picture), and hemisphere (left, right) as within-participant factors. The N170 to face

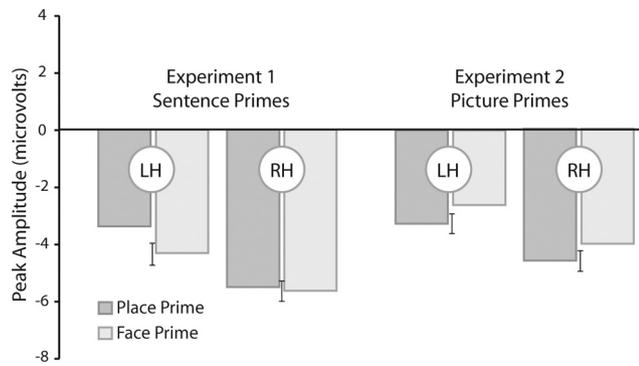


Figure 3. Peak amplitude of the N170 response to face probes in experiment 1 (left side) and experiment 2 (right side). Within each experiment, the data are plotted for the left (LH) and right (RH) hemispheres, as a function of whether the prime referred to a face or place. Error bars correspond to SE of the difference between the two conditions for each type of prime.

probes was larger than that elicited to place probes ($F_{(1,12)} = 38.51, p < 0.0001$). Central to our hypothesis, the two-way interaction of probe type and prime type ($F_{(1,12)} = 9.55, p < 0.01$) and the three-way interaction of these factors with hemisphere were significant ($F_{(1,12)} = 5.25, p < 0.05$). To identify the nature of the three-way interaction, we separately compared for each hemisphere the N170 with face probes after sentence primes about faces or places. For left hemisphere electrodes, the N170 was larger after face sentences compared to place sentences (-4.3 vs $-3.4 \mu\text{V}$; $t_{(12)} = 2.49, p < 0.05$). In contrast, for right hemisphere electrodes, the N170 to faces was of similar magnitude after face and place sentences (-5.56 vs $-5.69 \mu\text{V}$; $t < 1$).

Although our experimental hypotheses were focused on the N170 response, we conducted exploratory analyses over the full montage of electrodes to determine whether other early ERP components (< 200 ms) showed a differential response to the two types of primes. We failed to observe any differences in electrodes that were outside those showing an N170 response. The P1, an ERP component that exhibits early perceptual and attentional processing (Mangun and Hillyard, 1991), has been reported to exhibit face sensitivity (Itier and Taylor, 2004; Thierry et al., 2007) (but see Bentin et al., 2007; Rossion and Jacques, 2008). Given this, we assessed how the P1 amplitude was influenced by the different conditions. The P1 amplitude was larger over the right hemisphere compared to the left hemisphere ($F_{(1,18)} = 20.23, p < 0.01$), and face probes elicited a larger P1 component compared to place probes ($F_{(1,18)} = 30.81, p < 0.001$). However, the P1 response to the probes was not affected by the type of prime (face or place sentence), and two- and three-way interactions were also not reliable. Thus, the modulation elicited by the sentence primes was very specific to those electrodes that showed face selectivity in the N170 ERP response.

These data indicate that an electrophysiological marker associated with the early stages of face perception can be modulated by a preceding linguistic prime. Two features of this modulation are especially noteworthy. First, the modulation was dependent on the semantic relationship between the prime and the probe; thus, the modulation observed here cannot be attributed to a nonspecific effect related to the presentation of auditory stimuli before the probes (i.e., both probe conditions compared in these analyses followed an auditory prime). Rather, the semantic content of the primes led to a content-specific change in the physiological response to the probes. Although the N170 was larger after the face primes, we do not claim that this indicates a stronger neural response to the face probes. We will return to this issue and

possible underlying mechanisms in Discussion; our emphasis at this point is that the N170 was differentially affected by the face and place primes, indicating that the early perceptual process marked by this ERP component is influenced by language.

Second, we observed a laterality effect in terms of the modulatory effects of language. Although the N170 response to face stimuli is typically larger from electrodes positioned over the right hemisphere (Carmel and Bentin, 2002), the semantic modulation in the current study was only present in electrodes over the left hemisphere. This asymmetry is also consistent with the hypothesis that the modulation is language related.

The current results provide a novel demonstration of how linguistic primes can influence subsequent perceptual processing, with this interaction reflecting a high degree of content, or category specificity. These priming effects were observed in a physiological response, the N170, which has traditionally been associated with early stages of perceptual encoding. The early timing, as well as posterior locus of these effects, is consistent with the hypothesis that the linguistic primes modulated activity in neural regions recruited during relatively early stages of face perception.

There are various ways in which this modulatory effect can come about. The influence of linguistic input might be direct. By this view, the regions that generate the N170 would be recruited as part of the processes associated with understanding the semantics of a sentence describing a face (Rossion et al., 2003). Alternatively, the influence may be indirect. For example, comprehension of the linguistic primes might trigger visual images, and it is the difference between the residual activation of facial and nonfacial images that produces the modulation of the N170 response to the subsequent face probes. Indeed, face-sensitive regions in fusiform cortex can be activated when participants visualize faces (O'Craven and Kanwisher, 2000).

We note here that an imagery-based account is at odds with the laterality effects observed in experiment 1. fMRI activation patterns to faces, whether visually presented or imagined, are almost always larger in the right hemisphere compared to the left hemisphere (Puce et al., 1995, 1996; Bentin et al., 1996; Dien, 2009). Similarly, the N170 response to faces is typically larger over right hemisphere electrodes (Bentin et al., 1996; Carmel and Bentin, 2002; Jacques and Rossion, 2007). Consistent with these previous studies, we observed a larger N170 response to face probes over the right hemisphere. However, the category-specific linguistic modulation of the N170 was only evident over left hemisphere electrodes. Given the prominent role of the left hemisphere in linguistic comprehension, these data would suggest that the modulation is likely driven by the activation of linguistic representations rather than indirectly from linguistically triggered visual images. Indeed, if the modulation was attributable to mental imagery evoked by the linguistic primes, then we would expect bilateral or right lateralized modulation effect.

Experiment 2

We conducted a second experiment to directly assess whether the sentences might induce visual images, and it is these images that produce the modulation of the N170 response. To this end, we replaced the linguistic primes with picture primes. We reasoned that if the sentence effects are related to the generation of visual images, then the same pattern of results should be obtained if the primes are visual pictures. Specifically, the imagery account would predict that the picture primes should produce a similar enhancement of the N170 to face probes after picture primes of faces compared to picture primes of scenes.

The picture primes produced category-specific modulation of the N170 to face probes. However, unlike the enhanced N170 response observed in experiment 1 after face sentence primes, the face picture primes led to an attenuation of the N170 response (Figs. 2, 3 right). We evaluated the N170 data with a three-way ANOVA involving the factors hemisphere (left, right), prime type (face, place), and probe type (face, place). As expected, the N170 response to faces was greater than the N170 elicited to places ($F_{(1,18)} = 179.54, p < 0.001$). There was also a significant main effect of hemisphere ($F_{(1,18)} = 4.56, p < 0.05$), with the N170 amplitude greater (i.e., more negative) for electrodes positioned over the right hemisphere compared to the left hemisphere ($F_{(1,18)} = 4.56, p < 0.05$). Unlike experiment 1, the interaction of prime \times probe was not significant in the overall ANOVA ($F_{(1,18)} = 1.8, p = 0.19$).

As can be seen in a comparison of Figures 1–3, the N170 response was lower in experiment 2, perhaps because of the fact that both the prime and probe were visual stimuli. Given that our focus in experiment 2 was on the influence of the primes on the face probes, we conducted a two-way ANOVA limited to these trials with hemisphere (left, right) and prime type (face, place) as within-subject factors. Similar to experiment 1, there was a main effect of hemisphere, with the N170 response larger over the right hemisphere ($F_{(1,18)} = 4.97, p < 0.05$). In addition, the main effect of prime type was reliable ($F_{(1,18)} = 5.67, p < 0.05$). Importantly, the direction of this effect was different than that observed in experiment 1: face probes after picture primes of faces elicited a smaller N170 compared to face probes after picture primes of places. Unlike experiment 1, the prime \times hemisphere interaction was not significant ($F < 1$). The attenuation of the N170 response after face primes was similar from electrodes positioned over both hemispheres. Prime-related modulations were not observed in other early ERP components.

Face probes again elicited a larger P1 amplitude than place probes ($F_{(1,18)} = 13.3, p < 0.01$), and the P1 amplitude was equivalent between the hemispheres ($F_{(1,18)} = 3.06, p = 0.09$). Importantly, unlike the N170, the P1 was not modulated by prime type ($F_{(1,18)} = 0.06, p = 0.8$).

The results obtained in experiment 2 are in line with previous findings reporting suppression of N170 effect attributable to repetition (Jacques and Rossion, 2004, 2007; Kovács et al., 2005). Similarly, when people are asked to actively maintain an image of a previously presented picture of a face, the N170 to a subsequent face probe is reduced (Sreenivasan and Jha, 2007). Together, the results of experiments 1 and 2 argue strongly against an imagery-based interpretation of the priming effects observed with linguistic stimuli in experiment 1. Whereas face-related sentence primes led to an enhancement of the N170 response to subsequent pictures of faces, face-related picture primes led to an attenuation of the N170 to the same probes.

We recognize that there are important methodological differences between the two experiments; in particular, the primes in experiment 2 were visual, whereas they were auditory in experiment 1. However, the modulation of the ERP response within each experiment was measured within modality (i.e., in experiment 1, we measured a modulation of the visual response by different primes from the same auditory modality; in experiment 2, we measured a modulation of the visual response by different primes from the same visual modality). It seems unlikely that a change in prime modality would produce a reversal of the category-specific modulation of the N170 response. That is, compared to place primes, face sentence primes produced an enhancement of the N170 response, whereas face picture primes

produced an attenuation of the N170 response. Moreover, the modulation with linguistic primes was lateralized to the left hemisphere electrodes, whereas the modulation with picture primes was bilateral. As such, it would appear that the modulation of the N170 generators in the left hemisphere from linguistic input is not mediated by visual-based representations.

The attenuation of the N170 response after face picture primes can be viewed as a form of repetition suppression. Indeed, the category-specific effects observed here have been documented previously in both fMRI and EEG experiments (Kovács et al., 2005; Grill-Spector et al., 2006; Harris and Nakayama 2007). Various accounts of the neural mechanisms underlying repetition suppression effects have been proposed. By one account, the suppression reflects more efficient neural processing (for review, see Grill-Spector et al., 2006). Neural responses to the prime alter the dynamics of a network such that it can more readily respond to subsequent inputs that are related. An alternative hypothesis is that suppression reflects neural fatigue. A network that has recently been engaged is consequently less responsive to subsequent stimuli (Grill-Spector et al., 2006). For the present purposes, the most important point to be made is that the repetition suppression effect observed with picture primes was not observed with sentence primes.

Experiment 3

Experiments 1 and 2 reveal that linguistic and picture primes both modulate the N170 response in a content-specific manner but in opposite directions: sentences about faces produce a subsequent enhancement of the N170 to faces, whereas pictures of faces produce a subsequent attenuation of the N170. Given this dissociation, we would expect to observe divergent behavioral consequences of these two types of primes, assuming that the amplitude of the N170 is related to face perception. The behavioral task used in experiments 1 and 2 was not sufficiently sensitive to detect any functional consequences of the processing of primes; performance was near ceiling on the plausibility judgments and speed was not emphasized.

In experiment 3, participants performed a basic categorization task, indicating whether a picture depicted a male or female face (Schyns and Oliva, 1999; Landau and Bentin, 2008). The instructions emphasized response speed, thus providing a more sensitive behavioral assay on the processing consequences of the priming stimuli (Landau et al., 2007; Landau and Bentin, 2008). The presentation of the target stimulus was preceded by one of four types of primes, created by the factorial combination of two variables. First, the primes were either sentences or pictures. Second, the primes were either face related or place related. If the attenuation of the N170 response to face stimuli after visual face primes in experiment 2 reflects fatigue of processes associated with face perception, RTs on the gender discrimination task should be slower after face primes compared to place primes. Alternatively, if the attenuation reflects some form of enhanced neural efficiency, then RTs should be faster after the face primes. Independent of the outcome of these predictions, we predicted that the linguistic and visual primes would produce the opposite pattern of results given that the modulation of the N170 response was reversed for linguistic and visual primes.

All participants but one performed the gender discrimination task very accurately (96% correct). The one participant that was unable to perform sufficiently above chance was removed from additional analysis. In addition, trials with response times longer than 1500 ms and shorter than 100 ms were removed from addi-

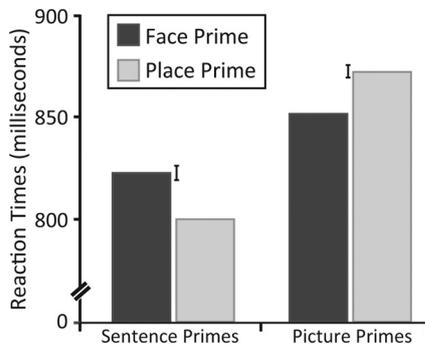


Figure 4. Reaction time data for experiment 3. Participants made a speeded response, classifying a face probe as male or female. These pictures were preceded by sentence (auditory) or picture (visual) primes that either referred to or depicted a face or place. Error bars correspond to SE of the difference between the two conditions for each type of prime.

tional analysis (under 8% of trials were eliminated from additional analyses).

Performance on the gender discrimination task was modulated by the primes (Fig. 4). Interestingly, the priming effect was category specific for both the sentence and picture primes but in opposite directions. In the sentence condition, RTs on the gender task were slower after a prime that described a face compared to when the priming sentence described a place. In the picture condition, RTs on the gender task were faster after picture primes of faces compared to picture primes of scenes. When evaluated with a repeated-measures ANOVA, the interaction of prime modality and prime category was significant ($F_{(1,25)} = 37.60$, $p < 0.001$).

Planned pairwise contrasts were performed, comparing the two prime categories for each prime modality. For the sentence primes, RTs were significantly slower after face primes compared to place primes (23.3 ms difference; $t_{(25)} = 3.97$, $p < 0.01$). For the picture primes, RTs were significantly faster after face primes (20.5 ms difference; $t_{(25)} = -2.82$, $p < 0.01$). The omnibus ANOVA also indicated a main effect of prime modality ($F_{(1,25)} = 9.80$, $p < 0.01$). RTs on the gender discrimination task were slower overall after (visual) picture primes compared to (auditory) sentence primes.

The principal results of experiment 3 converge with the ERP results obtained in experiments 1 and 2: in accord with the opposite effects found for the linguistic and visual primes on the N170 response, we also observed divergent priming effects with our behavioral assay. Most important, these parallel findings were category specific. Sentence primes describing faces that led to an enhancement of the N170 response to face probes produced a cost in reaction time when participants made a perceptual judgment of the same probes. In contrast, picture primes of faces that led to an attenuation of the N170 response to face probes facilitated RTs to the probes.

The specific pattern of the dissociation sheds light on the functional consequences of the physiological changes observed in experiments 1 and 2. RTs were faster after face primes, a finding consistent with the enhanced neural efficiency hypothesis. The recent activation of face processing mechanisms by the visual face primes renders these mechanisms more sensitive to subsequent facial stimuli. In contrast, face-related auditory primes delayed performance on the subsequent face discrimination task compared to place-related auditory primes. This dissociation suggests that the interpretation of the enhanced N170 response after linguistic primes needs to be reevaluated. It would appear that face-

processing mechanisms are disrupted by linguistic inputs that describe faces, an issue we return to below. At this point, we want to emphasize that the current results raise a cautionary note about the interpretation of changes in an ERP response with respect to underlying processes. Superficially, one would expect that an enhanced ERP marker, presumably reflecting a larger neural response, would be associated with more efficient processing (but see Kutas and Hillyard, 1980). The RT data observed here, however, suggest that the enhanced N170 response is associated with less efficient face perception.

Although the experiment was designed such that the key comparisons were within modality, there was a pronounced main effect in that linguistic primes led to faster RTs than picture primes. It is difficult to interpret this main effect given the difference in stimulus modality for the two primes. The participants' attentional set may differ after an auditory prime compared to a visual prime. Related to this, refractory effects to a visual stimulus may differ after a visual prime compared to an auditory prime.

Discussion

The present study was designed to investigate the effects of semantic processing on an electrophysiological marker associated with face processing. We exploited the fact that faces elicit a robust and reliable evoked response, the N170, a negative deflection that peaks at ~170 ms after stimulus onset. We observed that the magnitude of the N170 was modulated in a content-specific manner when the face stimuli were preceded by auditory sentences. Specifically, the amplitude of the N170 to pictures of faces was larger after sentences that described faces compared to sentences that described scenes. In addition to the content specificity of this modulation, the N170 enhancement was only reliable for electrodes positioned over posterior regions of the left hemisphere. This hemispheric asymmetry is especially interesting given that, similar to previous studies, the N170 response recorded from right hemisphere electrodes was larger than that recorded from the left hemisphere. The content specificity and laterality pattern provide a novel demonstration of the interaction of language and perception.

In experiment 2, we considered the hypothesis that this interaction was mediated by visual imagery. Given that the sentence primes preceded the visual probes by 3 s, it would be possible for participants to have generated visual images to the linguistic primes and that the content-specific modulation reflected the recent history of these imagery processes. By this account, the interaction of language and perception would be indirect, with sentence comprehension, at least for concrete concepts, involving the generation of images associated with those concepts. The lingering effects of these (visual) images in turn could lead to the modulatory effects observed when participants were subsequently presented with the visual probe stimuli.

However, when the semantic primes were replaced by visual primes, the direction of the content-specific modulation of the N170 reversed. Face probes elicited a smaller N170 when preceded by a face prime compared to when preceded by a place prime. Moreover, this effect was bilateral, observed in both left hemisphere and right hemisphere electrodes. We recognize that there are important differences between the two types of primes. First, the sentences were auditory, whereas the picture primes were visual. Second, the time course to process the sentence primes may be quite different than that required for the picture primes. Nonetheless, it would be difficult to modify an imagery hypothesis to account for the divergent pattern of results seen for sentence and visual primes, especially when considering both the

modulatory effects on the N170 and the differences in laterality effects.

The behavioral results in experiment 3 also argue against an imagery account and provide converging evidence that linguistic and visual primes produce qualitatively different effects on the perception of subsequent visual probes. Response times on a gender discrimination task were influenced in a content-specific manner by linguistic and visual primes. However, the direction of this effect was again reversed for the two types of primes. With linguistic primes, reaction times were slower after face primes compared to place primes; with visual primes, reaction times were slower after place primes compared to face primes.

The divergent effects of linguistic and visual primes have implications for conceptualizing the interactions of language and perception, as well as for providing a cautionary note with respect to the interpretation of electrophysiological effects. At first blush, it would be tempting to interpret the larger N170 amplitude to faces that follow face-related sentences as evidence of shared representations for language and perception. That is, one might infer that linguistic comprehension of sentences about faces engages neural regions that are recruited when perceiving faces. Applied to the current results, the interpretation would be that the face sentences have primed these regions, producing a stronger response to a subsequent picture of a face. This hypothesis would be consistent with the left lateralized effect observed in experiment 1 given that linguistic primes may evoke stronger representations in the left hemisphere (Bentin et al., 1999; Rossion et al., 2003).

However, this interpretation is challenged by the fact that visual primes in experiment 2 produced opposite effects on the N170 response. Here the N170 response to a face probe was attenuated when preceded by a face-picture prime, a repetition suppression effect. Various mechanisms of repetition suppression effects have been proposed, including ones based on the idea of neural fatigue or improved neural efficiency (for review, see Grill-Spector et al., 2006). Although this study was not designed around the question of mechanisms of repetition suppression, the RT priming data in experiment 3 provides indirect support in favor of the latter hypothesis. Visual primes of faces engage neural regions associated with face perception, leading to more efficient (i.e., faster) processing of faces to subsequent probes.

Applying the same logic to the effects of the linguistic primes leads to the counterintuitive proposal that sentences describing faces actually interfere with the subsequent perception of faces. That is, linguistic processing may involve the disruption or inhibition of early perceptual systems that are associated with conceptually related representations. Similar interference effects have been observed in several behavioral studies (Richardson et al., 2003; Kaschak et al., 2005, 2006). Of particular relevance to the current discussion is the work on verbal overshadowing (Engstler-Schooler, 1990; Fallshore and Schooler, 1995; Schooler and Schooler et al., 1997; Meissner and Brigham, 2001). When people are instructed to verbally describe a face, their ability to subsequently discriminate that face from a set of distracter faces is impaired. Several, nonexclusive hypotheses have been offered to account for this effect. One idea is that verbalization focuses processing on features, whereas face processing is primarily holistic, with a corresponding hemispheric distinction (feature-based processing associated with left hemisphere, holistic-based with right hemisphere). This processing incompatibility creates competition, resulting in verbal interference on face processing (Fallshore and Schooler, 1995; Macrae and Lewis, 2002; Schooler, 2002). Although the overshadowing literature has emphasized interference in terms of memory retrieval, the current results

suggest that the engagement of linguistic processes may inhibit or interfere with perceptual processes in a domain-specific manner.

An alternative hypothesis is based on the idea of resource competition or prime-probe semantic incompatibility. If sentences describing faces engage mechanisms required for face perception, the cost in RT could reflect a competition for similar resources or incompatibility between the prime and probe. This idea is especially reasonable within the current experimental context given that we only varied categorical overlap (face or place primes) but not the relationship of the specific exemplars. That is, the specific content of the priming sentences was not related to the face probes. A similar incompatibility was also present between visual primes and the face probes. However, the visual and linguistic primes produced opposite effects both in terms of the N170 and in the RT study. Thus, the resource competition idea advanced here emphasizes a competition at the domain level (i.e., language vs perception) rather than in terms of a competition between specific, conceptually overlapping representations.

The preceding discussion addressed the psychological implications of our results. In what follows, we consider neural mechanisms that could produce an enhanced N170 to faces after linguistic primes. If the generators had been selectively attenuated or inhibited by the face sentences, the N170 to a subsequent face picture would be larger even if the actual response to the face probes was invariant, assuming that the amplitude of an ERP is influenced by the background state of activity. An fMRI study conducted by Aziz-Zadeh et al. (2008) is consistent with the hypothesis that linguistic input may reduce activation in conceptually related perceptual areas. Though the effects were small, activation in a functionally defined FFA was lower when participants listened to face-related sentences compared to when they listened to place-related sentences. Moreover, this effect was limited to the left hemisphere, similar to the present results. Alternatively, if the processing of face-related sentences disrupted activity in the neural generators of the N170, an increase in this ERP component could result from increased processing efforts required to overcome this interference (Tomasi et al., 2004). This hypothesis bears some similarity to previous accounts of the enhanced N170 response observed with atypical faces, for example, if the face is presented upside down (Rossion et al., 1999; Freire et al., 2000) or with certain features misaligned (Letourneau and Mitchell, 2008). As in the current study, conditions associated with an enhanced N170 were those in which face perception was disrupted. Either of these mechanisms would be consistent with the behavioral data showing that gender discrimination was actually slower after face-related sentences compared to place-related sentences.

In summary, the current results provide compelling and convergent evidence of the influence of language on perception. Surprisingly, the results suggest that language about faces may be disruptive to systems associated with the perception of visually presented faces. It is useful to consider why there may be a bias for online mechanisms to separate linguistic and perceptual representations. When asked to provide verbal descriptions of scenes, people certainly do look at the object being described, a condition in which language and perception would be highly correlated (Tanenhaus and Brown-Schmidt, 2008). However, an important feature of language is that it provides a medium to describe people, places, and events that are spatially and temporally displaced from the current context. When conversing with someone, we typically do not describe the objects that are the focus of our current visual attention. Rather, we describe the vivid cinematography of the film we just attended or the beautiful sunset we saw

last night. In such situations, the current contents of the visual system (e.g., the other person's face, the late night cafe!) are generally uncorrelated with the contents of language. It is possible that the natural statistics of the world may impose a bias that segregates, to some extent, language and perception to attenuate interference that could arise during online processing.

References

- Aziz-Zadeh L, Wilson SM, Rizzolatti G, Iacoboni M (2006) Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Curr Biol* 16:1818–1823.
- Aziz-Zadeh L, Fiebach CJ, Naranayan S, Feldman J, Dodge E, Ivry RB (2008) Modulation of the FFA and PPA by language related to faces and places. *Soc Neurosci* 3:229–238.
- Bentin S, Allison T, Puce A, Perez E, McCarthy G (1996) Electrophysiological studies of face perception in humans. *J Cogn Neurosci* 8:551–565.
- Bentin S, Mouchetant-Rostaing Y, Giard MH, Echallier JF, Pernier J (1999) ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. *J Cogn Neurosci* 11:235–260.
- Bentin S, Taylor MJ, Rousselet GA, Itier RJ, Caldara R, Schyns PG, Jacques C, Rossion B (2007) Controlling interstimulus perceptual variance does not abolish N170 face sensitivity. *Nat Neurosci* 10:801–802; author reply 802–803.
- Bötzel K, Grüsser OJ (1989) Electric brain potentials evoked by pictures of faces and non-faces: a search for “face-specific” EEG-potentials. *Exp Brain Res* 77:349–360.
- Carmel D, Bentin S (2002) Domain specificity versus expertise: factors influencing distinct processing of faces. *Cognition* 83:1–29.
- Dien J (2009) A tale of two recognition systems: implications of the fusiform face area and the visual word form area for lateralized object recognition models. *Neuropsychologia* 47:1–16.
- Fallshore M, Schooler JW (1995) Verbal vulnerability of perceptual expertise. *J Exp Psychol Learn Mem Cogn* 21:1608–1623.
- Feldman JA (2006) From molecule to metaphor: a neural theory of language. Cambridge, MA: Massachusetts Institute of Technology.
- Freire A, Lee K, Symons AL (2000) The face-inversion effect as a deficit in the encoding of configural information: direct evidence. *Perception* 2000:159–170.
- Gallese V, Lakoff G (2005) The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cogn Neuropsychol* 22:455.
- Glenberg AM, Kaschak MP (2002) Grounding language in action. *Psychonom Bull Rev* 9:558–565.
- Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci* 10:14–23.
- Harris A, Nakayama K (2007) Rapid face-selective adaptation of an early extrastriate component in MEG. *Cereb Cortex* 17:63–70.
- Hauk O, Pulvermüller F (2004) Neurophysiological distinction of action words in the fronto-central cortex. *Hum Brain Mapp* 21:191–201.
- Hauk O, Johnsrude I, Pulvermüller F (2004) Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41:301–307.
- Itier RJ, Taylor MJ (2004) Source analysis of the N170 to faces and objects. *NeuroReport* 15:1261–1265.
- Jacques C, Rossion B (2004) Concurrent processing reveals competition between visual representations of faces. *Neuroreport* 15:2417–2421.
- Jacques C, Rossion B (2007) Electrophysiological evidence for temporal dissociation between spatial attention and sensory competition during human face processing. *Cereb Cortex* 17:1055–1065.
- Jeffreys DA (1989) A face-responsive potential recorded from the human scalp. *Exp Brain Res* 78:193–202.
- Kaschak MP, Madden CJ, Theriault DJ, Yaxley RH, Aveyard M, Blanchard AA, Zwaan RA (2005) Perception of motion affects language processing. *Cognition* 94:B79–B89.
- Kaschak MP, Zwaan RA, Aveyard M, Yaxley RH (2006) Perception of auditory motion affects language processing. *Cogn Sci* 30:733–744.
- Kovács G, Zimmer M, Harza I, Antal A, Vidnyánszky Z (2005) Position-specificity of facial adaptation. *Neuroreport* 16:1945–1949.
- Kutas M, Hillyard SA (1980) Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207:203–205.
- Landau A, Bentin S (2008) Attentional and perceptual factors affecting the attentional blink for faces and objects. *J Exp Psychol Hum Percept Perform* 34:818–830.
- Landau AN, Esterman M, Robertson LC, Bentin S, Prinzmetal W (2007) Different effects of voluntary and involuntary attention on EEG activity in the gamma band. *J Neurosci* 27:11986–11990.
- Letourneau SM, Mitchell TV (2008) Behavioral and ERP measures of holistic face processing in a composite task. *Brain Cogn* 67:234–245.
- Macrae CN, Lewis HL (2002) Do I know you? Processing orientation and face recognition. *Psychol Sci* 13:194–196.
- Mangun GR, Hillyard SA (1991) Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *J Exp Psychol Hum Percept Perform* 17:1057–1074.
- Meissner CA, Brigham JC (2001) A meta-analysis of the verbal overshadowing effect in face identification. *Appl Cogn Psychol* 15:603–616.
- Meteyard L, Bahrami B, Vigliocco G (2007) Motion detection and motion verbs: language affects low-level visual perception. *Psychol Sci* 18:1007–1013.
- O'Craven KM, Kanwisher N (2000) Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J Cogn Neurosci* 12:1013–1023.
- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9:97–113.
- Puce A, Allison T, Gore JC, McCarthy G (1995) Face-sensitive regions in human extrastriate cortex studied by functional MRI. *J Neurophysiol* 74:1192–1199.
- Puce A, Allison T, Asgari M, Gore JC, McCarthy G (1996) Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *J Neurosci* 16:5205–5215.
- Richardson DC, Spivey MJ, Barsalou LW, McRae K (2003) Spatial representations activated during real-time comprehension of verbs. *Cogn Sci* 27:767–780.
- Rossion B, Jacques C (2008) Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage* 39:1959–1979.
- Rossion B, Delvenne JF, Debatiste D, Goffaux V, Bruyer R, Crommelinck M, Guérit JM (1999) Spatio-temporal localization of the face inversion effect: an event-related potentials study. *Biol Psychol* 50:173–189.
- Rossion B, Joyce CA, Cottrell GW, Tarr MJ (2003) Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage* 20:1609–1624.
- Schooler JW (2002) Verbalization produces a transfer inappropriate processing shift. *Appl Cogn Psychol* 16:989–997.
- Schooler JW, Engstler-Schooler TY (1990) Verbal overshadowing of visual memories: some things are better left unsaid. *Cogn Psychol* 22:36–71.
- Schooler JW, Fiore SM, Brandimonte MA, Douglas LM (1997) At a loss from words: verbal overshadowing of perceptual memories. *Psychology of learning and motivation*, pp 291–340. New York: Academic.
- Schyns PG, Oliva A (1999) Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition* 69:243–265.
- Sreenivasan KK, Jha AP (2007) Selective attention supports working memory maintenance by modulating perceptual processing of distractors. *J Cogn Neurosci* 19:32–41.
- Tanenhaus MK, Brown-Schmidt S (2008) Language processing in the natural world. *Philos Trans R Soc B Biol Sci* 363:1105–1122.
- Tettamanti M, Buccino G, Saccuman MC, Gallese V, Danna M, Scifo P, Fazio F, Rizzolatti G, Cappa SF, Perani D (2005) Listening to action-related sentences activates fronto-parietal motor circuits. *J Cogn Neurosci* 17:273–281.
- Thierry G, Martin CD, Downing P, Pegna AJ (2007) Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nat Neurosci* 10:505–511.
- Tomasi D, Ernst T, Caparelli EC, Chang L (2004) Practice-induced changes of brain function during visual attention: a parametric fMRI study at 4 Tesla. *Neuroimage* 23:1414–1421.
- Zwaan RA, Yaxley RH (2004) Lateralization of object-shape information in semantic processing. *Cognition* 94:B35–B43.