

# First- and Second-language Phonological Representations in the Mental Lexicon

Núria Sebastian-Gallés<sup>1,2</sup>, Antoni Rodríguez-Fornells<sup>2,3</sup>,  
Ruth de Diego-Balaguer<sup>4</sup>, and Begoña Díaz<sup>1,2</sup>

## Abstract

■ Performance-based studies on the psychological nature of linguistic competence can conceal significant differences in the brain processes that underlie native versus nonnative knowledge of language. Here we report results from the brain activity of very proficient early bilinguals making a lexical decision task that illustrates this point. Two groups of Spanish–Catalan early bilinguals (Spanish-dominant and Catalan-dominant) were asked to decide whether a given form was a Catalan word or not. The nonwords were based on real words, with one vowel changed. In the experimental stimuli, the vowel change involved a Catalan-specific contrast that previous research had shown to be difficult for Spanish natives to perceive. In the control stimuli, the vowel switch involved contrasts common to Spanish and Catalan. The results indicated that the groups of bilinguals did not differ in their behavioral and event-related brain potential measurements for the control stimuli; both

groups made very few errors and showed a larger N400 component for control nonwords than for control words. However, significant differences were observed for the experimental stimuli across groups: Specifically, Spanish-dominant bilinguals showed great difficulty in rejecting experimental nonwords. Indeed, these participants not only showed very high error rates for these stimuli, but also did not show an error-related negativity effect in their erroneous nonword decisions. However, both groups of bilinguals showed a larger correct-related negativity when making correct decisions about the experimental nonwords. The results suggest that although some aspects of a second language system may show a remarkable lack of plasticity (like the acquisition of some foreign contrasts), first-language representations seem to be more dynamic in their capacity of adapting and incorporating new information. ■

## INTRODUCTION

Learning a second language as an adult is a difficult task. In fact, it is so hard that most human beings, in spite of their efforts, fail to attain native performance levels. Among the most difficult aspects of a second language to be mastered is its sound system. It is very difficult (maybe impossible) to both perceive and produce a foreign language with a native accent. Common experience indicates the prevalence of these difficulties, in particular when the second language is learned late in life. However, training studies with adult participants have shown that significant improvements are possible with very low proficiency, or even monolingual, participants (see, for reviews, Sebastián-Gallés, 2005; Sebastián-Gallés & Kroll, 2003; Strange, 1995). The existence of brain plasticity for speech sounds has been attested in a range of studies addressing the question of learning

difficult L2 phonemic contrasts. In most of these studies, along with behavioral measures, the mismatch negativity (MMN) component has been used as an index of phonological discrimination. This measure shows a high sensitivity to the physical properties of the stimuli and, interestingly, it is observed in the absence of conscious realization of the contrast. Several studies have revealed significant differences (an increase in the amplitude) when the MMN is elicited in the presence of a linguistic (phonological) contrast, when compared with a foreign linguistic phonetic contrast (Sharma & Dorman, 2000; Näätänen et al., 1997; Phillips et al., 1995). In addition, several training studies have shown that electrophysiological changes can be induced through short but intensive programs (McClelland, Fiez, & McCandliss, 2002; Tremblay, Kraus, & McGee, 1998; Tremblay, Kraus, Carrell, & McGee, 1997). One aspect of the Tremblay, Kraus, and McGee (1998) study was that significant changes in the electrophysiological signatures were observed, before any behavioral improvement could be detected. This result could be taken as an indication that although no behavioral discriminations may be observed for nonnative contrasts, they may actually be detected at a subconscious level, and so behavioral measures may

---

<sup>1</sup>GRNC, Parc Científic Universitat de Barcelona & Hospital Sant Joan de Déu, Spain, <sup>2</sup>Universitat de Barcelona, Spain, <sup>3</sup>Institució Catalana de Recerca i Estudis Avançats (ICREA), Spain, <sup>4</sup>École Normale Supérieure, France

not be sensitive enough to reveal the real sensitivity to nonnative contrasts.

However, these training studies present certain limitations, and any generalization to normal bilingual processing should be made with caution. First, these studies used synthetic stimuli in highly simplified linguistic environments; in most of the cases, isolated syllables were employed. Second, participants were induced to pay attention to very low level acoustic properties that may go unnoticed in the normal course of speech perception (see Toro, Sinnett, & Soto-Faraco, 2005, for the importance of attention in speech perception tasks). Although it is vital to acquire reliable knowledge about the perceptual basis of speech processing, it has to be remembered that the ultimate goal of the language system is to efficiently extract the meaning of utterances. Thus, it is of utmost importance to gather knowledge about the consequences of these potentially difficult initial nonnative phonemic discriminations. To our knowledge, no electrophysiological data are available on the consequences of these difficulties at the level of the lexicon. However, some behavioral data regarding this issue are available (Weber & Cutler, 2004; Marian & Spivey, 2003; Schulpen, Dijkstra, Schriefers, & Hasper, 2003; Pallier, Colomé, & Sebastián-Gallés, 2001; Spivey & Marian, 1999). Taken together, these studies indicate that less precise prelexical processing in L2 leads to the activation of more candidates for consideration in the L2 lexicon and thus to less efficient lexical access (although see Ju & Luce, 2004).

A second related issue, also dealing with the plasticity of the speech system, is how the first language may be modified by exposure to a particular dialect. Indeed, it is commonly reported that natives, after living for some time in another community, change the way they speak their own language. For instance, it is commonly observed that American speakers who have spent some time in Britain often speak with a British accent on their return to the United States. The few reports of this issue have referred mostly to perception of dialect variation, namely, identifying dialects (see, for a review, Clopper & Pisoni, 2005). However, it would be reasonable to expect the L1 speech processing system to be modified after relatively extensive exposure to a “foreign” dialect. Recently, we have reported some behavioral data addressing these two related questions.

Sebastián-Gallés, Echeverría, and Bosch (2005) used a lexical decision task to address the issues of the phonological representation of L1 and L2 words. In this study, Catalan-dominant and Spanish-dominant Catalan-Spanish bilinguals were asked to make lexical decisions in response to Catalan stimuli. Crucially, nonwords were made by changing one vowel from an existing Catalan word. In some cases, the vowel change involved a difficult contrast for Spanish natives (experimental stimuli). Catalan and Spanish are two Romance languages differ-

ing in the number of vowels; whereas Spanish is a five-vowel system (/a, e, i, o, u/), Catalan has eight vowels (/a, e, ε, i, o, ɔ, u, θ/). Past research (Pallier, Colomé, & Sebastián-Gallés, 2001; Bosch, Costa, & Sebastian-Gallés, 2000; Sebastián-Gallés & Soto-Faraco, 1999; Pallier, Bosch, & Sebastián, 1997) has shown that some Catalan-specific contrasts (like the /e-ε/ contrast) are particularly difficult for Spanish natives to perceive. In the Sebastián-Gallés et al. study, the Catalan word “finestra,” meaning “window,” pronounced /finestrə/ was transformed into \*/finestrθ/. In other cases, the change involved a vowel contrast common to Spanish and Catalan, and was thus easily perceived by both groups of bilinguals (for instance, the Catalan word “cadira” pronounced /kədirə/, meaning “chair” was transformed into \*/kədurə/, the contrast /i-u/ being common to both languages). The results showed that although Spanish-dominant bilinguals had no problems in rejecting nonwords made by changing a common vowel contrast, they had substantial difficulty when the change involved a Catalan-specific one. Indeed, they showed a strong bias toward considering experimental nonwords as real words. There are two possible explanations for the high percentage of misidentifications in the experimental stimuli by Spanish-dominant bilinguals. First, these bilinguals really mapped the experimental nonwords into a single lexical entry; that is, after prelexical analyses, they both processed /finestrə/ and \*/finestrθ/ as homophones and, consequently, they treated both stimuli as real words (allophones). The second possibility is that participants were unconsciously able to detect the differences between experimental words and nonwords, but that the differences were too subtle to allow for conscious decisions. Under time pressure, as in Sebastián-Gallés et al., participants may have accepted many pseudowords as real words. In this study, it was not possible to distinguish between the two possibilities. Recently, it has been shown that event-related potentials (ERPs) can be effectively used to uncover unconscious differential processing of L2 words and nonwords. Indeed, McLaughlin, Osterhout, and Kim (2004) obtained different ERP signatures for second-language words and nonwords, even when participants were at chance levels in a behavioral lexical decision task.

In the Sebastián-Gallés et al. (2005) study, it was also observed that the performance of Catalan-dominant participants was slightly impaired with stimuli, where the change involved the /e-ε/ contrast, than with common contrasts. The explanation given for this decrease in performance in L1 processing was exposure to the Catalan “dialect” that Spanish-dominant bilinguals produce. Many Spanish-dominant bilinguals use the phonology of Spanish when producing Catalan words; that is, they do not reproduce the /e-ε/ contrast. In this situation, Catalan natives may have two representations for some stimuli (the “correct” Catalan one and the

mispronunciation of the dialect of Spanish natives). The use of ERP recordings may help to clarify this possibility.

### **ERPs: The N400 and the Error-related Negativity**

One way of addressing the present questions is by measuring scalp recorded ERPs that are used as an on-line direct manifestation of brain activity with millisecond temporal resolution (Müntz, Urbach, Düzel, & Kutas, 2000) and allow a chronometric analysis of language processes (Kutas & Federmeier, 2000; Kutas, 1997; Osterhout, McLaughlin, & Bersick, 1997). Two interesting components will be explored in the present study: (a) the N400 component, which is sensitive to lexical activations and (b) the error-related negativity (ERN), which is sensitive to the level of uncertainty when responding.

The first component, the well-known N400, is an ERP index that is regarded as sensitive to meaning integration and semantic processing. For example, when a semantically incongruous word is presented at the end of a sentence, a larger N400 component is elicited than if the final words are congruous (McCallum, Farmer, & Pocock, 1984; Kutas & Hillyard, 1980). In lexical decision tasks, an increased N400 is also encountered for pronounceable nonwords compared to words and consonant strings when presented in isolation or in pairs (Ziegler, Besson, Jacobs, Nazir, & Carr, 1997; Chwilla, Brown, & Hagoort, 1995; Holcomb, 1993, 1988; Bentin, 1987; Rugg & Nagy, 1987; Smith & Halgren, 1987; Bentin, McCarthy, & Wood, 1985). In the auditory domain, Holcomb and Neville (1990) have also shown that only nonwords (pseudowords) that have similar orthographic and phonological characteristics to real words elicit an N400, but that this effect is not encountered with unpronounceable nonwords. The same result was obtained in the visual domain by Ziegler, Besson, et al. (1997). This N400 lexicality effect has been interpreted as an index of the lexical search process, in which an orthographical and pronounceable nonword is compared to possible lexical candidates in order to find associated lexical-semantic information in long-term memory. The N400 component has also been considered to reflect the degree of lexical-semantic activation. Holcomb, Grainger, and O'Rourke (2002) observed that words and nonwords with many lexical neighbors generated larger N400 than similar items with relatively fewer lexical neighbors. In this view, the closer a nonword is to a real word (the more "unique" the activation of a lexical candidate by a nonword), the smaller the difference between words and nonwords in the N400 component should be.

In the context of the present research, it is expected that Spanish-dominant bilinguals, because of their lack of perceptual discrimination between the /e-ε/ Catalan contrast, will show no differences in the N400 component for experimental words and nonwords because both acoustic stimuli will activate the same lexical rep-

resentation. The predictions for the Catalan-dominant bilinguals are less clear. In principle, these bilinguals should show larger N400 for nonwords. Nevertheless, as mentioned above, it might be the case that the extended exposure to the particular Catalan dialect of Spanish-dominant bilinguals (in which the e-ε contrast is neutralized), as suggested in Sebastián-Gallés et al. (2005) may have the consequence of creating secondary phonological representations; if this occurs, experimental nonwords could also activate the corresponding word. In this case, we would also expect reduced or no differences in the N400 component for experimental words and nonwords for this population. However, the behavioral data of Sebastián-Gallés et al. showed that these participants made few errors with experimental stimuli. So, if no significant differences in lexical activation between words and nonwords (as indexed by the N400 component) were observed between the two groups of bilinguals, but both populations clearly show different error patterns, then differences should arise in other processing stages. Given the nature of the lexical decision task, a good candidate for assessment is the decision stage, that is, the moment when participants decide if the stimulus presented is a real word or not.

Different response-locked ERPs, which are specific components associated with the onset of response-related processes (Osman, Moore, & Ulrich, 1995) have been previously described. Of particular interest is the ERN (Falkenstein, Hohnsbein, & Hoormann, 1995; Gerrig, 1993). The peak amplitude of the ERN is normally observed approximately 60–100 msec after an erroneous response. Its amplitude is increased when accuracy is stressed (Gerrig, 1993) and with error awareness (Scheffers & Coles, 2000). Although the ERN was initially associated to the conscious detection of errors (Christ, Falkenstein, Heuer, & Hohnsbein, 2000; Falkenstein, Hoormann, Christ, & Hohnsbein, 2000; Luu, Collins, & Tucker, 2000; Scheffers, Coles, Bernstein, Gehring, & Donchin, 1996; Falkenstein, Hohnsbein, & Hoormann, 1995; Gehring, Coles, Meyer, & Donchin, 1995) this hypothesis has been challenged by Nieuwenhuis, Ridderinkhof, Blom, Band, and Kok (2001). Interestingly for the goals of the present research, these authors showed that the ERN was present even in the case of unperceived errors.

Although the ERN was first considered to be associated only with erroneous processing, several studies have already observed ERN-like components associated with correct responses (Rodríguez-Fornells, Kofidis, & Münte, 2004; Nessler & Mecklinger, 2003; Swick & Turken, 2002; Luu et al., 2000; Scheffers & Coles, 2000; Vidal, Hasbroucq, Grapperon, & Bonnet, 2000; but see Coles, Scheffers, & Holroyd, 2001). The ERN has a focal midline frontocentral maximum and probably arises from the anterior cingulate gyrus (ACG) (Dehaene, Posner, & Tucker, 1994) with additional contributions from the lateral prefrontal cortex (van Veen & Carter,

2002; Luu & Tucker, 2001). Functional magnetic resonance imaging studies (fMRI) have localized error processing in the anterior cingulate and lateral inferior frontal cortex extending to the bilateral insular cortex (see Laurens, Ngan, Bates, Kiehl, & Liddle, 2003; Garavan, Ross, Murphy, Roche, & Stein, 2002; Carter, MacDonald, Ross, & Stenger, 2001; Menon, Adelman, White, Glover, & Reiss, 2001; Ullsperger & von Cramon, 2001; Kiehl, Liddle, & Hopfinger, 2000; Carter, Braver, et al., 1998).

An alternative account of the ERN component has also been put forward. The conflict monitoring account of the ERN considers that this component indexes the amount of response conflict (Yeung, Botvinick, & Cohen, 2004). This could explain the presence of ERN-like activity also in correct trials in certain studies (Carter, MacDonald, et al., 2001; Luu et al., 2000; Scheffers & Coles, 2000; Vidal et al., 2000), when responses involved a higher amount of motor conflict (Gehring & Fencsik, 2001; van Veen, Cohen, Botvinick, Stenger, & Carter, 2001; Barch, Braver, Sabb, & Noll, 2000) and when the on-line corrective response is being prepared (Rodríguez-Fornells, Kurzbuch, & Münte, 2002). This conflict monitoring account contrasts with the original interpretation of the ERN component according to which the ERN is the outcome of a mismatch process elicited when detecting that the actual and the intended responses do not coincide. In contrast, the conflict proposal considers that the ERN reflects the conflict that develops in the period following the errors as a consequence of the continued processing of information, which at the same time activated the correct response that competes with the incorrect response. Regardless of the origin of the ERN component, in the context of the present study it is viewed as an index of uncertainty in lexical decisions (Scheffers & Coles, 2000). When participants are fairly sure that their response is an error, a clear ERN-like component would be expected. In contrast, for insecure or doubtful decisions, the differences between correct and error responses would be less evident. Although to our knowledge the ERN component has not been used in language research, some studies have already been presented in the memory domain, specifically for old/new decisions (Rodríguez-Fornells, Kofidis, & Münte, 2004; Nessler & Mecklinger, 2003). An interesting finding in these studies is that correct recognition of old words elicited a larger ERN than correct rejections of new words. These results suggest that the ERN component may be modulated regardless of the correctness status of the response emitted. In the present context, the ERN-like component will be explored for the first time in a lexical decision task. Interestingly, although no clear differences for experimental words and nonwords between Spanish- and Catalan-dominant bilinguals might be observed in the N400 component, the populations may differ in terms of the ERN component. If, as stated above, Spanish-dominant bilinguals really activate the same lexical entry from experimental words and nonwords

(/finestrə/ and \*/finestrə/ would both activate the same lexical entry), no differences would be expected for correct and erroneous responses in this group of participants. In contrast, Catalan-dominant bilinguals might show a clear modulation of the ERN component regarding its correctness because they would be able to detect whether or not the proper pronunciation was presented.

## METHODS

### Participants

Thirty right-handed Spanish–Catalan bilinguals participated in this experiment. All were born in Catalonia (most of them in Barcelona or its metropolitan area). Half were raised as Spanish monolinguals until the age of three at the latest (when schooling started). During the first years of their lives, they had only occasional contact with Catalan. The other half was the mirror image, with Catalan as their first language. All participants had received a bilingual education and claimed to be very fluent in the two languages in both listening and reading. Furthermore, all participants had passed the mandatory examination to enter Spanish universities in Catalonia, meaning that they had proven their proficiency not only in oral and written skills but also in their formal knowledge of both Catalan and Spanish phonology, morphology, and syntax (the requirements are the same as those that monolingual Spanish students have to meet in Spanish before entering any Spanish university). Only participants who reported that their current dominant language was the language learned at home participated in this experiment.

Participants were undergraduate Psychology students at the University of Barcelona, where, according to official statistics, 60% of the courses are taught in Catalan. Participants received economic compensation for their participation in the experiment. None of the participants reported any hearing problems. Two participants, one from each group, were rejected due to excessive ocular movements. In the final sample there were 21 women (11 in the Catalan-dominant group and 10 in the Spanish-dominant group) and 7 men (3 in the Catalan-dominant group and 4 in the Spanish-dominant group). Age range was 19–24 years.

### Materials

One hundred twenty Catalan words containing the vowel /e/ and 120 Catalan words containing the vowel /ɛ/ were selected. Most words were nouns and a few were verbs (18 verbs in their citation form—infinite) were also included. Words varied in length (from one to four syllables). Both types of words were equal in terms of frequency (written word frequency per million for e words: 102.98, *SD* 205.88; for ε words: average 172.07, *SD* 510.33; *t* test, *p* < .17) (Rafel i Fontanals, 1998),

length (for e words, average 656 msec, *SD* 119; for ε words, average 649 msec, *SD* 105; *t* test,  $p < .39$ ) and loudness (for e words, average RMS power:  $-20.16$  dB, *SD* 1.35; for ε words, average:  $-20.42$  dB, *SD* 1.02; *t* test,  $p < .12$ ). The corresponding nonwords were created by replacing the vowel /e/ with /ε/ and vice-versa. Thus, for example, the word “galleda” (meaning “bucket”), pronounced /gəλɛdə/, generated the nonword \*/gəλedə/, and the word “ulleres” (meaning “glasses”), pronounced /uλerəs/, generated the nonword \*/uλerəs/. Because Catalan has vowel reduction, /e/ and /ε/ can only occur in stressed positions, these changes were restricted to stressed syllables. Nonwords were also equated in length (for e nonwords, average 644 msec, *SD* 121; for ε nonwords, average 637 msec, *SD* 101; *t* test,  $p < .46$ ) and loudness (for e nonwords, average RMS power  $-19.92$  dB, *SD* 1.33; for ε nonwords, average  $-19.84$  dB, *SD* 1.32; *t* test,  $p < .48$ ; both nonwords were also matched with their corresponding words; *t* test, both  $ps > .3$ ). The percentage of noncognate words containing vowel /e/ was 49.16% and of words containing vowel ε was 53.33% (these two percentages were not significantly different). Although not all words were noncognates, no stimuli sounded the same in both languages, even accepting some perceptual assimilations made by Spanish speakers (such as perceiving Catalan /ə/ as /a/). For instance, the Catalan word “convent” is cognate with the Spanish word “convento,” but the pronunciation (mostly because of vowel reduction in Catalan) makes the acoustic realization of the two words quite different: /kumβen/ in Catalan and /komβento/ in Spanish. Even in some cases in which the number of phonemes and orthography is the same in both languages, for instance, the word “portera” (meaning “doorman,” feminine), the words are not pronounced the same in the two languages, /purterə/ in Catalan and /portera/ in Spanish.

One hundred twenty control words were selected. These words also varied in length between one and five syllables (average length 629 msec, *SD* 112). They were also equated in frequency (average 185.82, *SD* 377.5) and loudness (average RMS power:  $-20.21$  dB, *SD* 1.32) with experimental words. One hundred twenty nonwords were created from words similar to the controls by replacing one vowel. These words did not contain vowels /e/ or /ε/. In addition, in this case, the vowel replacement always corresponded to the stressed one. An example of control words and their corresponding pseudowords is the word “llençol” (meaning “sheet”), pronounced /λənsol/, which had a vowel changed to make \*/λənsal/. Control nonwords did not differ from experimental ones or from control words in length (average 627 msec, *SD* 124) and loudness (average RMS power:  $-20.327$  dB, *SD* 1.32). Twenty-five percent of the control words were cognates. Taking together all conditions, a total of 360 nouns and verbs of mid to high frequency were used in the present experi-

ment. Two lists were created: in each list, half of the words and half of the nonwords appeared. Of each word–nonword pair, one of the members appeared in List 1 and the other in List 2. Order of presentation of stimuli within each list was randomized for each subject. Stimuli (recorded by a male native speaker of Catalan, aged 25) were digitalized and down-sampled to 16 kHz. All stimuli were recorded in a single session. Two experienced Catalan-native independent judges checked for the correct pronunciation of every stimulus. Stimuli were edited with Cool Edit (Syntrillium Software Corp., Phoenix, AZ), and individual stimuli files were created for each word. No silences were left at the beginning or end of each file.

### Procedure

Participants were asked to perform a lexical decision task on the stimuli presented. Half of the participants were asked to press a button with their right hand whenever they heard a word and to press a button with their left hand if they heard a nonword. Response hand assignments were changed for the other half of the participants. They were instructed to respond as fast as possible, but to try to avoid errors. The instructions specified that changes always involved vowels and that in many cases they involved the vowels /e/ and /ε/.

The participants were tested in two sessions and with both lists (half of them started with List 1 and the other half with List 2). The presentation of the stimuli was controlled by personal computers equipped with Pro-Audio 16 sound cards. Auditory stimuli were presented through Sennheiser HMD224x headphones. The experimental situation was controlled by the program EXPE (Pallier, Dupoux, & Jeannin, 1997). Reaction times were measured from stimuli onset.

### Electrophysiological Recording

The ERPs were recorded from the scalp using tin electrodes mounted in an electrocap (Electro-Cap International, Eaton, OH) and located at 29 standard positions (Fp1/2, Fz, F7/8, F3/4, Fc1/2 Fc5/6, Cz, C3/4, T3/4, Cp1/2, Cp5/6, Pz, P3/4, T5/6, Po1/2, O1/2). Biosignals were re-referenced off-line to the mean of the activity at the two mastoid processes. Vertical eye movements were monitored with an electrode at the infraorbital ridge of the right eye. Electrode impedances were kept below 5 kΩ.

The electrophysiological signals were filtered with a band pass of 0.01–50 Hz (half amplitude cutoffs) and digitized at a rate of 250 Hz. Trials with base-to-peak electrooculogram (EOG) amplitude of more than 50 ZV, amplifier saturation, or a baseline shift exceeding 200 ZV/sec were automatically rejected off-line. No significant differences were observed for the percentage of rejected trials in the groups ( $p > .13$ ; 9.3% in the Spanish-dominant group and 13.6% in the Catalan-dominant group).

## Data Analyses

Stimulus-locked ERPs were averaged for epochs of 1700 msec starting 200 msec prior to the stimulus. In addition, response-locked averages starting 400 msec before and extending 624 msec beyond the button-press responses were obtained. The baseline used for response-locked averages was  $-200$  until  $-50$  msec before the onset of the response. All response-locked averages and topographical maps were band-pass filtered (1–8 Hz half amplitude cutoff) in order to compensate for the positive trend in which they are superimposed (Rodríguez-Fornells, Kurzbuch, & Münte, 2002).

Several repeated measures analyses of variance (ANOVAs) were conducted for the evaluation of stimulus-locked ERPs (specified in each case in the Results section) and including Lexicality (word, non-word), Stimulus Type ( $\epsilon$ ,  $e$ , control words) and Electrode (midline locations, Fz, Cz, Pz). All statistical tests comprised mean amplitudes for the different time windows specified in the corresponding contrast. In addition, in all ANOVAs, Bilingual Group (Spanish-dominant, Catalan-dominant) was introduced as a between-subjects factor. For the resulting interactions including Group, Lexicality, or Stimulus Type, additional ANOVAs were carried out, which were restricted to specific electrode sites. The same analyses were carried out at parasagittal and temporal electrode locations. These results did not differ from those found at midline locations.

A similar statistical evaluation was carried out in the response-locked ERPs, focusing on the ERN (mean amplitude in the interval between 0 and 100 msec after the response) and restricted to midline locations, where the ERN is known to be maximum (especially with fronto-

central electrodes). In the ANOVAs, the Type of Trial (correct vs. error) was introduced as a within-subject factor. ERNs were accurately computed due to the large amount of errors committed in both groups of subjects in the nonword decisions ( $\sim 63\%$  for Spanish-dominant and  $\sim 33\%$  for Catalan-dominant). Also, the correct-related negativity (CRN) (Falkenstein, Hoormann, et al., 2000; Vidal et al., 2000) was analyzed, which is an ERN-like component associated with correct trials. Although a fixed baseline was used for all the statistical computations of the ERN (see above), the results presented were contrasted with different baselines in order to assess the stability of the results (Picton et al., 2000).

For all statistical effects involving two or more degrees of freedom in the numerator, the Greenhouse–Geisser epsilon was used to correct for possible violations of the sphericity assumption (Jennings & Wood, 1976). The exact  $p$  value after the correction is reported. Tests involving Electrode  $\times$  Condition interactions were carried on data corrected using the vector normalization procedure described by McCarthy and Wood (1985).

## RESULTS

### Behavioral Performance

#### Percentage of Errors

Error rates for the  $\epsilon$  and  $e$  types were very high, in particular for the Spanish-dominant bilinguals (see Table 1). These participants showed a strong bias to respond to both words and nonwords as if they were real words. Because of this response bias, it was decided to carry out the accuracy analyses using the  $A'$  statistics (a non-

**Table 1.** Performance of Both Groups in the Auditory Lexical Decision Task

	Words			Nonwords		
	$\epsilon$ Type "galleda"	$e$ Type "finestra"	Control "llençol"	$\epsilon$ Type "galleda"	$e$ Type "finestra"	Control "llençal"
<i>Percentage errors (SEM)</i>						
Spanish-dominant	6.7 (1.3)	7.2 (1.1)	6.2 (1.0)	69.9 (5.3)	63.3 (7.5)	6.3 (0.9)
Catalan-dominant	3.9 (0.7)	2.9 (0.5)	2.3 (0.8)	33.8 (7.0)	17.9 (4.3)	5.2 (1.3)
<i>F</i> value	<i>ns</i>	12.1*	7.8*	16.8**	27.2**	<i>ns</i>
<i>Reaction time correct responses (SEM)</i>						
Spanish-dominant	1186 (33)	1167 (26)	1132 (30)	1344 (40)	1372 (30)	1196 (28)
Catalan-dominant	1102 (37)	1085 (40)	1054 (38)	1147 (26)	1232 (26)	1123 (33)
<i>F</i> value	<i>ns</i>	<i>ns</i>	<i>ns</i>	16.8**	12.5*	<i>ns</i>

Values are percentages of errors and reaction times (milliseconds) and standard error of the mean (in parentheses). In all cases,  $F(1,26)$ . *ns* = nonsignificant.

\* $p < .01$ .

\*\* $p < .001$ .

parametric unbiased index with 0.5 indicating response at chance level and 1 perfect discrimination; McNichol, 1972). Averages for each stimulus type and group are displayed in Table 2.

An ANOVA was performed for the  $A'$  values introducing Stimulus Type ( $\epsilon$  type,  $e$  type, and control) as within-subject factor and Group as a between-subjects factor. All main effects were significant: Group,  $F(1,26) = 38.5$ ,  $p < .001$ ; Stimulus Type,  $F(2,52) = 59.2$ ,  $p < .001$ , as well as the corresponding interaction,  $F(2,52) = 19.7$ ,  $p < .001$ . Catalan-dominant bilinguals performed better than Spanish-dominant ones in all stimulus types (all  $p < .001$ )

### Reaction Times

The same ANOVA design was used for correct responses (range 200–2000 msec). The Catalan-dominant group was faster than the Spanish-dominant group: Group,  $F(1,26) = 7.3$ ,  $p < .02$ ; mean reaction time for Catalan-dominant bilinguals = 1121 msec; Spanish-dominant = 1236 msec; see also Table 1). Nonwords were faster than words: Lexicality,  $F(1,26) = 51.30$ ,  $p < .01$ ; words 1121, nonwords 1236. A main effect of stimulus type was also observed,  $F(2,52) = 41.7$ ,  $p < .01$ , which was modulated by the group, Group  $\times$  Type,  $F(2,52) = 4.6$ ,  $p < .02$ . The interaction Lexicality  $\times$  Stimulus Type was also significant,  $F(2,52) = 9.5$ ,  $p < .01$ , as well as the interaction between Group  $\times$  Lexicality  $\times$  Stimulus Type,  $F(2,52) = 4.19$ ,  $p < .05$ . Direct between-group comparisons (see Table 1) showed that  $e$  type and  $\epsilon$  type nonwords showed the largest differences between groups.

Furthermore, group effects on the mean reaction time for erroneous responses were inspected only for nonwords. No differences between groups were observed in any case:  $e$  type, Spanish-dominant =  $1264 \pm 40$  and Catalan-dominant =  $1248 \pm 35$ ,  $F(1,26) < 1$ ;  $\epsilon$  type, Spanish-dominant =  $1230 \pm 44$  and Catalan-dominant =  $1239 \pm 32$ ,  $F < 1$ ; control, Spanish-dominant =  $1290 \pm 41$  and Catalan-dominant =  $1367 \pm 56$ ,  $F = 1.2$ .

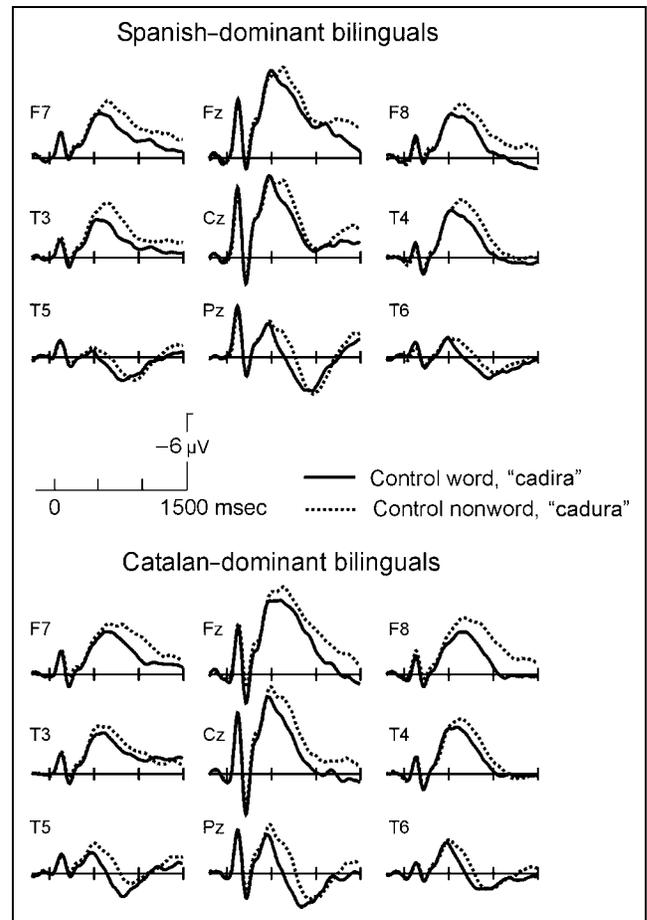
### Stimulus-locked ERPs

The stimulus-locked ERPs elicited for the control words and nonwords in both groups are illustrated in Figure 1.

**Table 2.** Mean  $A'$  and Standard Error of the Mean (in Parentheses) for Each Group of Bilingual and Each Stimulus Type

	$\epsilon$ Type	$e$ Type	Control
Spanish-dominant	.695 (.030)	.710 (.037)	.953 (.005)
Catalan-dominant	.887 (.020)	.942 (.010)	.982 (.003)
<i>F</i> value	26.9	34.9	18.3

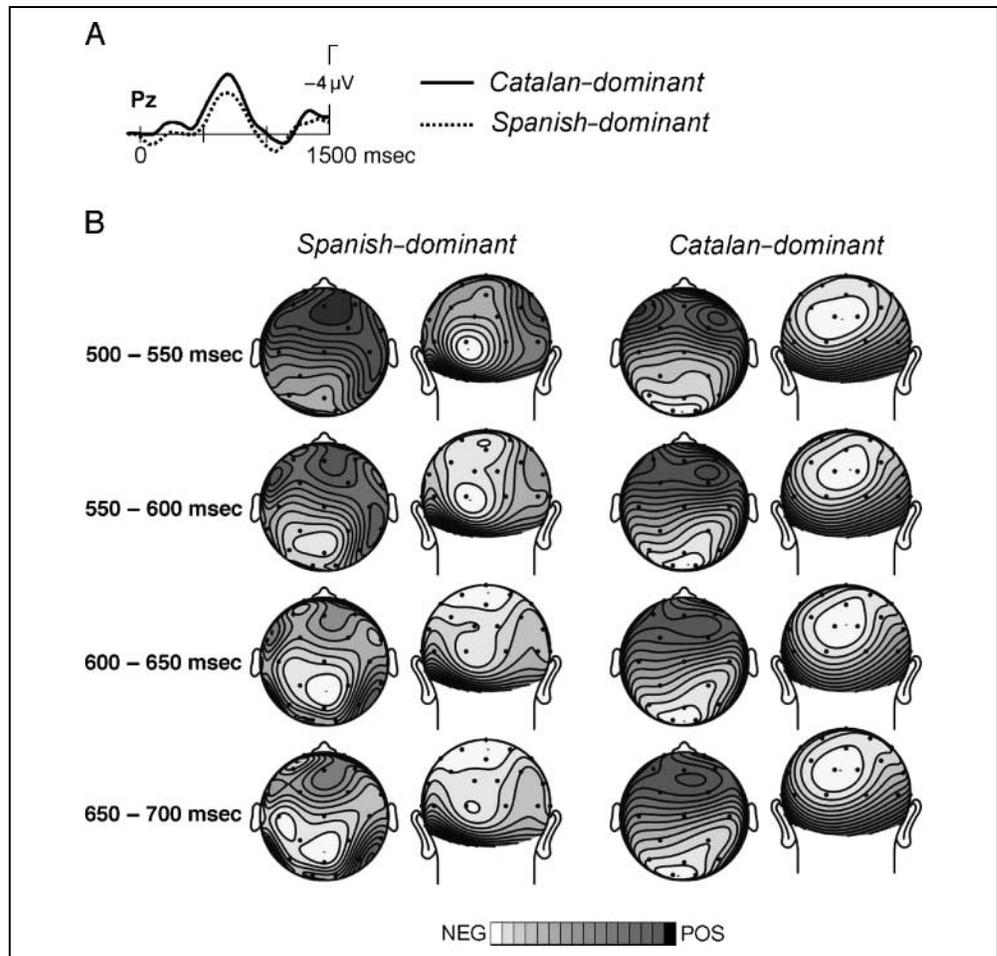
In all cases,  $F(1,26)$  and  $p < .001$ .



**Figure 1.** Stimulus-locked ERPs synchronized to the onset of the stimulus. ERPs for both bilingual groups illustrate the differences between word and nonword correct decisions (lexicality effect) in the control words. Notice the larger and broadly distributed negativity elicited by nonwords. Midline and temporal electrode locations are shown.

A classical centroparietal N100–P200 was observed, followed by a widespread negativity peaking at about 500–600 msec at central and frontal sites. This negative component is larger for nonword decisions and corresponds to the N400 enhancement observed for nonword decisions. An ANOVA was performed for this control condition at midline (Fz, Cz, Pz) locations for the time window 500–750 msec. The mean amplitude of the N400 component was larger for nonwords than for control words in all ANOVAs,  $F(1,26) = 12.7$ ,  $p < .01$ . No significant differences were found for Group and the corresponding interactions of this variable and the other factors (in all cases,  $F < 1$ ). The N400 enhancement was larger at parietal locations: Lexicality  $\times$  Electrode,  $F(2,52) = 7.9$ ,  $p < .01$ . Because the N400 lexicality effect was observed as a long-lasting component in some electrodes, a further analysis was conducted on a larger time window (500–1500). The increased negativity attributed to nonwords extended significantly to 1500 msec,  $F(1,26) = 7.89$ ,  $p < .01$ . The

**Figure 2.** (A) ERP difference waveforms at parietal electrode location showing the subtraction between nonwords minus words in the control condition and for both bilingual groups. (B) Topographical maps for the difference waveforms created using isovoltage spline interpolation in the time windows indicated in each row. This topography showed a parieto-occipital distribution slowly shifting to centroparietal sites. Minimum and maximum values of the maps: Spanish-dominant (from top to bottom):  $-1.1/0.5 \mu\text{V}$ ,  $-1.4/0.8$ ,  $-1.8/1.2$ ,  $-1.8/-1.4$ ; Catalan-dominant,  $-1.5/0.5$ ,  $-1.9/0.2$ ,  $-2.4/0.2$ ,  $-2.8/0.3$ .

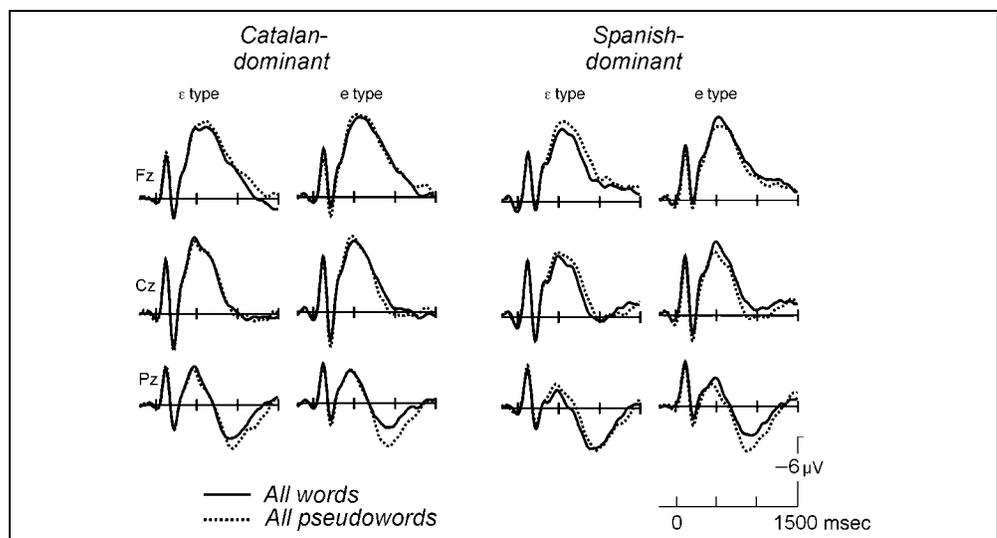


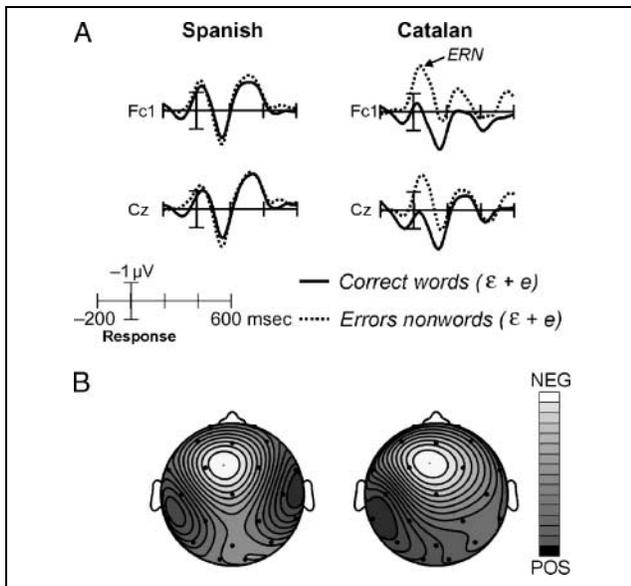
difference waveforms (nonword minus word) and their scalp distribution are depicted in Figure 2 for both groups. Notice the predominant occipitoparietal distribution of this component. Further analysis performed on the peak and latency of the N400 component showed

no significant differences between groups in these measures.

The grand averages for the two word and nonword experimental conditions ( $\epsilon$  and  $\varepsilon$  types) for the two groups are shown in Figure 3. Because of the small

**Figure 3.** ERPs elicited by  $\epsilon$  type and  $\varepsilon$  type word and nonword stimuli in both groups. Correct and incorrect trials have been collapsed in these averages. Midline electrode locations are depicted.





**Figure 4.** (A) Response-locked ERPs for both bilingual groups depicted for frontocentral electrodes and for correct words ( $\epsilon$  type +  $e$  type) and errors produced when responding to nonwords ( $\epsilon$  type +  $e$  type). Notice the increased negativity just after the commission of the responses (R) (labeled ERN component). The Catalan-dominant group showed a clear differentiation between correct and erroneous responses, which was not present in the Spanish-dominant group. In both conditions, subjects decided that the stimulus presented was a Catalan word. (B) Scalp distribution for the ERN component that appeared in the errors for nonwords ( $\epsilon$  type +  $e$  type) represented in (A). Topographical maps were created using isovoltage spline interpolation for the 30- to 80-msec interval. This topography showed the classical frontocentral negativity of the ERN component. Notice that relative scaling was used. Maximum and minimum values for each isovoltage map: Spanish group,  $-1.33 \mu\text{V}/0.54$ ; Catalan group,  $-2.75/0.27$ .

number of correct responses in the Spanish-dominant group, both correct and incorrect responses were included (the analyses of correct responses were performed for the Catalan-dominant group; no differences between these analyses and the ones reported here were observed). This analysis clearly suggests a similar ERP pattern for both language groups. An ANOVA was computed introducing Group and Stimulus Type ( $e$  type and  $\epsilon$  type) and Lexicality (word/nonword). At the N400 time window (500–750 msec) and for midline electrode locations no significant main effects were encountered: in all cases,  $F(1,26) < 0.5$ . Its corresponding interactions were not significant either: Group  $\times$  Lexicality and Group  $\times$  Stimulus Type, in both cases  $F < 1$ ; Lexicality  $\times$  Stimulus Type and Group  $\times$  Lexicality  $\times$  Stimulus Type, in both cases  $F(1,26) < 2.4$ ,  $p > .14$ . This pattern of results thus shows that pseudowords did not elicit a N400 lexicality effect in either group.

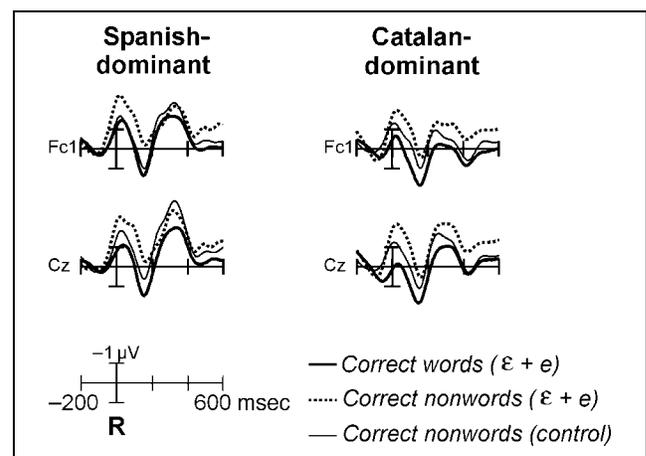
Also, inspection of Figure 3 showed that  $e$  type and  $\epsilon$  type nonwords elicited a late positive component (LPC) with a peak around 850–900 msec, which was

not present for the control words (see Figure 1). To encompass this positive component, an ANOVA was performed on the mean amplitude (800- to 1000-msec time window) in the experimental conditions at midline locations. Neither significant main effects, Group,  $F < 1$ ; Lexicality,  $F(1,26) = 1.13$ ; Stimulus Type,  $F < 1$ , nor the interactions Group  $\times$  Lexicality and Group  $\times$  Stimulus Type ( $F < 1$  in both cases) yielded significant results. A significant interaction appeared between Lexicality and Stimulus Type,  $F(1,26) = 4.3$ ,  $p < .5$ , and Lexicality and Electrode,  $F(2,52) = 10.3$ ,  $p < .002$ . These interactions reflect the increased LPC at parietal locations for  $e$  type nonwords in both groups compared to the  $\epsilon$  type nonwords.

### Response-locked ERPs

A classical ERN component was obtained in different kinds of trials for both groups (see Figures 4 and 5). An ANOVA was performed introducing Type of Trial (correct trials,  $\epsilon$  type and  $e$  type stimuli vs. erroneous responses, which were the nonwords misclassified as words) and Group at midline locations on the mean amplitude between 0 and 100 msec. A significant interaction was obtained between Group and Type of Trial,  $F(1,26) = 10.3$ ,  $p < .1$  (see Figure 4).

A main effect of Type of Trial was also obtained,  $F(1,26) = 9.36$ ,  $p < .1$ . The interaction between Group and Type of Trial reflects that for Catalan subjects the production of an erroneous response elicited a clear increase in the ERN component,  $F(1,13) = 14.2$ ,  $p < .1$ ; correct trials,  $1.6 \pm 2.5 \mu\text{V}$ ; erroneous trials,  $-0.4 \pm 2.6 \mu\text{V}$ . In contrast, in the Spanish group, no clear differences



**Figure 5.** Response-locked ERPs for both bilingual groups depicted for frontocentral electrodes. Three correct types of responses are illustrated showing a clear modulation of the ERN component (known as correct-related negativity, CRN). Notice the increased CRN in nonword  $\epsilon$  type +  $e$  type decisions compared to the control nonwords and words ( $\epsilon$  type +  $e$  type).

appeared between correct and erroneous responses,  $F(1,13) < 1$ ; correct trials,  $-0.66 \pm 2.3 \mu\text{V}$ ; erroneous trials,  $-0.61 \pm 1.85 \mu\text{V}$ . Notice that in both decisions subjects answered that the word or the nonwords were real words.

To test if Spanish subjects were responding at a higher degree of uncertainty, an ANOVA was also conducted introducing only correct trials for both  $\epsilon$  type and  $e$  type words and comparing the ERN at midline location (0–100 msec). A significant effect of group was obtained,  $F(1,26) = 8.37, p < .076$ . For the Spanish group, a larger ERN was obtained for correct trials ( $-0.66 \pm 2.4 \mu\text{V}$ ) compared to the Catalan group ( $1.58 \pm 2.6 \mu\text{V}$ ). No differences were obtained between  $\epsilon$  type and  $e$  type correct trials,  $F(1,26) < 1$ , or for electrode locations,  $F(1,26) = 2.64, p > .1$ . The other interactions between Group and Factors were also nonsignificant. The iso-voltage spline interpolated maps are also depicted for the errors in nonword conditions showing the classical frontocentral negativity attributed to the ERN component (Figure 4B).

Finally, in Figure 5, the response-related ERP components associated only with correct responses were studied. Correct responses associated with an experimental nonword decision (giving a negative response to a nonword) showed a clear ERN-like component in spite of the correctness of the response. This component has been observed previously and is termed correct-related negativity (CRN). An ANOVA conducted at midline locations for the time window 0–100 msec showed that correct nonword decisions ( $\epsilon + e$  type) increased the CRN component compared to correct word decisions ( $\epsilon + e$  type),  $F(1,26) = 17.2, p < .01$ ; correct word decisions,  $0.46 \pm 2.6 \mu\text{V}$ ; correct nonword decisions,  $-1.37 \pm 3.5 \mu\text{V}$ . The interaction between Type of Decision and Electrode Location was also significant,  $F(2,52) = 3.6, p < .5$ . The effect was larger at central locations.

The correct nonword decisions for the control condition are also depicted in Figure 5. Control nonword correct decisions had a reduced amplitude compared to correct experimental nonword decisions,  $F(1,26) = 6.34, p < .5$ . Correct control nonwords also tended to have a larger CRN than correct word decisions, although they failed to reach a significant value,  $F(1,26) = 2.9, p > .09$ ; correct control words mean amplitude,  $-0.28 \pm 3.2 \mu\text{V}$ .

A direct pairwise test of the amplitude of the experimental words ( $\epsilon + e$  type) and control nonwords comparing both CRNs in each group of participants showed a significant difference in the Catalan-dominant group,  $F(1,13) = 6.1, p < .284$ . In contrast, no differences were observed in the Spanish-dominant group,  $F(1,13) < 1$  (applied to midline locations at the 0- to 100-msec time window). Finally, when both correct nonwords ( $\epsilon + e$  type and control) were compared separately in each group, there was only a marginal effect for the Spanish-dominant group:  $F(1,13) = 3.6, p < .78$ ; for the Catalan-dominant group,  $F(1,13) = 2.7, p > .12$ .

## DISCUSSION

In the present study two groups of highly proficient Spanish–Catalan bilinguals were assessed using a lexical decision task in Catalan. The samples were carefully selected, bearing in mind the age of acquisition of, and proficiency in, the two languages. The phonological representation of Catalan words was evaluated using behavioral and ERP measures. Spanish-dominant bilinguals showed a large impairment in nonword decision when the nonwords were constructed by changing one vowel from an existing Catalan word that involved a difficult vowel contrast (i.e., /finestrə/ vs. \*/finestrə/: experimental stimuli). In contrast, this group had no difficulty with replacements involving a vowel contrast existing in both languages (e.g., /kədirə/ vs. /kədurə/: control stimuli). The major findings of the present study were as follows: (a) Catalan-dominant bilinguals did not show a standard N400 lexicality effect for the experimental nonword decisions, and (b) Spanish-dominant bilinguals did not show an ERN effect in experimental nonword decisions. In addition, a larger CRN was found in lexical decisions involving difficult decisions. In what follows we first discuss the results obtained and then consider the major implications of these findings in the context of bilingualism research.

### Words versus Nonwords: The N400 Component

The behavioral results showed that Spanish-dominant bilinguals had great difficulty in rejecting nonwords made by changing a single vowel involving a Catalan-specific contrast (and that they found difficult to perceive); however, these bilinguals had no difficulty in rejecting nonwords when the vowel change involved a common contrast. This pattern was observed both in the reaction times and in the error analyses. One important requirement for our present goals was that Spanish-dominant and Catalan-dominant bilinguals should not differ in terms of their general knowledge of the Catalan lexicon; otherwise, any potential difference between the groups with the experimental stimuli could be attributable to differences in their competence in Catalan. The results of the control condition clearly rule out this possibility. Indeed, both the behavioral and the ERP data show that the two bilingual populations were equivalent in this respect. Figures 1 and 2 show that control nonwords induced a classical enhanced N400 component compared to real words. More importantly, this effect and its topography did not differ across groups.

However, a different picture emerged when experimental stimuli ( $e$  type and  $\epsilon$  type) were analyzed. Spanish-dominant participants showed very high error rates in these conditions. In fact, there was a bias toward producing behavioral responses to experimental nonwords as if they were real words. The high error rates made it impossible to properly interpret the reaction

times. For the ERP analyses, we collapsed the correct and erroneous responses to compare both groups (Figure 3). In spite of large differences in the discrimination scores ( $A'$ ), the patterns of the N400 component were equivalent for both bilingual groups. This apparent contradiction can be explained by considering the N400 as an index of lexical activation. It is clear that for all participants, both experimental words and nonwords activated lexical entries. The lack of perception of the /e-ɛ/ contrast explains the data of the Spanish-dominant bilingual (also supported by the ERN data). The exposure to phonological variants explains the results of the Catalan-dominant bilinguals. As mentioned in the Introduction, the /e-ɛ/ contrast is allophonic in Spanish: that is, although there is a single /e/ category in Spanish, its acoustic realization depends on the phonetic environment, with the result that it approaches the values of Catalan /e/ in most contexts and the values of Catalan /ɛ/ in others (this acoustic asymmetry mapping accounts for the unevenness in accuracy rates for both experimental nonword categories). Because of the bilingual nature of Catalan society, many Spanish-dominant bilinguals currently use Catalan but pronounce it with a Spanish accent, using Spanish categories to utter Catalan words. Thus, most individuals are exposed not only to correct pronunciations (uttered by Catalan-dominant bilinguals), but also to incorrect ones (uttered by many Spanish-dominant bilinguals). In this way, two acoustic realizations for some Catalan words might coexist in the lexicon of Catalan-dominant bilinguals. In this scenario, the reduced N400 negativity associated with experimental nonwords may be related to the existence of an alternative phonological representation for some words. It seems likely that the pattern observed in this specific ERP component reflects the similar lexical activation generated by both stimuli types. Note that this account assumes that lexical entries may include different phonological representations. This kind of hypothesis has also been postulated for phonological variants that coexist within a single language, for instance, in the case of the different pronunciations of the English word “pretty” that can be pronounced either with a flap or a [t] variant (Connine, 2004). The continuous exposure to the mispronunciations of Spanish-dominant bilinguals may have created phonological variants for some Catalan words.

Finally, a second factor that may have contributed to differences between experimental and control nonwords has to do with the role of orthography in auditory word recognition. The higher activation of lexical entries by experimental nonwords is also supported by their potentially larger (incorrect) activation of a lexical entry through the orthographic codes, when compared with control stimuli. Many reports in the literature show that auditory presentation of words automatically activates orthographic representations (for instance, Ziegler, Ferrand, & Montant, 2004). Both Catalan phonemes /e/ and /ɛ/ are represented by the same single letter “e”

(diacritics are sometimes used, basically to differentiate between minimal pairs). As a result, the only way of knowing the corresponding sound of most written word representations containing the letter “e” is by inspecting the mental lexicon. When hearing experimental nonwords, participants in our experiment probably activated the orthographic representation of the corresponding original word. The activation of the orthographic representation of control nonwords would be smaller because replaced vowels in this stimulus type did not share the same letter.

Summarizing, the lack of differences between experimental words and nonwords in the N400 component for both populations cannot be considered as indicating that both populations process experimental stimuli in an equivalent way. Although both showed equivalent N400 patterns, the substantial differences in discrimination rates clearly rule out this possibility: Whereas Catalan-dominant bilinguals differentiate between correct phonological representations and incorrect ones; Spanish-dominant bilinguals do not. This difference between the two populations clearly emerged in the analyses of the ERN component.

Before turning to the ERN data, a few words about the LPC are in order. An increased late positivity (LPC) was observed in the e type nonword judgments, equivalent in both populations. A first interpretation of this result might be to relate it to an extra effort associated with these stimuli, as is the case in the P300 component (Johnson, 1986). In this study, this extra process could be related to a second judgment on phonology congruence that participants had to make after the incorrect activation of the target word. However, at present, we cannot provide any clear explanation of why this particular stimulus type should be more difficult to process than the other.

### **Error Detection: The ERN Component**

The fundamental question of the present research was to assess whether Spanish-dominant bilinguals were unable to distinguish between experimental words and nonwords, or whether they were merely hesitant in their response. The analyses of the ERN (and CRN) helped us to shed light on this issue. In addition, the analysis of this component lent support to the explanation of the lack of differences between words and nonwords in the N400 component proposed above for Catalan-dominant bilinguals.

Figure 4 presents the major findings of this study. In these analyses “word” responses to real words and nonwords (in the experimental conditions) were explored in response-locked ERPs. As Figure 5A shows, Catalan-dominant bilinguals showed ERN differences between erroneous experimental nonword decisions and correct experimental word decisions. Crucially, Spanish-dominant bilinguals showed no differences between

the two types of responses; that is, they did not process their erroneous responses to experimental nonwords as such. In fact, as discussed further below, their correct responses to real words and their incorrect responses to nonwords showed the same degree of uncertainty. Furthermore, as this figure also indicates, Catalan-dominant bilinguals showed an increased ERN component for erroneous experimental nonword decisions when compared with Spanish-dominant ones.

The analysis of the correct responses to experimental words and nonwords and control nonwords indicated a significant CRN. Figure 5 demonstrates that Spanish-dominant bilinguals showed a higher CRN component for correct responses to experimental words than Catalan-dominant participants. In fact, both bilingual populations showed an increased CRN for correct responses to experimental nonwords. A clear association between the level of uncertainty in decision making and the amplitude of the CRN has also been recently reported (Pailing & Segalowitz, 2004). An integrative account of the ERN and CRN in the present context could be as follows. In general, we should find that correct word responses lead to smaller CRN components than correct nonword responses. It is likely that in some cases participants may hesitate when accepting that some nonwords are actually nonwords (it could be that they are words that they do not know). This is, indeed, what is observed for Catalan-dominant participants (see Figure 5); peak values at central locations are near zero for correct experimental word responses but not for correct nonword responses. However, for nonwords, there are differences between correct responses for experimental and control stimuli. When interpreting the stimulus-locked analyses, we postulated that experimental nonwords were more likely to activate real words than control nonwords. Accordingly, it could be the case that participants were more uncertain about the correctness of their responses for experimental nonwords. This uncertainty was reflected in the response-locked analyses with higher amplitudes for these stimuli.

The pattern of results of Spanish-dominant bilinguals can also be accounted for in analogous terms. In this group, reductions (or no differences) in amplitudes were observed between correct experimental words and correct control nonwords. As already mentioned, this reduction in the difference is due to the increase in uncertainty for correct experimental words, measured with an increased CRN for correct experimental words (compared with Catalan-dominant bilinguals, both populations showed equivalent responses for correct control nonwords, but differed for correct experimental words). As mentioned, Spanish-dominant participants showed the same ERN responses for correct experimental words and erroneous responses to nonwords (Figure 4A), indicating that they were equally uncertain of their “yes” responses to experimental words (correct responses) and of their “yes” responses to experimental nonwords

(incorrect responses). This explanation is also consistent with the increased CRN observed for correct experimental nonword responses. Although it was a correct response, participants might have believed they were making a mistake, and, accordingly, a larger amplitude was observed for correct nonword experimental responses. It has to be remembered that these participants treated experimental nonwords as real words, as revealed by the low discrimination scores in these conditions (0.710 and 0.695 for e type and ε type stimuli, respectively). Although more studies need to be conducted, the CRN could be used as an index of uncertainty or difficulty of the lexical decisions under conditions with small differences between words and nonwords.

### Implications for First- and Second-language Phonological Representation

What are the limits of second-language phonological acquisition? The present results are consistent with previous studies indicating that, at least in some circumstances, persistent difficulties in second-language phonetic perception prevail (Takagi, 2002) even in the case of early, intensive exposure (Cutler, Mehler, Norris, & Seguí, 1989). In fact, the ERN results show that the erroneous behavioral responses of Spanish-dominant bilinguals to experimental nonwords actually reflect their inability to differentiate real words from mispronounced ones and do not reflect the existence of uncertainty and a bias toward giving positive responses. However, it has to be borne in mind that the current results reflect an extreme difficulty in L2 phoneme contrast perception; there is evidence that the relative distribution of L1 and L2 phonemic repertoires has an influence on the way a second language is learned and represented in the brain (Minagawa-Kawai, Mori, Sato, & Koizumi, 2004; Best, 1995).

However, our results also indicate that substantial plasticity is observed for *first language* phonological representations. Indeed, it is commonly reported that natives are able to easily adapt their phonological system to difficult situations (such as distorted speech, background noise) and dialects; this ability seems to be less present in the second language. The present results show clear electrophysiological evidence of long-term first-language adaptation of phonological lexical representations. Catalan-dominant bilinguals showed equivalent lexical activation for real words and for mispronunciations that could coincide with a particular phonological activation. Although the influence of the orthographic codes cannot be ruled out at present, it is reasonable to assume that the lack of differences in the N400 component partially reflects lexical activations for both experimental words and nonwords. The combined use of the N400 and the ERN components has identified the postlexical nature of the behavioral responses of the Catalan-dominant participants.

## Conclusion

The present experiment aimed to explore the extent to which highly proficient bilinguals are sensitive to subtle mispronunciations of L1 and L2 words. The results are consistent with previous data showing that when bilinguals fail to perceive an L2 contrast, there seems to be a single phonological representation for minimal pairs (Pallier, Colomé, & Sebastián-Gallés, 2001) or for difficult to perceive mispronunciations (Sebastián-Gallés et al., 2005). The present results also indicate that the L1 lexicon may incorporate phonological variants. The lack of differences in the N400 component between experimental words and nonwords for Catalan-dominant natives, together with the large CNR component for correct experimental nonword responses in this population, can be taken as support for this explanation. There is a substantial debate on the degree of plasticity of the different brain structures subserving the language function (see, for instance, McLaughlin et al., 2004). Our results agree with the generalized notion that some parts of the language system, in our case the acquisition of an L2 contrast, may show a striking lack of plasticity. Our results show that although L2 may, in some cases, show a striking lack of plasticity, the first language seems to be more dynamic in its capacity of adapting and incorporating new information.

## Acknowledgments

We thank L. Bosch, A. Costa, and G. Deco for their helpful suggestions, and X. Mayoral and E. Azañon for their technical support and help with testing participants. This study was supported by the James S. McDonnell Foundation (JMSF 20002079) and the Spanish Ministerio de Educación y Ciencia (with EC Fondos FEDER SEJ2004-07680-C02-01/PSIC). A. R. F. was supported by the program Ramon y Cajal and B. D. by a predoctoral fellowship (AP 2003-5091) both from the Spanish Ministerio de Educación y Ciencia.

Reprint requests should be sent to Núria Sebastián-Gallés, GRNC, Parc Científic, Edifici Docent, Carrer Santa Rosa 39-57, 08950 Esplugues del Llobregat, Spain, or via e-mail: nsebastian@ub.edu.

## REFERENCES

- Barch, D. M., Braver, T. S., Sabb, F. W., & Noll, D. C. (2000). Anterior cingulate and the monitoring of response conflict: Evidence from an fMRI study of overt verb generation. *Journal of Cognitive Neuroscience, 12*, 298–309.
- Bentin, S. (1987). Event-related potentials, semantic processes, and expectancy factors in word recognition. *Brain and Language, 31*, 308–327.
- Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology, 60*, 343–355.
- Best, C. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 171–206). Baltimore, MD: York Press.
- Bosch, L., Costa, A., & Sebastián-Gallés, N. (2000). First and second language vowel perception in early bilinguals. *European Journal of Cognitive Psychology, 12*, 189–222.
- Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280*, 747–749.
- Carter, C. S., MacDonald, A. W., III, Ross, L. L., & Stenger, V. A. (2001). Anterior cingulate cortex activity and impaired self-monitoring of performance in patients with schizophrenia: An event-related fMRI study. *American Journal of Psychiatry, 158*, 1423–1428.
- Christ, S., Falkenstein, M., Heuer, H., & Hohnsbein, J. (2000). Different error types and error processing in spatial stimulus-response-compatibility tasks: Behavioural and electrophysiological data. *Biological Psychology, 51*, 129–150.
- Chwill, D. J., Brown, C. M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology, 32*, 274–285.
- Clopper, C. G., & Pisoni, D. B. (2005). Perception of dialect variation: Some implications for current research and theory in speech perception. In D. B. Pisoni & R. E. Remez (Eds.), *Handbook of Speech Perception*. Oxford, UK: Blackwell.
- Coles, M. G., Scheffers, M. K., & Holroyd, C. B. (2001). Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. *Biological Psychology, 56*, 173–189.
- Connine, C. M. (2004). It's not what you hear but how often you hear it: On the neglected role of phonological variant frequency in auditory word recognition. *Psychonomic Bulletin & Review, 11*, 1084–1089.
- Cutler, A., Mehler, J., Norris, D., & Seguí, J. (1989). Limits on bilingualism. *Nature, 340*, 229–230.
- Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science, 5*, 303–305.
- Falkenstein, M., Hohnsbein, J., & Hoormann, J. (1995). Event-related potential correlates of errors in reaction tasks. *Electroencephalography and Clinical Neurophysiology, Supplement, 44*, 287–296.
- Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology, 51*, 87–107.
- Garavan, H., Ross, T. J., Murphy, K., Roche, R. A., & Stein, E. A. (2002). Dissociable executive functions in the dynamic control of behavior: Inhibition, error detection, and correction. *Neuroimage, 17*, 1820–1829.
- Gehring, W. J., Coles, M. G., Meyer, D. E., & Donchin, E. (1995). A brain potential manifestation of error-related processing. *Electroencephalography and Clinical Neurophysiology, Supplement, 44*, 261–272.
- Gehring, W. J., & Fencsik, D. E. (2001). Functions of the medial frontal cortex in the processing of conflict and errors. *Journal of Neuroscience, 21*, 9430–9437.
- Gerrig, R. J. (1993). *Experiencing Narrative Worlds*. New Haven, CT: Yale University Press.
- Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic processing. *Brain and Language, 35*, 66–85.
- Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology, 30*, 47–61.

- Holcomb, P. J., Graiger, J., & O'Rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience*, *14*, 938–950.
- Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes*, *5*, 281–312.
- Jennings, J. R., & Wood, C. C. (1976). Letter: The epsilon-adjustment procedure for repeated-measures analyses of variance. *Psychophysiology*, *13*, 277–278.
- Johnson, R. J. (1986). A triarchic model of P300 amplitude. *Psychophysiology*, *23*, 367–384.
- Ju, M., & Luce, P. A. (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science*, *15*, 314–318.
- Kiehl, K. A., Liddle, P. F., & Hopfinger, J. B. (2000). Error processing and the rostral anterior cingulate: An event-related fMRI study. *Psychophysiology*, *37*, 216–223.
- Kutas, M. (1997). Views on how the electrical activity that the brain generates reflects the functions of different language structures. *Psychophysiology*, *34*, 383–398.
- Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*, 463–470.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.
- Laurens, K. R., Ngan, E. T., Bates, A. T., Kiehl, K. A., & Liddle, P. F. (2003). Rostral anterior cingulate cortex dysfunction during error processing in schizophrenia. *Brain*, *126*, 610–622.
- Luu, P., Collins, P., & Tucker, D. M. (2000). Mood, personality, and self-monitoring: Negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *Journal of Experimental Psychology and Genetics*, *129*, 43–60.
- Luu, P., & Tucker, D. M. (2001). Regulating action: Alternating activation of midline frontal and motor cortical networks. *Clinical Neurophysiology*, *112*, 1295–1306.
- Marian, V., & Spivey, M. (2003). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics*, *24*, 173–193.
- McCallum, W. C., Farmer, S. F., & Pocock, V. P. (1984). The effects of physical and semantic incongruities on auditory event-related potentials. *Electroencephalography, and Clinical Neurophysiology*, *59*, 477–488.
- McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology*, *62*, 203–208.
- McClelland, J. L., Fiez, J. A., & McCandliss, B. D. (2002). Teaching the /r-/l/ discrimination to Japanese adults: Behavioral and neural aspects. *Physiology & Behavior*, *77*, 657–662.
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, *7*, 703–704.
- McNichol, D. (1972). *A primer of signal detection theory*. London: Allen & Unwin.
- Menon, V., Adelman, N. E., White, C. D., Glover, G. H., & Reiss, A. L. (2001). Error-related brain activation during a Go/NoGo response inhibition task. *Human Brain Mapping*, *12*, 131–143.
- Minagawa-Kawai, Y., Mori, K., Sato, Y., & Koizumi, T. (2004). Differential cortical responses in second language learners to different vowel contrasts. *NeuroReport*, *15*, 899–903.
- Münte, T. F., Urbach, T. P., Düzel, E., & Kutas, M. (2000). Event-related brain potentials in the study of human cognition and neuropsychology. In F. Boller, J. Grafmann, & G. Rizzolatti (Eds.), *Handbook of neuropsychology* (Vol. 1, pp. 139–235). Amsterdam: Elsevier.
- Nääätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Livonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*, 432–434.
- Nessler, D., & Mecklinger, A. (2003). ERP correlates of true and false recognition after different retention delays: Stimulus- and response-related processes. *Psychophysiology*, *40*, 146–159.
- Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology*, *38*, 752–760.
- Osman, A., Moore, C. M., & Ulrich, R. (1995). Bisecting RT with lateralized readiness potentials: Precue effects after LRP onset. *Acta Psychologica*, *90*, 111–127.
- Osterhout, L., McLaughlin, J., & Bersick, M. (1997). Event-related brain potentials and human language. *Trends in Cognitive Sciences*, *1*, 203–209.
- Pailing, P. E., & Segalowitz, S. J. (2004). The effects of uncertainty in error monitoring on associated ERPs. *Brain and Cognition*, *56*, 215–233.
- Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in vowel acquisition. *Cognition*, *64*, B9–B17.
- Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based vs. abstract lexical entries. *Psychological Science*, *12*, 445–449.
- Pallier, C., Dupoux, E., & Jeannin, X. (1997). EXPE: An expandable programming language for on-line psychological experiments. *Behavior Research Methods, Instruments, & Computers*, *29*, 322–327.
- Phillips, C., Marantz, A., McGinnis, M., Pesetsky, D., Wexler, K., Yellin, A., Poeppel, D., Roberts, T., & Rowley, H. (1995). Brain mechanisms of speech perception: A preliminary report. *MIT Working Papers in Linguistics*, *26*, 125–163.
- Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Jr., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*, 127–152.
- Rafel i Fontanals, J. (1998). *Diccionari de freqüencies*. Barcelona: Institut d'Estudis Catalans.
- Rodríguez-Fornells, A., Kofidis, C., & Münte, T. F. (2004). An electrophysiological study of errorless learning. *Cognitive Brain Research*, *19*, 160–173.
- Rodríguez-Fornells, A., Kurzbuch, A. R., & Münte, T. F. (2002). Time course of error detection and correction in humans: Neurophysiological evidence. *Journal of Neuroscience*, *22*, 9990–9996.
- Rugg, M. D., & Nagy, M. E. (1987). Lexical contribution to nonword-repetition effects: Evidence from event-related potentials. *Memory & Cognition*, *15*, 473–481.
- Scheffers, M. K., & Coles, M. G. (2000). Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 141–151.

- Scheffers, M. K., Coles, M. G., Bernstein, P., Gehring, W. J., & Donchin, E. (1996). Event-related brain potentials and error-related processing: An analysis of incorrect responses to go and no-go stimuli. *Psychophysiology*, *33*, 42–53.
- Schulpen, B., Dijkstra, A., Schriefers, H., & Hasper, M. (2003). Recognition of interlingual homophones in bilingual auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 1155–1178.
- Sebastián-Gallés, N. (2005). Cross-language speech perception. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 546–566). Oxford, UK: Blackwell.
- Sebastián-Gallés, N., Echeverría, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language*, *52*, 240–255.
- Sebastián-Gallés, N., & Kroll, J. (2003). Phonology in bilingual language processing: Acquisition, perception, and production. In A. Meyer & N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 279–318). Berlin: Mouton de Gruyter.
- Sebastián-Gallés, N., & Soto-Faraco, S. (1999). On-line processing of native and non-native phonemic contrasts in early bilinguals. *Cognition*, *72*, 112–123.
- Sharma, A., & Dorman, M. F. (2000). Neurophysiologic correlates of cross-language phonetic perception. *Journal of the Acoustical Society of America*, *107*, 2697–2703.
- Smith, M. E., & Halgren, E. (1987). Event-related brain potentials during lexical decision: Effects of repetition, word-frequency, pronounceability, and concreteness. *Electroencephalography and Clinical Neurophysiology, Supplement*, *40*, 417–421.
- Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10*, 281–284.
- Strange, W. (Ed.) (1995). *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press.
- Swick, D., & Turken, A. U. (2002). Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *99*, 16354–16359.
- Takagi, N. (2002). The limits of training Japanese listeners to identify English /r/ and /l/: Eight case studies. *Journal of the Acoustical Society of America*, *111*, 2887–2896.
- Toro, J. M., Sinnott, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, *97*, B25–B34.
- Tremblay, K., Kraus, N., Carrell, T., & McGee, T. (1997). Central auditory system plasticity: Generalization to novel stimuli following listening training. *Journal of the Acoustical Society of America*, *6*, 3762–3773.
- Tremblay, K., Kraus, N., & McGee, T. (1998). The time course of auditory perceptual learning: Neurophysiological changes during speech-sound training. *NeuroReport*, *9*, 3557–3560.
- Ullsperger, M., & von Cramon, D. Y. (2001). Subprocesses of performance monitoring: A dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *Neuroimage*, *14*, 1387–1401.
- van Veen, V., & Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology & Behavior*, *77*, 477–482.
- van Veen, V., Cohen, J. D., Botvinick, M. M., Stenger, V. A., & Carter, C. S. (2001). Anterior cingulate cortex, conflict monitoring, and levels of processing. *Neuroimage*, *14*, 1302–1308.
- Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the ‘error negativity’ specific to errors? *Biological Psychology*, *51*, 109–128.
- Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*, 1–25.
- Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931–959.
- Ziegler, J. C., Besson, M., Jacobs, A. M., Nazir, T. A., & Carr, T. H. (1997). Word, pseudoword, and nonword processing: A multitask comparison using event-related brain potentials. *Journal of Cognitive Neuroscience*, *9*, 758–775.
- Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory and Cognition*, *32*, 732–741.