Seeing Objects through the Language Glass

Bastien Boutonnet[1], Benjamin Dering[2], Nestor Viñas-Guasch[1],
and Guillaume Thierry[1]

Abstract

■ Recent streams of research support the Whorfian hypothesis according to which language affects one's perception of the world. However, studies of object categorization in different languages have heavily relied on behavioral measures that are fuzzy and inconsistent. Here, we provide the first electrophysiological evidence for unconscious effects of language terminology on object perception. Whereas English has two words for cup and mug, Spanish labels those two objects with the word "taza." We tested native speakers of Spanish and English in an object detection task using a visual oddball paradigm, while measuring event-related brain potentials. The early deviant-related negativity elicited by deviant stimuli was greater in English than in Spanish participants. This effect, which relates to the existence of two labels in English versus one in Spanish, substantiates the neurophysiological evidence that language-specific terminology affects object categorization. ■

INTRODUCTION

The question of language–thought interactions has recently become a major topic of interest in cognitive neuroscience. It has become essential because of the debate on language encapsulation and on the potential effects of language on other cognitive processes (Fodor, 1975, 2008; Chomsky, 2000). The linguistic relativity hypothesis has undergone several interpretations since its inception by Whorf, Carroll, and Chase (1956). One early (misleading) interpretation of the hypothesis contends that language determines thought and, therefore, that without language thought is impossible. In light of compelling evidence that high-level cognitive operations are possible without language, this position has simply become untenable (e.g., number cognition in primates [Gallistel, 1989], infants [Feigenson, Dehaene, & Spelke, 2004], and in languages that do not have a complex lexicalized umber system [Gordon, 2004]). On the other hand, recent theoretical reconceptualizations (e.g., Gentner & Goldin-Meadow, 2003; Gumperz & Levinson, 1996) have put forward nondeterministic versions of the hypothesis, according to which language influences (rather than determines) thought. The linguistic relativity debate has therefore moved toward the question of interaction between language representations and perception rather than that of determinism (Lucy, 1992a). However, this reconceptualization lacks psychological and physiological underpinning.

Here, we aimed at testing the validity of the most recent theoretical take on the Whorfian hypothesis, which does away with a "strong/weak" distinction (Klemfuss, Prinzmetal, & Ivry, 2012; Lupyan, 2012) and offers researchers clearer working hypotheses regarding language–thought interactions. For instance, based on interactive-processing models such as those developed by McClelland and Rumelhart (1981), the label–feedback hypothesis (Lupyan, 2012) proposes that language is highly interconnected with other cognitive processes such as vision and categorization and that it produces transient modulations of on-going perceptual (and higher level) processing. Whorfian effects can therefore arise from interactions among distributed brain regions, as in the case of prefontal areas preparing the visual cortex to perceive particular dimensions of stimuli before they are actually displayed (Lamme & Roelfsema, 2000). This model therefore allows for nontrivial linguistic relativity effects to arise but is not tied in a deterministic view where perceptual areas are functionally structured by language (for an exhaustive explanation of the hypothesis and a review of the experimental literature, see Lupyan, 2012).

Previous studies have highlighted areas where lexical and grammatical information affect domain-general cognitive processes. For instance, lexicalization constraints on spatial representation and event conceptualization (e.g., focus on manner vs. end point of motion) have been shown to affect speakers' event description and recollection (Papafragou & Selimis, 2010; Majid, Bowerman, Kita, Haun, & Levinson, 2004; Bowermann & Choi, 1991) or to elicit different gaze patterns when exploring scenes depicting events (Flecken, 2010). Studies investigating grammatical number (i.e., language with classifier systems) reveal a tendency to categorize objects on the basis of

---

[1]Bangor University, UK, [2]The Salk Institute for Biological Studies, La Jolla, CA

substance rather than shape when classifiers put the focus on substance (Saalbach & Imai, 2007; Zhang & Schmitt, 1998; Lucy, 1992b). In a similar vein, grammatical gender has also been shown to affect speakers' object categorization in covert gender assignment tasks (Kurinski & Sera, 2010; Forbes, Poulin-Dubois, Rivero, & Sera, 2008; Bassetti, 2007; Sera et al., 2002), judgment and adjective-association tasks (Boroditsky, Schmidt, & Phillips, 2003; Phillips & Boroditsky, 2003), and priming paradigms (Boutonnet, Athanasopoulos, & Thierry, 2012; Cubelli, Paolieri, Lotto, & Job, 2011). Finally, differences in color terminology have been shown to affect color perception in behavioral (Athanasopoulos, 2009; Franklin et al., 2008; Roberson, Davidoff, Davies, & Shapiro, 2005; Ozgen, 2004) and neurophysiological (Athanasopoulos, Dering, Wiggett, Kuipers, & Thierry, 2010; Clifford, Holmes, Davies, & Franklin, 2010; Liu et al., 2010; Thierry, Athanasopoulos, Wiggett, Dering, & Kuipers, 2009) investigations.

Despite the evidence in favor of the existence of Whorfian effects, it remains that studies in the field have mostly relied on behavioral measures. The problem is that such measures are open to contamination by explicit strategies used by participants to resolve the tasks, a process likely to involve language processing. Here, following neurophysiological investigations in the domain of color (Liu et al., 2010; Thierry et al., 2009; Franklin et al., 2008; Roberson, Pak, & Hanley, 2008; Gilbert, Regier, Kay, & Ivry, 2006), we investigated whether language-specific terminology also constrains object categorization (Gilbert, Regier, Kay, & Ivry, 2008). For instance, Thierry and colleagues (2009) recorded ERP correlates of color change detection in Greek–English bilinguals who have two color terms for blue (*ble* = "dark blue" and *ghalazio* = "light blue") and a control group of English monolinguals. They found that native speakers of Greek exhibited a greater visual MMN (vMMN) elicited by blue deviants than English controls. On the basis of this paradigm, we chose to extend the evidence from the domain of color perception to that of object categorization. We chose drinking vessels because they have been examined thoroughly in previous cross-linguistic naming studies (Pavlenko & Malt, 2010; Ameel, Malt, Storms, & Van Assche, 2009; Ameel, Storms, Malt, & Sloman, 2005). These studies suggest that bilingual speakers' categorical boundaries shift through exposure to their second language—a phenomenon that has also been reported for colors (Athanasopoulos et al., 2010).

We recorded brain potentials from Spanish and English native speakers while they performed an object detection task within an oddball paradigm to test the extent to which unconscious aspects of visual object processing are modulated by one's language.

Spanish differs from English in the way some objects are labeled. Whereas English has two words to refer to a cup and a mug, Spanish only uses one label for these two objects: "taza." In this experiment, participants were presented with three stimuli within an oddball paradigm (one of high local probability, i.e., standard; and two of low local probability, i.e., deviants). Participants were instructed to detect a particular deviant stimulus or target (a bowl) in each of two experimental blocks. In one block, the nontarget deviant was a cup and the standard was a mug, and in the other block, the nontarget deviant was a mug and the standard was a cup.

We expected nontarget deviants to spontaneously elicit a deviant-related negativity (DRN) regardless of a response from the participants (Winkler, Czigler, Sussman, Horváth, & Balázs, 2005; Czigler, Balázs, & Pató, 2004; Czigler, Balázs, & Winkler, 2002; Turatto, Angrilli, Mazza, Umilta, & Driver, 2002; Csibra, Czigler, & Ambro, 1994). Because of the terminological difference between English and Spanish, we expected that the change from cup to mug would elicit a greater DRN in English than Spanish participants.

METHODS

Participants

Participants were 13 native speakers of Spanish (10 women, 3 men; M_{Age} = 21 years, SD = 1.6 years) tested in Spain and 14 native speakers of English tested in Wales (8 women, 6 men; M_{Age} = 20 years, SD = 0.6 years). Spanish participants were recruited from a database filtered to have a level no higher than A2 in English and a daily use of English lower than 5%. As part of the normal education curriculum in Spain, all Spanish participants received some exposure to English, but all reported having a limited knowledge of the language as well as a rare use of it. None of the Spanish participants had spent more than 2 weeks in an English-speaking country. The language usage background data used to filter the database were collected from self-reports from the participants before entry in the database.

Some of the Spanish speakers were also fluent in Catalan. This was not considered a problem because Catalan and Spanish are matched with respect to object denomination for cups and mugs. Some of the English participants reported having basic knowledge of other languages (including Spanish) but had self-reported very low proficiency and were not using any of their other languages on an everyday basis.

Materials

Three grayscale photographs of a cup, a mug, and a bowl subtending approximately 8° of visual angle were presented in the middle of a white background square in the center of a CRT monitor.

Procedure

Participants viewed two blocks of 450 stimuli. Within each block, a standard stimulus was presented with a high local

probability (either a cup or a mug, 80%). Deviant stimuli, presented with a low local probability, were either to be ignored (a mug or a cup, depending on the nature of the standard, 15%) or to be reported (bowl target, 5%). Presentation order was pseudorandomized such that two deviants or targets never appeared in immediate succession, and there were at least three standards in a row between two deviants. Stimuli were presented for 300 msec with a random variable ISI of 400, 450, 500, 550, and 600 msec, averaging to 500 msec. Participants were instructed to detect the target object (bowl) by pressing a button on a response box as quickly as possible. Block order was fully counterbalanced between participants.

Electrophysiological Recording

Electrophysiological data were recorded in two different laboratories. The Spanish participants were tested in Barcelona, Spain (Pompeu Fabra University). EEG was recorded (BrainVision Recorder 1.10, Charlotte, NC) in reference to the left mastoid electrode at the rate of 1 kHz from 34 tin electrodes placed according to the 10–20 convention. Impedances were kept below 5 kΩ for electrodes on the cap and below 10 kΩ for external electrodes. The English participants were tested in Bangor, Wales (Bangor University). EEG was recorded (Neuro-Scan 4.4, Charlotte, NC) in reference to the left mastoid electrode at the rate of 1 kHz from 34 Ag–Cl electrodes placed according to the 10–20 convention. All impedances were kept below 5 kΩ for electrodes on the cap and below 10 kΩ for external electrodes. Both data sets were analyzed using BrainVision Analyzer 2. EEG activity was filtered off-line with a high-pass 0.1-Hz filter (slope of 12 dB/oct) and a low-pass 30-Hz filter (slope of 48 dB/oct).

Data Analysis

Accuracy scores and RTs were submitted to independent samples t tests between groups (t_1 and t_2, respectively). Eye blinks were mathematically corrected using the Gratton, Coles, and Donchin (1983) algorithm provided in Brain Vision Analyzer 2, and epochs with activity exceeding ±75 μV at any electrode site were automatically discarded. Epochs ranged from −100 to 600 msec after stimulus onset. Baseline correction was performed in reference to prestimulus activity, and individual averages were rereferenced to the left and right mastoid off-line. ERPs time-locked to the onset of the pictures were visually inspected, and mean amplitudes were measured in temporal windows determined based on variation of the mean global field power measured across the scalp (Picton et al., 2000). ERPs elicited by standard stimuli were averaged across blocks as were ERPs elicited by deviants; therefore, comparisons between standard and deviants did not reflect inherent perceptual differences between cups and mugs but only the deviancy effect.

Potential perceptual differences between the cup and mug objects were also investigated by analyzing amplitude and latency of the P1 peak from ERPs computed from standard stimuli, separately for each of the two experimental blocks. The P1 was maximal at parietal sites and was measured in the 100- to 150-msec range. Mean amplitude and latency of the P1 collected from a linear derivation of the five electrodes of interest (PO1, PO2, O1, OZ, and O2) were submitted to a 2 within-subject × 2 between-subject ANOVA with Standard Object (cup/mug) as a within-subject factor and Language Group (Spanish/English) as a between-subject factor.

The DRN was defined as the earliest modulation of the negative component following the P1 over occipital recording sites. DRN analysis was conducted on individual ERPs elicited by standards and nontarget deviants, was maximal over the parieto-occipital scalp, and was studied in the 145- to 180-msec range at electrodes PO1, PO2, O1, OZ, and O2, predicted to be the electrodes of maximal sensitivity for the effect measured (Liu et al., 2010; Thierry et al., 2009). Mean amplitudes of ERPs from standard and deviant stimuli were subjected to a mixed repeated-measures ANOVA with Deviancy (deviant/standard) and Electrode (five levels) as a within-subject factors and Language Group (Spanish/English) as a between-subject factor. In addition, paired sample t tests were conducted between the standard and deviant conditions millisecond-by-millisecond to determine the onset of differences between conditions (using a linear derivation of the five electrodes used in the mean amplitude analysis).

Furthermore, the latency of the N1 elicited by nontarget deviants was compared with that of the N1 elicited by the standards, measured at the electrode of maximal amplitude (O2). Peak latencies were submitted to a 2 within-subject × 2 between-subject ANOVA with Deviancy (standard/deviant) as a within-subject factor and Language Group (Spanish/English) as a between-subject factor.

Because some native speakers of Spanish were also Spanish–Catalan bilinguals, we investigated potential differences in attention allocation between groups by comparing ERPs elicited by mug standards and bowl targets on the one hand and cup standards and bowl targets on the other hand, because these comparisons always involved objects that have different names in both of the languages. P1s and DRNs elicited by "cup," "mug," and "bowl" (in identical time windows and the same electrodes as the analyses above) were subjected to repeated-measures ANOVAs with Object (cup–bowl/mug–bowl) as within-subject factor and Language Group (Spanish/English) as a between-subject factor. Because of the very high level of repetition involved in the oddball paradigm used here, we expected potential differences in attention to have a negligible impact on basic object discrimination as indexed by DRN. We therefore expected to find no interaction between object type and group in these comparisons.

RESULTS

Behavioral Data

Accuracy in the bowl detection task was above 90% in all participants and blocks ($M_{English} = 0.94$, $SD = 0.02$; $M_{Spanish} = 0.93$, $SD = 0.02$). There was no significant differences between groups on target detection accuracy nor RTs ($t_1(25) = .62$, $p > .05$; $t_2(25) = .29$, $p > .05$).

Electrophysiological Data

Critical Comparison: Standard (Cup/Mug) versus Passive Deviant (Cup/Mug)

As expected, nontarget deviants elicited a greater DRN as compared with standards. This difference was qualified by a significant main effect of Deviancy ($F(1, 25) = 10.3$, $p < .05$, $\eta_p^2 = 0.29$) with deviant stimuli eliciting more negative amplitudes than standard stimuli in the DRN window. The effect of Deviancy further interacted with Language Group ($F(1, 25) = 4.9$, $p < .05$, $\eta_p^2 = 0.16$), such that the deviancy effect was of significantly greater magnitude in English than Spanish participants (Figure 1A and B).

Post hoc tests showed that there was no significant DRN effect in the Spanish group ($F(1, 12) = .46$, $p > .05$, $\eta_p^2 = 0.04$) but a significant effect in the English group ($F(1, 13) = 16.31$, $p = .001$, $\eta_p^2 = 0.56$; Figure 1C). Furthermore, there was no significant difference between standard and deviant conditions at any point in time in the DRN window in the Spanish participants, but standard and deviant conditions differed significantly from 135 to 177 msec in the English group (lower part of Figure 1A and B). To reduce the risk of type I errors and given the high levels of autocorrelation of ERP time series, we followed the method advocated by Guthrie and Buchwald (1991) where only sequences with a minimum of 12 consecutive significant t tests were considered (see, for instance, Kuipers & Thierry, 2011).
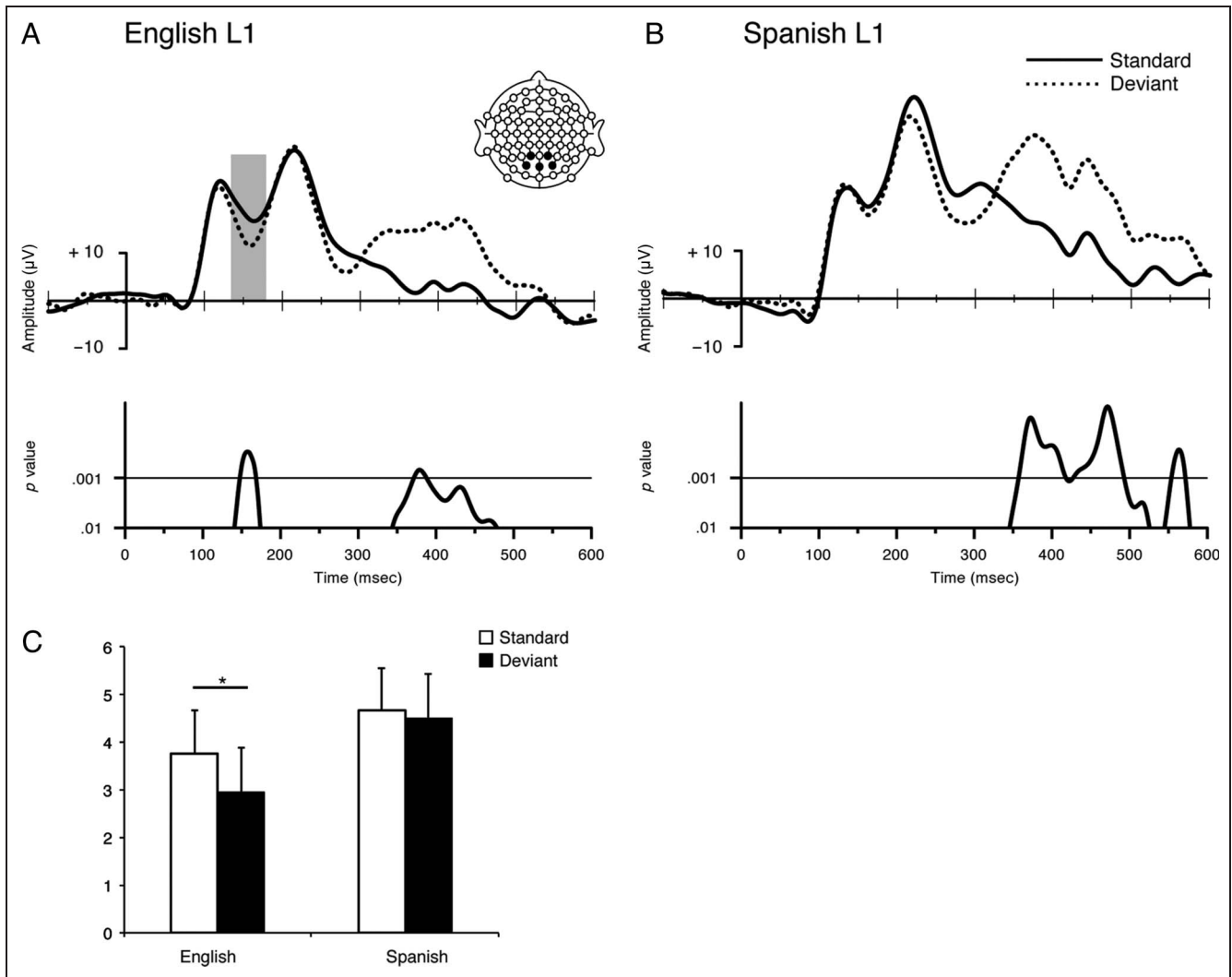

Figure 1. Event-related brain potentials elicited by standard and deviant stimuli averaged across blocks. ERPs and plots of p value of differences between conditions in (A) native speakers of English and (B) native speakers of Spanish. (C) Plot of DRN mean amplitude. Waveforms correspond to linear derivation of electrodes PO1, PO2, O1, OZ, and O2. Error bars depict *SEM*.

Figure 2. Event-related brain potentials elicited by cup and mug standards in each of the two experimental blocks. (A) Native speakers of English and (B) native speakers of Spanish. Waveforms correspond to linear derivation of electrodes PO1, PO2, O1, OZ, and O2.
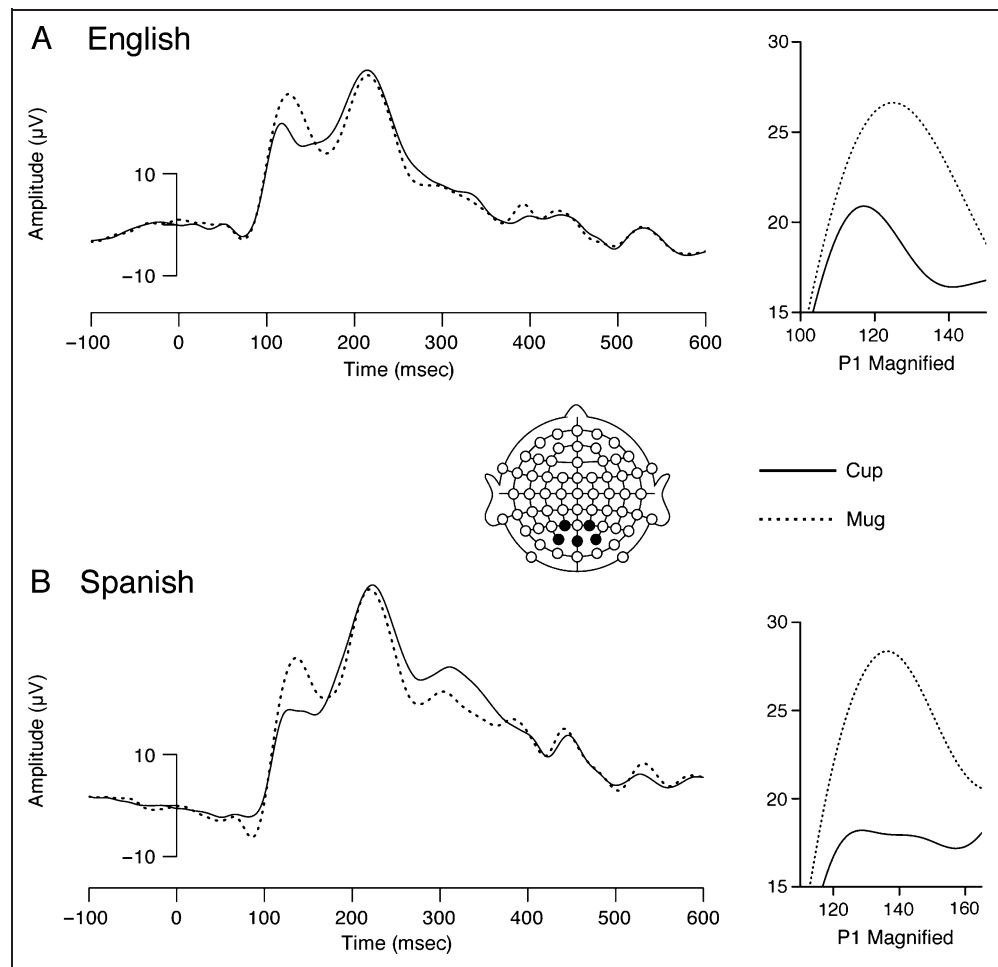
Latency analyses of the DRN revealed no significant differences between group or condition in the window of interest ($F(1, 24) = 1.53$, $p > .05$, $\eta_p^2 = 0.06$).

ERPs elicited by standard stimuli in each of the two blocks considered separately (Figure 2) displayed significant differences in P1 mean amplitude (F_1) and latency (F_2) between cup and mug ($F_1(1, 24) = 5.76$, $p < .05$, $\eta_p^2 = 0.19$; $F_2(1, 24) = 17.56$, $p < .001$, $\eta_p^2 = 0.42$). Critically, these effects did not interact with participant group ($F_1(1, 24) = 1.29$, $p > .05$, $\eta_p^2 = 0.05$; $F_2(1, 24) = 3.2$, $p > .05$, $\eta_p^2 = 0.12$).

Control Comparison: Standard (Cup/Mug) versus Target (Bowl)

ANOVAs on the P1 revealed a significant effect of Object Type in both the mug versus bowl comparison ($F(1, 25) = 50.32$, $p < .0001$, $\eta_p^2 = 0.69$) and the cup versus bowl comparison ($F(1, 25) = 40.28$, $p < .0001$, $\eta_p^2 = 0.62$). Critically, there was no interaction between Language Group and Object Type in either comparisons (both $ps > .1$).

ANOVAs on the DRN revealed a significant effect of Object Type in both the mug versus bowl comparison ($F(1, 25) = 40.28$, $p < .0001$, $\eta_p^2 = 0.62$) and the cup versus bowl comparison ($F(1, 25) = 48.57$, $p < .0001$,

$\eta_p^2 = 0.66$) (Figure 3). Again, there was no interaction between Language Group and Object Type in either comparisons (both $ps > .1$).

DISCUSSION

This study tested potential effects of language-specific terminology on early stages of visual perception and categorization based on the analysis of spontaneous modulations of the P1/N1 event-related brain potential complex. In a design controlling for perceptual features of the objects presented, ERPs successfully distinguished standards and deviants within the N1 range in native speakers of English but not in speakers of Spanish who name both these objects using the same noun. Moreover, when comparing the P1 elicited by the two objects presented as standards in each of the blocks, ERP differences were indistinguishable between groups.

The N1 range of ERPs is thought to index stages of visual processing beyond categorical discrimination (Dering, Martin, Moro, Pegna, & Thierry, 2011; Thierry, Martin, Downing, & Pegna, 2007a). Indeed, categorical effects have been reported in the domain of face processing in the P1 range and even earlier (Thierry et al., 2007a;
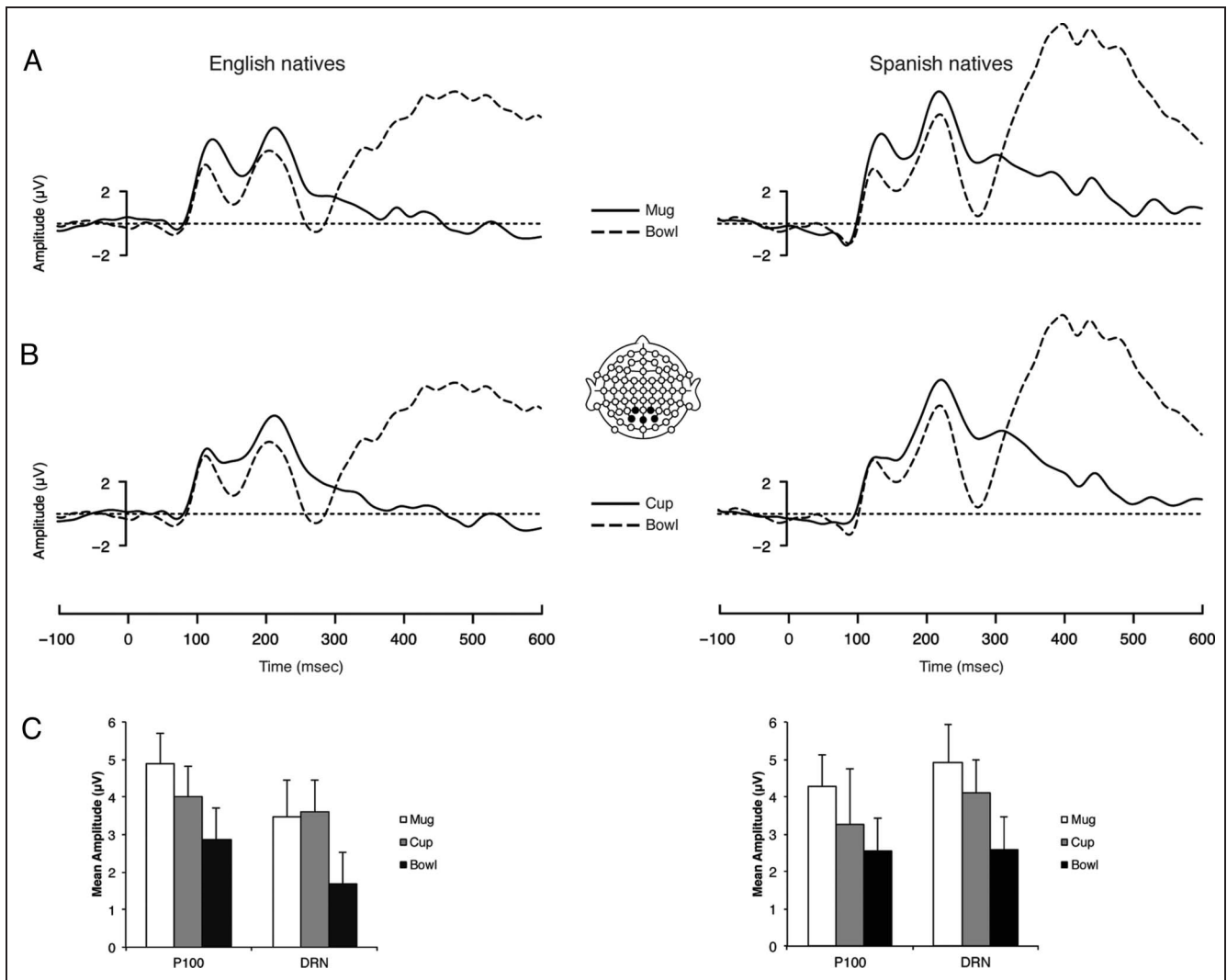
Figure 3. Event-related brain potentials elicited by (A) mug standards and bowl targets and (B) cup standards and bowl targets. (C) Plot of P1 and DRN mean amplitudes in both participant groups. Waveforms correspond to linear derivation of electrodes PO1, PO2, O1, OZ, and O2. Error bars depict *SEM*.

Thierry, Martin, Downing, & Pegna, 2007b; Seeck et al., 1997, 2001). Therefore, because it occurs beyond the P1 range, the DRN effect found here concerns relatively sophisticated levels of visual object processing—probably relating to object identity resolution. Critically, however, the DRN occurred before the temporal window in which lexical representations are considered to be accessed. Indeed, during practiced picture naming, Strijkers, Costa, and Thierry (2010) and Costa, Strijkers, Martin, and Thierry (2009) have established that lexical access occurs between 180 and 200 msec after picture onset. Here, significant differences were observed as early as 145 msec after picture onset. In addition, as shown by Strijkers and colleagues (Strijkers, Holcomb, & Costa, 2011), lexical access appears to be substantially delayed until ~350 msec after stimulus onset when there is no requirement to name the pictures (see also Blackford, Holcomb, Grainger, & Kuperberg, 2012). This was indeed the case here because participants were asked to press a button when they saw a specific

object and not instructed to name them. Thus, the influence of language-specific terminology on object processing does not merely result from online interaction with processes underlying lexical access. In other words, our finding is not simply an effect of language on language.

We report the N1 modulation recorded here as a DRN rather than a vMMN (the visual counterpart of the auditory MMN; Winkler et al., 2005; Czigler et al., 2002) because the vMMN proper is supposedly only elicited by visual stimuli presented outside the focus of attention, for example, in peripheral vision rather than fixation (Clifford et al., 2010). However, (a) the latency of the DRN effect we reported here is similar to that previously reported in vMMN studies (Pazo-Alvarez, Cadaveira, & Amenedo, 2003); (b) like our effect, the vMMN has a parieto-occipital topography with a right hemispheric predominance. Because the DRN in this study (peak time: ~160 msec at electrode O2) peaked substantially earlier and was observed at a different scalp location than N2 modulations

elicited by overt cognitive control (Folstein & Van Petten, 2007), we interpret this effect as an index of automatic, preattentional, and, crucially, prelexical cognitive mechanism (Strijkers et al., 2010, 2011; Costa et al., 2009).

The P1 results further suggest that Spanish and English participants perceptually discriminated cup and mug pictures in a similar fashion. These two objects are indeed ostensibly different, and P1 amplitude has been shown to distinguish different object types previously (Dering et al., 2011; e.g., Thierry et al., 2007a). Therefore, the DRN effect observed in the N1 window cannot be explained by differences arising at more elementary stages of perceptual analysis preceding the N1 window. Furthermore, we consider the absence of between-group differences in the P1 range to be of fundamental importance because they could be underpinned by differences in cultural background or ethnic origin or even genetic factors and would therefore invalidate our results as merely stemming from different perceptual grooming in different environments.

Differences between groups in the P1 range could have been expected because our group has already reported such differences in a previous study of color perception (Thierry et al., 2009). However, it must be noted that the relationship between color terminology and P1 measurement was not trivial in that it did not yield a P1 amplitude by language group interaction. Expecting a reduction or cancellation of P1 differences between cups and mugs in the Spanish participants here would assume that perceptual differences between a cup and a mug are even more subtle than perceptual differences between two neighboring shades of blue, which have been shown to occur between 100 and 200 msec after stimulus onset (Fonteneau & Davidoff, 2007). We contend that cups and mugs are more discriminable at a perceptual level (at least by shape, size, and luminance) than two discs of the same size and color saturation, differing exclusively by their relative luminance. For example, people will argue indefinitely about color names at the green–blue or the navy–indigo border, but the same individuals will hardly argue as to what differentiates a mug and a cup shape. Therefore, it is reasonable to assume that P1 differences indexing early perceptual distinctions should effectively discriminate cups and mugs in both groups but that orientation responses measured by the DRN would be selectively affected by language terminology.

The fact that differences occur only in the N1 range and based on standard–deviant comparisons is essential to demonstrate an effect of language terminology on high-level perceptual processing. Additionally, these differences arising beyond the P1 range are consistent with an interactional account of linguistic relativity effects (Lupyan, 2012) because basic perception need not be changed for such effects to arise.

Our experimental design also allowed us to investigate potential attentional differences between the Spanish–Catalan speakers and English monolinguals. Indeed, one could argue that the interaction on the DRN could be a result of better inhibition/monitoring mechanisms in the bilinguals. As suggested by our results, this was not the case because, when the items both had a different label in Spanish and English, the DRN elicited between target and standard had the same magnitude in the two groups. If Spanish participants had different attentional skills, and if such skills were generically reflected in DRN modulation, we would have expected the interaction observed in the critical comparison (mug/cup) to carry over to the case of comparisons with the target (bowl).

To our knowledge, this is the first neurophysiological demonstration of a relationship between native language and spontaneous object identity discrimination during visual perception, which goes beyond the observation of overt effects on object categorization (Pavlenko & Malt, 2010; Ameel et al., 2005, 2009). Furthermore, these findings generalize the linguistic relativity effects previously reported in the case of color perception (Liu et al., 2010; Thierry et al., 2009; Franklin et al., 2008) to the domain of object identity processing (Gilbert et al., 2008; arguably affecting higher-level cognitive representations). Overall, our results are incompatible with the view that language is functionally encapsulated in the human brain and fundamentally independent of, for example, visual cognition (Fodor, 1975, 2008; Chomsky, 2000; Pinker, 1995, 2007). On the contrary, they support an interactive conceptualization of the brain where language is highly integrated and can modulate ongoing cognitive processes such as object categorization and perception (Lupyan, 2012). Future studies will determine whether the effects reported here are confined to interactions within the left hemisphere (Mo, Xu, Kay, & Tan, 2011; Regier & Kay, 2009; Franklin et al., 2008; Roberson et al., 2008) and the extent to which they are adaptable over time (Athanasopoulos et al., 2010).

REFERENCES

Ameel, E., Malt, B. C., Storms, G., & Van Assche, F. (2009). Semantic convergence in the bilingual lexicon. *Journal of Memory and Language, 60,* 270–290.

Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language, 53,* 60–80.

Athanasopoulos, P. (2009). Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition, 12,* 83.

Athanasopoulos, P., Dering, B., Wiggett, A., Kuipers, J.-R., & Thierry, G. (2010). Perceptual shift in bilingualism: Brain potentials reveal plasticity in pre-attentive colour perception. *Cognition, 116,* 437–443.

Bassetti, B. (2007). Bilingualism and thought: Grammatical gender and concepts of objects in Italian–German bilingual children. *International Journal of Bilingualism, 11,* 251–273.

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

Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, syntax, and semantics. In D. Centner & S. Goldin-Meadow, *Language in Mind: Advances in the Study of Language and Cognition* (pp. 61–79). Cambridge, MA: MIT Press.

Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research, 1479,* 72–79.

Bowermann, M., & Choi, S. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition, 41,* 83–121.

Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.

Clifford, A., Holmes, A., Davies, I. R. L., & Franklin, A. (2010). Color categories affect pre-attentive color perception. *Biological Psychology, 85,* 275–282.

Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proceedings of the National Academy of Sciences, U.S.A., 106,* 21442–21446.

Csibra, G., Czigler, I., & Ambro, A. (1994). Effects of stimulus alternation, repetition and response requirements on event-related potentials to patterned visual stimuli. *Biological Psychology, 37,* 115–132.

Cubelli, R., Paolieri, D., Lotto, L., & Job, R. (2011). The effect of grammatical gender on object categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 449–460.

Czigler, I., Balázs, L., & Pató, L. G. (2004). Visual change detection: Event-related potentials are dependent on stimulus location in humans. *Neuroscience Letters, 364,* 149–153.

Czigler, I., Balázs, L., & Winkler, I. (2002). Memory-based detection of task-irrelevant visual changes. *Psychophysiology, 39,* 869–873.

Dering, B., Martin, C. D., Moro, S., Pegna, A. J., & Thierry, G. (2011). Face-sensitive processes one hundred milliseconds after picture onset. *Frontiers in Human Neuroscience, 5,* 93.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8,* 307–314.

Flecken, M. (2010). Event conceptualization by early Dutch–German bilinguals: Insights from linguistic and eye-tracking data. *Bilingualism: Language and Cognition, 14,* 61–77.

Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Fodor, J. A. (2008). *LOT 2: The language of thought revisited*. Oxford: Clarendon Press.

Folstein, J. R., & Van Petten, C. (2007). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology, 45,* 152–170.

Fonteneau, E., & Davidoff, J. (2007). Neural correlates of colour categories. *NeuroReport, 18,* 1323–1327.

Forbes, J. N., Poulin-Dubois, D., Rivero, M. R., & Sera, M. D. (2008). Grammatical gender affects bilinguals' conceptual gender: Implications for linguistic relativity and decision making. *Open Applied Linguistics Journal, 1,* 68–76.

Franklin, A., Drivonikou, G., Clifford, A., Kay, P., Regier, T., & Davies, I. (2008). Lateralization of categorical perception of color changes with color term acquisition. *Proceedings of the National Academy of Sciences, U.S.A., 105,* 18221.

Gallistel, C. R. (1989). Animal cognition: The representation of space, time and number. *Psychology, 40,* 155–189.

Gentner, D., & Goldin-Meadow, S. (2003). *Language in mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press.

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 489–494.

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2008). Support for lateralization of the Whorf effect beyond the realm of color discrimination. *Brain and Language, 105,* 91–98.

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science, 306,* 496–499.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55,* 468–484.

Gumperz, J. J., & Levinson, S. C. (1996). *Rethinking linguistic relativity*. Cambridge: Cambridge University Press.

Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology, 28,* 240–244.

Klemfuss, N., Prinzmetal, W., & Ivry, R. B. (2012). How does language change perception: A cautionary note. *Frontiers in Psychology, 3,* 78.

Kuipers, J.-R., & Thierry, G. (2011). N400 amplitude reduction correlates with an increase in pupil size. *Frontiers in Human Neuroscience, 5,* 61.

Kurinski, E., & Sera, M. D. (2010). Does learning Spanish grammatical gender change English-speaking adults' categorization of inanimate objects? *Bilingualism: Language and Cognition, 14,* 203–220.

Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences, 23,* 571–579.

Liu, Q., Li, H., Campos, J. L., Teeter, C., Tao, W., Zhang, Q., et al. (2010). Language suppression effects on the categorical perception of colour as evidenced through ERPs. *Biological Psychology, 85,* 45–52.

Lucy, J. A. (1992a). *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.

Lucy, J. A. (1992b). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label–feedback hypothesis. *Frontiers in Psychology, 3,* 54.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences, 8,* 108–114.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88,* 375.

Mo, L., Xu, G., Kay, P., & Tan, L. H. (2011). Electrophysiological evidence for the left-lateralized effect of language on preattentive categorical perception of color. *Proceedings of the National Academy of Sciences, U.S.A., 108,* 14026–14030.

Ozgen, E. (2004). Language, learning, and color perception. *Current Directions in Psychological Science, 13,* 95–98.

Papafragou, A., & Selimis, S. (2010). Event categorisation and language: A cross-linguistic study of motion. *Language and Cognitive Processes, 25,* 224–260.

Pavlenko, A., & Malt, B. C. (2010). Kitchen Russian: Cross-linguistic differences and first-language object

naming by Russian–English bilinguals. *Bilingualism: Language and Cognition, 14,* 19–45.

Pazo-Alvarez, P., Cadaveira, F., & Amenedo, E. (2003). MMN in the visual modality: A review. *Biological Psychology, 63,* 199–236.

Phillips, W., & Boroditsky, L. (2003). Can quirks of grammar affect the way you think? Grammatical gender and object concepts. *Proceedings of the 25th Annual Meeting of the Cognitive Science Society,* 928–933.

Picton, T., Bentin, S., Berg, P., Donchin, E., Hillyard, S., Johnson, R., Jr., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology, 37,* 127–152.

Pinker, S. (1995). *The language instinct: The new science of language and mind*. London: Penguin.

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. New York: Viking.

Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences, 13,* 439–446.

Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology, 50,* 378–411.

Roberson, D., Pak, H., & Hanley, J. R. (2008). Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition, 107,* 752–762.

Saalbach, H., & Imai, M. (2007). Scope of linguistic influence: Does a classifier system alter object concepts? *Journal of Experimental Psychology: General, 136,* 485–501.

Seeck, M., Michel, C. M., Blanke, O., Thut, G., Landis, T., & Schomer, D. L. (2001). Intracranial neurophysiological correlates related to the processing of faces. *Epilepsy & Behavior, 2,* 545–557.

Seeck, M., Michel, C. M., Mainwaring, N., Cosgrove, R., Blume, H., Ives, J., et al. (1997). Evidence for rapid face recognition from human scalp and intracranial electrodes. *NeuroReport, 8,* 2749–2754.

Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodríguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General, 131,* 377–397.

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

Strijkers, K., Holcomb, P. J., & Costa, A. (2011). Conscious intention to speak proactively facilitates lexical access during overt object naming. *Journal of Memory and Language, 65,* 345–362.

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J. R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences, U.S.A., 106,* 4567–4570.

Thierry, G., Martin, C. D., Downing, P., & Pegna, A. J. (2007a). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nature Neuroscience, 10,* 505–511.

Thierry, G., Martin, C., Downing, P., & Pegna, A. (2007b). Is the N170 sensitive to the human face or to several intertwined perceptual and conceptual factors? *Nature Neuroscience, 10,* 802–803.

Turatto, M., Angrilli, A., Mazza, V., Umilta, C., & Driver, J. (2002). Looking without seeing the background change: Electrophysiological correlates of change detection versus change blindness. *Cognition, 84,* B1–B10.

Whorf, B. L., Carroll, J. B., & Chase, S. (1956). *Language, thought and reality*. New York: Wiley.

Winkler, I., Czigler, I., Sussman, E., Horváth, J., & Balázs, L. (2005). Preattentive binding of auditory and visual stimulus features. *Journal of Cognitive Neuroscience, 17,* 320–339.

Zhang, S., & Schmitt, B. (1998). Language-dependent classification: The mental representation of classifiers in cognition, memory, and ad evaluations. *Journal of Experimental Psychology: Applied, 4,* 375.