

# Lexical Competition in Nonnative Speech Comprehension

Ian FitzPatrick<sup>1,2</sup> and Peter Indefrey<sup>1,2,3</sup>

## Abstract

■ Electrophysiological studies consistently find N400 effects of semantic incongruity in nonnative (L2) language comprehension. These N400 effects are often delayed compared with native (L1) comprehension, suggesting that semantic integration in one's second language occurs later than in one's first language. In this study, we investigated whether such a delay could be attributed to (1) *intralingual* lexical competition and/or (2) *interlingual* lexical competition. We recorded EEG from Dutch–English bilinguals who listened to English (L2) sentences in which the sentence-final word was (a) semantically fitting and (b) semantically incongruent or semantically incongruent but initially congruent due to sharing initial phonemes with

(c) the most probable sentence completion within the L2 or (d) the L1 translation equivalent of the most probable sentence completion. We found an N400 effect in each of the semantically incongruent conditions. This N400 effect was significantly delayed to L2 words but not to L1 translation equivalents that were initially congruent with the sentence context. Taken together, these findings firstly demonstrate that semantic integration in nonnative listening can start based on word initial phonemes (i.e., before a single lexical candidate could have been selected based on the input) and secondly suggest that spuriously elicited L1 lexical candidates are not available for semantic integration in L2 speech comprehension. ■

## INTRODUCTION

As Grosjean (1989) rightly pointed out, “The bilingual is not two monolinguals in one person,” alluding to the fact that there may be qualitative differences between how one produces and comprehends language in a second language (L2) and how a monolingual native (L1) speaker of that language would do so.

Important qualitative differences are evident in the domain of bilingual speech comprehension. For instance, when tasked with identifying spoken words in their second language, bilinguals are slower (e.g., Scarborough, Gerard, & Cortese, 1984; Soares & Grosjean, 1984), less proficient, and less confident in their identification than monolinguals (Schulpen, Dijkstra, Schriefers, & Hasper, 2003). Notably, bilinguals consistently take longer to process semantic anomalies in sentence contexts (e.g., “He spread the warm bread with socks.”) than monolinguals. That is, although bilinguals exhibit the same N400 effect (Kutas & Hillyard, 1980) to semantic incongruity in sentences as monolinguals, the effect is often delayed (Hahne, 2001; Weber-Fox & Neville, 1996; for a review, see Moreno, Rodriguez-Fornells, & Laine, 2008). The functional interpretation of this delay is far from clear, and a number of possible accounts have recently been put forward (see also Rueschemeyer, Nojack, & Limbach, 2008). These

accounts center on the notion that bilinguals, despite knowing fewer words in their second language, have to identify words from among a larger pool of concurrently activated candidates than monolinguals. In other words, in bilinguals, more lexical candidates compete for recognition than in monolinguals.

Two sources for this enhanced competition have been postulated: (1) due to less efficient phonological processing and/or confusable phonemes between languages, bilinguals may experience greater competition from *intralingual* lexical candidates (e.g., Broersma, 2005; Weber & Cutler, 2004), and (2) shared lexical storage systems between a bilingual's languages may cause concurrent activation of word candidates from both of the bilingual's languages (e.g., Weber & Cutler, 2004; Marian & Spivey, 2003a, 2003b; Marian, Spivey, & Hirsch, 2003; Schulpen et al., 2003; Spivey & Marian, 1999); hence, bilinguals may experience greater competition from *interlingual* lexical candidates. To earmark either of these possible sources of enhanced competition as plausible causes for N400 delays in bilinguals, it is important to establish whether concurrently activated intralingual and/or interlingual lexical items are actively evaluated during word recognition. Consequently, the main aim of the present study is to investigate what influence intralingual and/or interlingual lexical competition has on the time course of semantic processing in nonnative speech comprehension. Expanding on previous studies that have shown intralingual and/or interlingual competition in single word (e.g., Schulpen et al., 2003) and invariant sentence contexts

<sup>1</sup>Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, <sup>2</sup>Radboud University Nijmegen, Nijmegen, The Netherlands, <sup>3</sup>Heinrich Heine University Düsseldorf, Germany

(e.g., Weber & Cutler, 2004; Spivey & Marian, 1999), this is the first study to investigate nonnative lexical competition using semantically rich sentences.

### **Monolingual Word Recognition**

The concept of multiple lexical activation plays an important role in how we currently conceive of the word recognition process. Models of monolingual word recognition agree that multiple lexical candidates that match the input to a certain extent are briefly active during word recognition (Norris, 1994; Goldinger, Luce, & Pisoni, 1989; McClelland & Elman, 1986; Marslen-Wilson & Tyler, 1980). For example, hearing the word *box* would briefly activate words such as *bottle*, *boss*, or *body*. The collection of candidates is sometimes referred to as a cohort (Marslen-Wilson & Tyler, 1980) or shortlist (Norris, 1994). Concurrent activation of multiple lexical candidates is not simply an epiphenomenon of the speech recognition process; rather, cohort members are thought to actively compete with each other for recognition. In fact, recognition of a particular lexical item becomes progressively harder the more lexical candidates are concurrently active (Soto-Faraco, Sebastián-Gallés, & Cutler, 2001; Norris, McQueen, & Cutler, 1995; Vroomen & De Gelder, 1995; McQueen, Norris, & Cutler, 1994). As the speech signal unfolds, fewer and fewer candidates in the cohort will match the input and the size of the cohort will shrink (lexical selection). Words in the cohort must activate their semantic features (lexical access), and the activated semantic features can then be checked against the sentence context in a process called semantic integration.

Early electrophysiological studies of written language comprehension have identified an ERP component that is sensitive to the process of semantic integration. This component has been designated the N400 to reflect the fact that it is a negative going component that peaks between 300 and 500 msec after critical word onset (cf. Kutas & Hillyard, 1980). The N400 is more negative to words that are semantically incongruent within a sentence context than to congruent words. Certain characteristics of the N400 have proven useful indicators of underlying cognitive processes. For instance, the amplitude of the N400 component is taken to reflect the ease of semantic integration (Brown & Hagoort, 1993; Kutas & Hillyard, 1984), and the peak and the onset latencies of the N400 component are sensitive to the point at which a semantic incongruity is detected (O'Rourke & Holcomb, 2002; Praamstra, Meyer, & Levelt, 1994).

The N400 arises under similar circumstances in auditory language comprehension (Diaz & Swaab, 2007; Van den Brink, Brown, & Hagoort, 2001, 2006; Van den Brink & Hagoort, 2004; Newman, Connolly, Service, & McIvor, 2003; Hagoort & Brown, 2000; Van Petten, Coulson, Rubin, Plante, & Parks, 1999; Connolly & Phillips, 1994; Connolly, Phillips, Stewart, & Brake,

1992; Holcomb & Neville, 1991; Connolly, Stewart, & Phillips, 1990; McCallum, Farmer, & Pocock, 1984). Studies of speech comprehension report the earliest point at which an incongruity effect manifests itself to be around 200 msec after critical word onset. This early effect is sometimes reported as being functionally distinct from the N400 (e.g., Van den Brink & Hagoort, 2004; Newman et al., 2003; Van den Brink et al., 2001; Connolly & Phillips, 1994; Connolly et al., 1990, 1992). Others regard the early negativity as an early manifestation of the N400 component (Diaz & Swaab, 2007; Van den Brink et al., 2006; Van Petten et al., 1999). The existence of an early negative effect demonstrates that monolinguals are capable of noticing incongruity of a spoken word in a sentence context after only 200 msec. As Marslen-Wilson and Tyler (1980) pointed out, after the first 200 msec of a word (roughly corresponding to its first two phonemes), tens of possible word candidates remain viable based on the input alone. This implies that detection of incongruity is initiated even before incoming words are fully recognized.

The fact that incongruity can be detected at such an early point in time begs the question to what extent concurrently activated lexical candidates are considered for semantic integration. That is, when we hear a sentence like "When we move house I have to put all my books in a box," would semantic integration be attempted for all activated candidates (i.e., *bottle*, *boss*, *body*, etc.) or would only the selected candidate be considered? To address this question, Van den Brink et al. (2006) examined the relationship between the point at which a stimulus word was isolated (i.e., recognized) and the onset of the N400 effect. They found that the latency of the N400 effect did not differ between words with a late isolation point and words with an early isolation point. The latency of semantic integration is thus independent of the moment at which words are recognized. This finding clearly speaks against the concept of a "magic moment" at which lexical selection ends and semantic integration can start. It implies that semantic integration is attempted for a number of concurrently active lexical candidates even before one candidate is uniquely identified (selected) based on the input.

### **Bilingual Word Recognition**

Bilinguals generally know fewer words in their second language than do monolingual speakers of that language (e.g., Vermeer, 1992; Verhoeven & Vermeer, 1985). Although this fact might be expected to restrict the number of concurrently activated word candidates and constitute an advantage for bilingual word recognition, accumulating evidence seems to indicate that bilinguals even have to contend with a larger amount of concurrent lexical activation than monolinguals.

Converging experimental findings suggest that less efficient prelexical processing could contribute to spurious

activation of additional intralingual competitors. One striking example of spurious competitor activation arises when nonnative listeners are confronted with lexical items that contain confusable phonemes. Dutch–English bilinguals, for instance, have difficulty perceiving the contrast between /æ/ as in *pan* and /ɛ/ as in *pen* (Schouten, 1975). Such phonemic confusability has been shown to cause bilinguals to erroneously consider nonwords that differ only on a (for them) perceptually ambiguous phonemic contrast (e.g., *lemp–lamp*) as words (e.g., Sebastián-Gallés, Echeverría, & Bosch, 2005; Broersma, 2002). Thus, whereas English monolinguals hearing the word *pan* might activate lexical competitors like *panda*, *panther*, or *pancake*, Dutch–English bilinguals may experience additional competition from words like *pen*, *pencil*, or *pentagon*. That this is indeed the case has been demonstrated by a number of recent experiments. For instance, in a cross-modal priming paradigm, Broersma (2005) presented Dutch–English bilinguals and English native speakers with identical, mismatching, or unrelated auditory prime/visual target pairs. Mismatching primes were partial words that initially differed from the target on the /æ/–/ɛ/ vowel contrast (e.g., *daffo* from *daffodil* or *defi* from *deficit*). Whereas recognition of visual targets was inhibited following mismatching primes in English native speakers, auditory presentation of either *daffo* (identity) or *defi* (mismatching) led to significant priming of *daffodil* in Dutch–English bilinguals. This consequence of phonemic confusability is not simply a reflection of increased tolerance for phonemic mismatches in nonnative speech comprehension as becomes evident when we consider findings from Weber and Cutler (2004, Experiments 1a and 1b). They used an eye-tracking paradigm, in which participants identify the visual referent of an auditory stimulus from among phonologically related and unrelated distractors. Dutch–English bilinguals fixated more often on intralingual competitors when they differed from the target by way of a confusable (e.g., *panda–pencil*) rather than an unconfusable (e.g., *bottle–beetle*) phonemic contrast. Cutler (2005) estimated that, due to the perceived /æ/–/ɛ/ ambiguity, considerable numbers of nonwords embedded in larger real words (e.g., *daf* in *daffodil*, *lem* in *lemon*) may be erroneously perceived as real words (e.g., *deaf*, *lamb*). Increased intralingual lexical activation may thus pose a substantial challenge and, unfortunately for the bilingual, the difficulties with nonnative listening may go even further. Recent findings suggest that they may also have to contend with spuriously activated words from their other language (i.e., interlingual lexical activation).

A great deal of evidence for cross-linguistic lexical activation originates from bilingual visual word recognition studies demonstrating between-language lexical neighborhood size effects. The speed at which you recognize a word from one language is directly related to the number of words from the other language with a similar orthography in your lexicon (e.g., Van Heuven, Dijkstra & Grainger, 1998). This language nonselective lexical activation has since become an important feature of many

influential models of bilingual word recognition such as the bilingual interactive activation (BIA; Dijkstra & Van Heuven, 1998; Grainger & Dijkstra, 1992) and the BIA+ models (Dijkstra & Van Heuven, 2002). However, although cross-linguistic lexical activation is a well-established phenomenon in visual word recognition, the case for auditory word recognition is slightly more complex.

At first glance, it may not seem obvious why cross-linguistic activation in auditory language comprehension would happen at all. Whereas an interlingual homograph (e.g., “brand,” which is the Dutch word for “fire”) in isolation provides little clue as to its language membership, a bilingual may rely on a multitude of subtle phonemic and subphonemic cues to distinguish between different tokens of an interlingual homophone. Indeed, bilinguals are able to accurately judge language membership of words based on their initial phonemes alone (Grosjean, 1988) and have been shown to be sensitive to fine-grained acoustic-phonetic between language differences (Ju & Luce, 2004). Nevertheless, a number of important findings make a strong case for the existence of cross-linguistic lexical activation. Using Dutch–English bilinguals, Schulpen et al. (2003) showed that both pronunciations of an auditorily presented interlingual homophone could prime its L2 orthographic form (e.g., *lief* “sweet”—leaf vs. *leaf*—leaf). The authors took this as evidence that hearing a homophone activates both its L1 and L2 forms simultaneously. Equally striking evidence for language nonselective access to the bilingual lexicon comes from eye-tracking paradigms that show that bilinguals fixate on both intralingual and interlingual competitors while listening to verbal instructions in their L2 (Weber & Cutler, 2004; Marian & Spivey, 2003a, 2003b; Marian et al., 2003; Spivey & Marian, 1999). Thus, it seems that cross-linguistic lexical activation is a phenomenon that is not restricted to the domain of visual word recognition but can also occur in spoken language comprehension.

## Objectives

Although it has been shown that bilinguals experience greater intralingual as well as interlingual lexical competition in auditory word recognition, we are still somewhat removed from establishing whether either (or indeed both) of these sources of lexical competition are plausible causes for delayed semantic processing in L2 speech comprehension. An important first step to investigate the possibility of such a causal relationship would be to establish whether, firstly, concurrently activated intralingual competitors are considered for semantic integration in nonnative speech comprehension and, secondly, whether the same holds for cross-linguistically activated interlingual competitors.

In the present study, we use a monolingual L2 experimental setting in which our L2 listeners are kept unaware that their L1 is also under investigation. This setting is intended to reflect a common situation in L2 listening,

namely, full immersion in an all L2 environment. Although mixed language contexts are also common in L2 listening, we explicitly chose a monolingual L2 context to avoid unintentionally inducing interlingual competition. This experimental setting constitutes a strong test for interlingual lexical competition in L2 comprehension. Thus, finding interlingual competition under these restrictive circumstances would also allow us to infer the availability of interlingual lexical candidates if the L1 would have been more salient.

We investigate the availability of intralingual and interlingual competitors for semantic integration by exploiting the sensitivity of the N400 to the time point at which a semantic incongruity arises. As has been shown previously (e.g., Van den Brink et al., 2001, 2006; Van Petten et al., 1999), it is possible to infer the time course of lexical selection and semantic integration by examining the latency of the N400 to congruent words, incongruent words, and words that are initially congruent but become incongruent after the first few phonemes (e.g., “It was a pleasant surprise to find that the car repair bill was only seventeen *dollars/scholars/dolphins*”). If only one lexical candidate is considered for semantic integration (in other words, if semantic integration occurs after lexical selection has occurred), the N400 should have the same time course for initially congruent words as for incongruent words. That is, the semantic features of the cohort would only be assessed at the moment that one (in this case incongruent) item remains in the cohort. A delay of the N400 to initially congruent words indicates that semantic integration has started before lexical selection has occurred. Thus, multiple candidates have been considered for semantic integration.

We presented Dutch (L1)-English (L2) bilinguals with spoken sentences in their L2. The participants were drawn from the same population as was used by Weber and Cutler (2004) and Schulpen et al. (2003). Using semantically constraining sentence contexts, we manipulated the semantic fit of sentence-final target words such that they were (a) fully congruent (e.g., “The goods from Ikea arrived in a large cardboard *box*”), (b) fully incongruent (e.g., “He unpacked the computer, but the printer is still in the *towel*”), (c) initially congruent within the L2 (e.g., “When we moved house, I had to put all my books in a *bottle*”), or (d) initially overlapping with a congruent L1 lexical item (e.g., “My Christmas present came in a bright-orange *doughnut*,” which shares phonemes with the Dutch *doos* “box”).

Firstly, in accordance with earlier studies (Hahne, 2001; Hahne & Friederici, 2001), we expect an N400 effect between the fully incongruent (FI) condition and the fully congruent (FC) condition. Secondly, if intralingual lexical candidates are not considered for semantic integration in L2 listening (i.e., L2 listeners wait for lexical selection to occur before initiating semantic integration), the peak and/or onset latency of the N400 effect should not differ between the fully incongruent and the initially

congruent condition. If L2 listeners can initiate semantic integration after the word initial phonemes, we expect a difference in the onset and/or peak latency of the N400 between the condition where the critical word is initially congruent with the sentence context (ICL2) and the FI condition. This would reflect the fact that during the initial phonemes, the congruent item is still in the cohort; thus, semantic integration at this stage would treat the cohort as congruent with the sentence. Lastly, if concurrently activated L1 lexical candidates are considered for semantic integration, this should also be reflected in the onset and/or peak latency of the N400 between the L1 overlap condition (ICL1) and the FI condition. If L1 candidates are not considered for semantic integration, the participant should treat words with initial overlap with L1 items as if they were any other semantically incongruent word, in which case there would be no difference in N400 peak and/or onset latency between the ICL1 and the FI conditions.

## METHODS

### Participants

Thirty right-handed, highly proficient, late onset (after age 10), Dutch-English bilinguals participated in the experiment, 24 of which were included in the final analysis (7 men; mean age = 23.7 years). The participants' English proficiency was assessed using 50 grammaticality judgment items of the Oxford Placement Test (Allan, 1992; mean score = 43.65, “advanced level,”  $SD = 2.68$ ; maximum score = 50) and a nonspeeded lexical decision test (60 items), created by Meara (1996) and later adapted by Lemhöfer, Dijkstra, and Michel (2004; mean score = 75% correct,  $SD = 10.37$ ). Participants were either paid a small fee or they received study credits. None of the participants had any neurological impairment. All participants gave their written informed consent.

### Materials

Participants listened to English sentences that belonged to one of four conditions. In the FC condition, sentences ended in a high cloze probability word; for example, “The goods from Ikea arrived in a large cardboard *box*.” In the FI condition, sentences ended in a semantically incongruent word; for example, “He unpacked the computer, but the printer is still in the *towel*.” In the ICL2 condition, the sentence-final word shared initial phonemes with the highest cloze probability word; for example, “When we moved house, I had to put all my books in a *bottle*” (initial overlap with *box*). In the ICL1 condition, the sentence-final word shared initial phonemes with the direct translation of the highest cloze probability word in the participant's L1; for example, “My Christmas present came in a bright-orange *doughnut*” (initial overlap with “doos” where *doos* is Dutch for *box*). We defined a number

**Table 1.** Phoneme Correspondences between Dutch and English

<i>Dutch</i>	<i>English</i>
a:	ɑ:
ɑ	ʌ
ɛ, ø	ɛ, æ
ɪ	ɪ
i	i:
ɔ	ɒ
o:	ɔ:, əʊ
u, ʏ	u:
ə	ə
ɔu	əʊ

Displayed using the International Phonetic Alphabet (IPA) (International Phonetic Association, 1999), used to define phonemic overlap.

of correspondences between Dutch and English vowels and diphthongs (see Table 1), which we considered to be sufficiently similar to constitute an overlap. In each case, the extent of the overlap was the initial consonant or consonant cluster and the vowel. This amount of phonological overlap has been shown to be sufficient to elicit lexical competition in monolingual speech comprehension (e.g., Van den Brink et al., 2001; Van Petten et al., 1999). The stimulus sentences were selected from among 414 sentences that had been cloze tested by an independent group of participants ( $n = 15$ ). Sentences with high-cloze alternatives that shared initial phonemes with the (semantically congruent) target word were discarded. The average cloze probability for the remaining sentences was 0.47.

Thirty-eight English target words (e.g., “box”; FC condition) that were semantically congruent with the sentence context were matched with 38 semantically incongruent words that shared initial phonemes (e.g., “bottle”; ICL2 condition) with congruent target words, 38 semantically incongruent words that shared initial phonemes with a translation equivalent (Dutch: *doos* “box”) of the congruent word (e.g., “doughnut,” ICL1 condition), and 38 semantically incongruent words that were phonologically unrelated to the congruent word (e.g., “towel”; FI condition). For each set of four target words (e.g., “box,” “bottle,” “doughnut,” and “towel”), four sentence frames were created that had the FC item (e.g., “box”) as the most plausible continuation. We created two stimulus lists in which the four target words of each set were randomly assigned to each of the four corresponding sentence frames. Each sentence frame occurred only once per stimulus list. Every participant thus heard four sentences that had “box” as the most plausible sentence-final word. One of the sentences actually ended with the word “box,” the other three with “bottle,” “doughnut,” and “towel.” Seventy-six semantically congruent filler sentences were created and added

to both lists to balance the number of sentences that were incongruent and congruent. One stimulus list thus consisted of 152 experimental sentences (38 sentences per condition) and 76 filler items for a total of 228 sentences. Half of the participants were presented with stimuli from the first list, and half were presented with stimuli from the second list.

To give us a clear marker of critical word onset for time locking the EEG, all critical words were chosen from English nouns that had either a plosive onset or a vowel onset with a glottal stop. The distribution of critical words with a voiced plosive, an unvoiced plosive, and a vowel onset was kept constant over conditions. Critical words were controlled across conditions with respect to the number of phonemes and word frequency (see Table 2). Word frequencies were taken from the CELEX English lemma database (Baayen, Piepenbrock, & van Rijn, 1993). None of the critical words were cognates or homophones between English and Dutch.

The experimental sentences, fillers, and practice items were spoken by a female English native speaker at a normal speaking rate and with normal intonation. The materials were digitally recorded in a sound attenuating booth and digitized at a rate of 44.1 kHz. Sound files were later equalized to eliminate any differences in sound level. A full list of experimental materials is available via [http://corpus1.mpi.nl/ds/imdi\\_browser?openpath=MPI691203%23](http://corpus1.mpi.nl/ds/imdi_browser?openpath=MPI691203%23).

## Procedure

Participants were exclusively addressed in English by an English native speaker, both preceding and during the experiment, to make certain they were in a monolingual L2 language mode (Grosjean, 1982). Participants were placed in a sound-attenuating booth and were instructed to listen attentively to the sentences, which were played over two loudspeakers at a distance of roughly 1.5 m, and to try to understand them. The sound level was kept constant over participants. To ensure that participants remained focused on the sentences, they were prompted to make an animacy decision regarding the previous sentence (i.e., “Was there anything living in the last sentence?”) at five randomly occurring time points during the experiment. On average participants gave 4.0 out of

**Table 2.** Mean Log Frequency Per Million and Mean Number of Phonemes

<i>Condition</i>	<i>Frequency (SD)</i>	<i>Phonemes (SD)</i>
FC	3.34 (0.95)	5.29 (2.30)
FI	3.061 (1.03)	5.17 (1.90)
ICL1	2.94 (1.19)	4.90 (1.55)
ICL2	2.89 (1.29)	5.60 (1.90)

Standard deviations are given in parentheses.

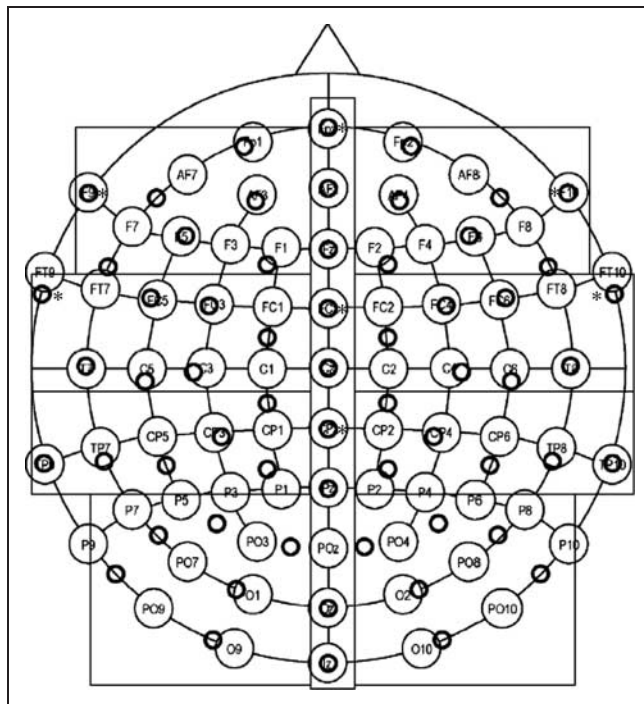
five correct responses, suggesting that they listened to the sentences attentively.

Each trial began with a 300-msec warning tone, followed by 1200 msec of silence, then a spoken sentence. The next trial began 4100 msec after the sentence offset. To ensure that participants did not blink during and shortly after presentation of the sentence, 1000 msec before the beginning of the sentence, a fixation point was displayed. Participants were instructed not to blink while the fixation point was on the screen. The fixation point remained until 1600 msec after the offset of the spoken sentence. Participants had a practice session with five sentences to familiarize themselves with the experimental setting.

After the EEG recording, the participants completed a word translation test on the critical items to verify that they were known and a cloze test on all the experimental sentence frames to check whether participants expected the sentence continuation that we had envisaged.

### EEG Recording

The EEG was recorded continuously from 64 sintered Ag/AgCl electrodes, each referred to an electrode on the nose of the participant. The electrodes were mounted in an equidistant elastic cap (www.easycap.de; for the electrode distribution, see Figure 1). The EEG and the EOG recordings were amplified with a BrainAmp DC amplifier (Brain Products, München, Germany) using a high cutoff



**Figure 1.** Electrode setup. Radial projection of electrode positions (small black circles) relative to the 10–20 system. Boxes denote regions for statistical analyses. Asterisks denote electrodes excluded from analysis.

of 200 Hz, a time constant of 10 sec (0.016 Hz), and a sampling rate of 500 Hz. Impedances were kept below 5 k $\Omega$ . Trials with eye blinks or deflections exceeding 70  $\mu$ V were rejected.

### ERP Analysis

Data from six participants were not analyzed. Four participants were excluded due to excessive alpha. Data from one participant were incomplete due to a technical malfunction. One other participant was left out due to failure to complete the posttests. The data were analyzed using the FieldTrip (<http://neuroimaging.ruhosting.nl/fieldtrip>) toolbox for Matlab (<http://www.mathworks.com>). EEG data were time locked to critical word onset. Average waveforms were calculated for each participant using a 150-msec pre-stimulus baseline. Grand average waveforms were calculated by averaging the individual average waveforms. Statistical analysis was performed by taking the mean amplitude per site (see Figure 1), in the N400 latency range (300–800 msec), from the grand averaged data. We used an omnibus ANOVA with condition (four levels) and site (nine levels) as within-subject factors. Seven electrodes were excluded from the analysis to have an equal number of electrodes in each site (see Figure 1). The latency range was chosen based on the previous literature and visual inspection of the grand average waveforms. All *p* values are reported after Greenhouse–Geisser correction (Greenhouse & Geisser, 1959). Contrasts between pairs of conditions were tested using a randomization approach that corrects for multiple comparisons (Maris, 2004; for a brief description, see Tuladhar et al., 2007; Takashima et al., 2006).

Cluster randomization was performed on the following pairs of conditions: FI versus FC, ICL1 versus FC, ICL2 versus FC, ICL1 versus FI, and ICL2 versus FI, using the same latency range as the ANOVA (300–800 msec).

To determine the peak and the onset latencies of the N400 in the three semantically incongruent conditions, we applied a low-pass filter at 5 Hz to the difference waveforms (FI-FC, ICL1-FC, and ICL2-FC) of the individual averages. We restricted our search to electrodes that show a significant N400 effect as determined by the cluster-randomization analysis. The peak of the N400 component was defined as the minimum of the filtered individual difference waveforms, in the 300- to 800-msec latency range. Visual quantification of onset latencies was complicated due to variability of individual averages. We therefore computed the mean amplitude values of the difference waveforms in 30-msec bins that shifted in steps of 10 msec in the latency range between critical word onset and 600 msec after critical word onset (cf., Hagoort & Brown, 2000). The values of these latency bins were tested against the null hypothesis that they did not differ from zero using *t* tests. We defined the onset latency of the N400 as the first bin at which three successive bins reached a significance threshold of *p* < .05.

## RESULTS

### Grand Averages

Figure 2 shows the grand average waveforms on 16 scalp electrodes and the topographical distribution of potentials in each condition. The waveforms for the three incongruent conditions (FI, ICL1, and ICL2) show an increased negativity in the 300- to 800-msec latency range relative to the fully congruent condition. This negativity is most pronounced on the centro-parietal electrodes. Figure 3 shows the difference waveforms of the incongruent conditions minus the fully congruent condition on 16 scalp electrodes.

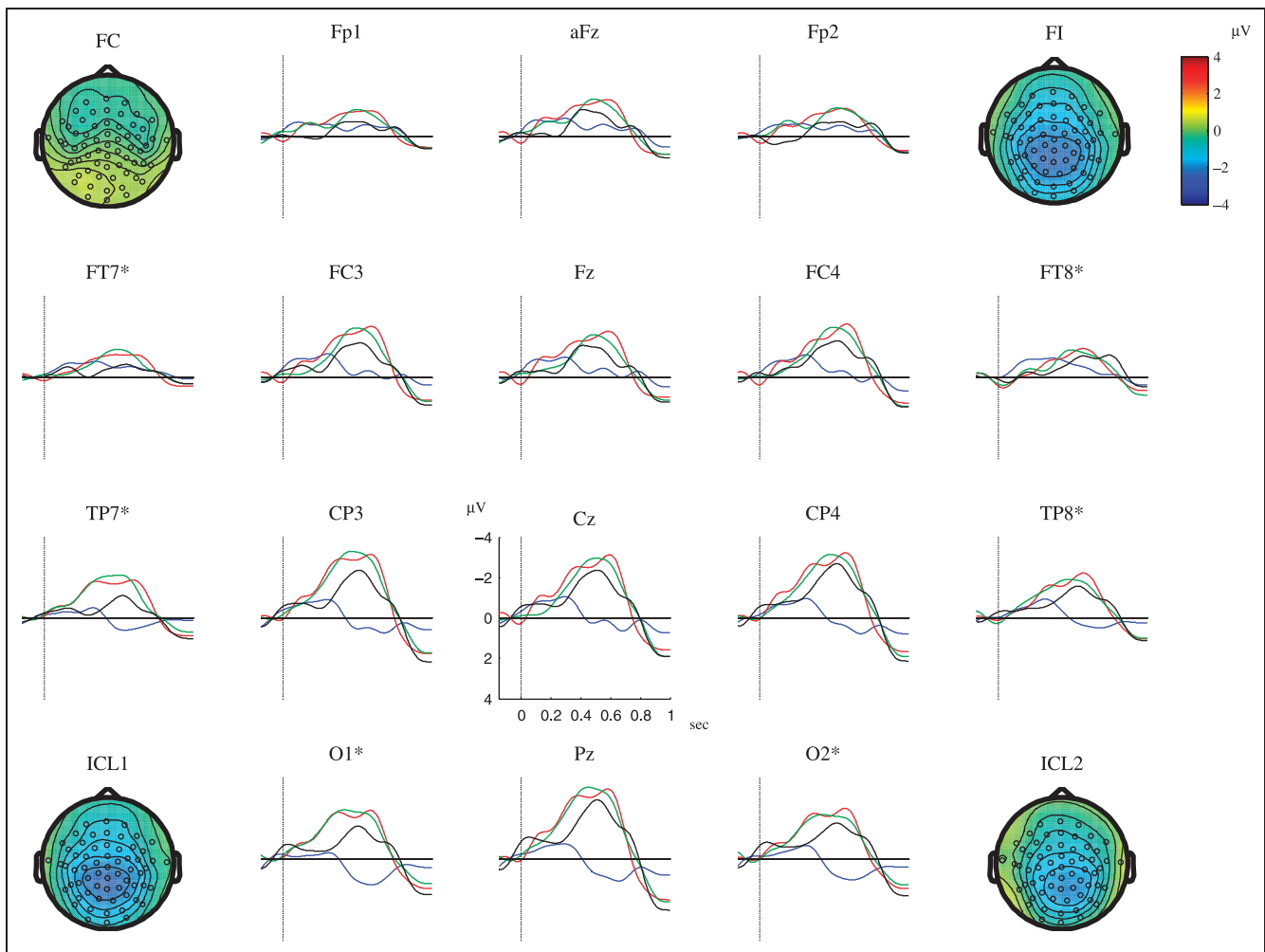
### Omnibus ANOVA

In the 300- to 800-msec latency range, the ANOVA yielded a significant main effect of condition,  $F(3, 69) = 6.128$ ,  $p_{GG} < .01$ ,  $\epsilon = .210$ . A priori contrasts revealed significant differences between the FC condition and the FI,  $F(1, 23) =$

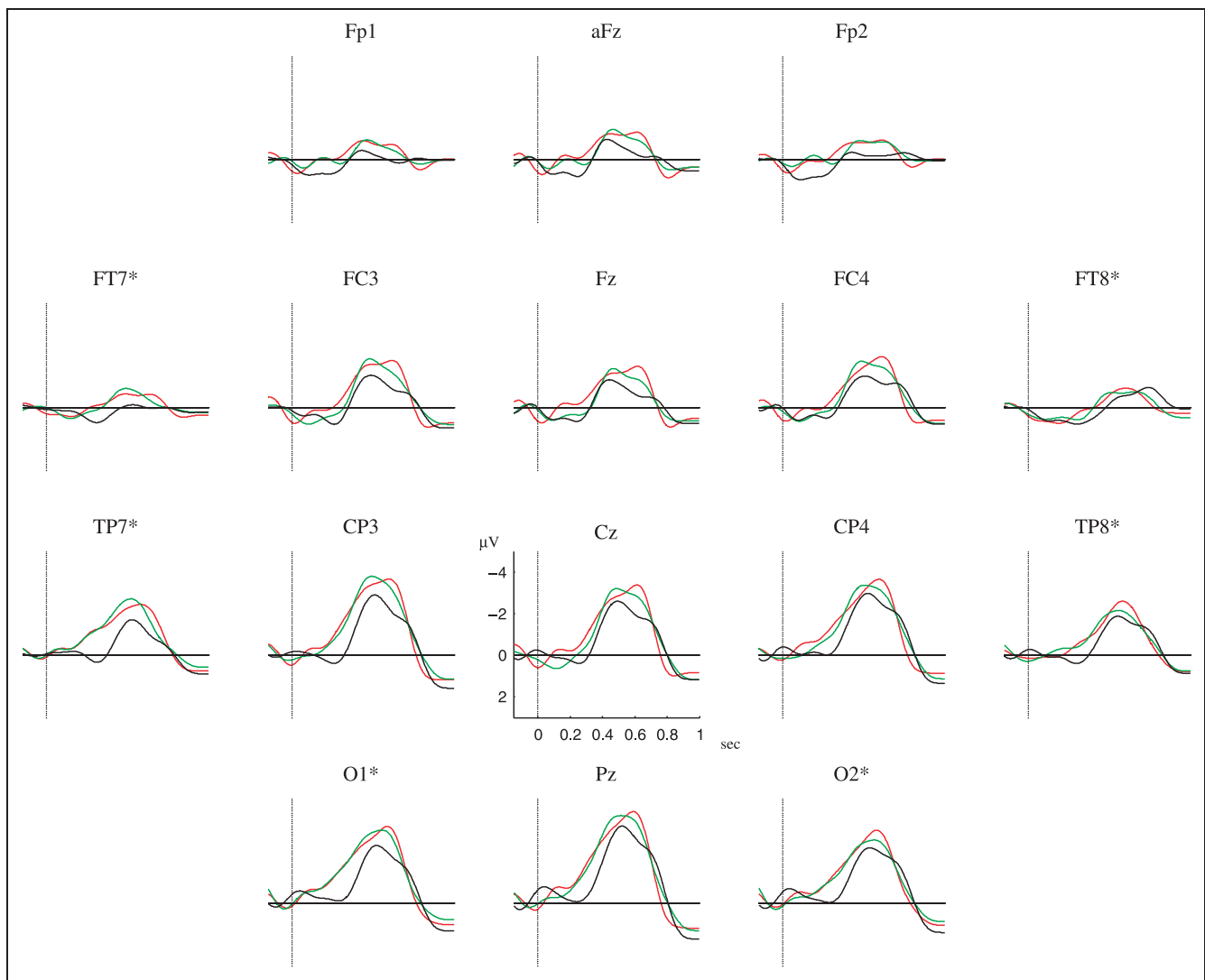
$10.507$ ,  $p < .01$ ,  $\epsilon = .314$ , ICL2,  $F(1, 23) = 6.448$ ,  $p < .05$ ,  $\epsilon = .219$ , and ICL1,  $F(1, 23) = 18.368$ ,  $p < .001$ ,  $\epsilon = .444$ .

There was also a significant main effect of site,  $F(8, 184) = 9.099$ ,  $p_{GG} < .001$ ,  $\epsilon = .283$ , with midline site,  $F(1, 23) = 46.762$ ,  $p < .001$ ,  $\epsilon = .670$ , right precentral site,  $F(1, 23) = 12.000$ ,  $p < .01$ ,  $\epsilon = .343$ , left postcentral site,  $F(1, 23) = 4.738$ ,  $p < .05$ ,  $\epsilon = .171$ , and right postcentral site,  $F(1, 23) = 13.403$ ,  $p < .01$ ,  $\epsilon = .368$ , showing the greatest negativity.

Finally, there was a significant interaction of condition with site,  $F(24, 552) = 5.596$ ,  $p_{GG} < .01$ ,  $\epsilon = .196$ , reflecting the fact that the greatest negativity in the FI versus FC, ICL2 versus FC, and ICL1 versus FC comparisons was found over the midline site: FI versus FC,  $F(1, 23) = 8.676$ ,  $p < .01$ ,  $\epsilon = .274$ ; ICL2 versus FC,  $F(1, 23) = 10.753$ ,  $p < .01$ ,  $\epsilon = .319$ ; ICL1 versus FC,  $F(1, 23) = 12.361$ ,  $p < .01$ ,  $\epsilon = .350$ ; the right postcentral site: FI versus FC,  $F(1, 23) = 9.972$ ,  $p < .01$ ,  $\epsilon = .302$ ; ICL2 versus FC,  $F(1, 23) = 11.446$ ,  $p < .01$ ,  $\epsilon = .332$ ; ICL1 versus FC,  $F(1, 23) = 7.457$ ,  $p < .05$ ,  $\epsilon = .245$ ; and the right occipital site, FI versus FC,



**Figure 2.** Grand average waveforms on 16 scalp electrodes and scalp topographies. (Blue lines, top left) Fully congruent; (red lines, top right) fully incongruent; (black lines, bottom right) initially congruent with the L2; and (green lines, bottom left) initially congruent with the L1. All waveforms were filtered with a 5-Hz low-pass filter for display purposes. Asterisk denotes nonstandard electrode location.



**Figure 3.** Difference waveforms on 16 scalp electrodes. (Red lines) Fully incongruent–fully congruent; (black lines) initially congruent with the L2-FC; and (green lines) initially congruent with the L1–fully congruent. All waveforms were filtered with a 5-Hz low-pass filter for display purposes. Asterisk denotes nonstandard electrode location.

$F(1, 23) = 8.047, p < .01, \epsilon = .259$ ; ICL2 versus FC,  $F(1, 23) = 5.034, p < .05, \epsilon = .180$ ; ICL1 versus FC,  $F(1, 23) = 4.425, p < .05, \epsilon = .161$ . In addition, the comparisons FI versus FC and ICL1 versus FC showed strong negativities over left postcentral site, FI versus FC,  $F(1, 23) = 17.670, p < .001, \epsilon = .434$ ; ICL1 versus FC,  $F(1, 23) = 42.691, p < .001, \epsilon = .650$ , and left occipital site, FI versus FC,  $F(1, 23) = 11.902, p < .01, \epsilon = .341$ ; ICL1 versus FC,  $F(1, 23) = 16.845, p < .001, \epsilon = .423$ .

### Fully Incongruent

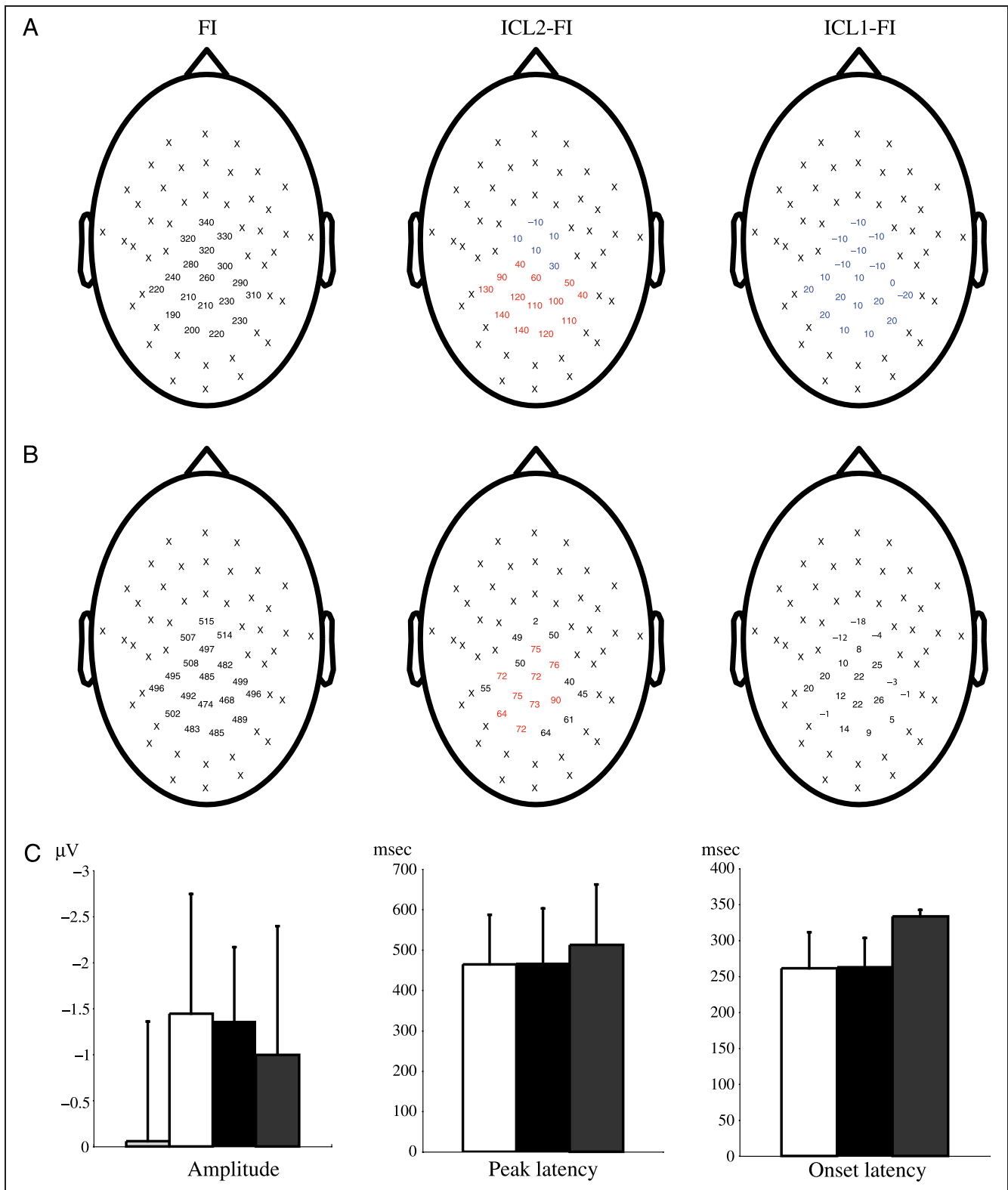
Relative to the FC condition, there was a significant negative cluster starting at 366 msec after critical word onset ( $p < .001$ , cluster size = 6516 data points) and lasting until 704 msec. Figure 4A shows the grand average onset latency of the negativity for each electrode that showed a significant negative effect as determined by the cluster-randomization analysis. Figure 4B shows the grand aver-

age peak latency of the negativity for each electrode that showed a significant negative effect as determined by the cluster-randomization analysis.

### Initially Congruent with the L2

No significant clusters were found in the comparison of the ICL2 condition with the FI condition. Relative to the FC condition, there was a negative cluster starting at 422 msec ( $p < .001$ , cluster size = 4136 data points) and lasting until 732 msec. The onset latency of the negativity, in the 300- to 800-msec time window, was substantially delayed compared with the corresponding negativity in the FI condition (see Figure 4A and C). To test whether the peak latency delay similarly delayed, we performed paired-samples  $t$  tests on the peak latencies of negativity in the ICL2 condition versus the FI condition (one tailed for ICL2 > FI) for each electrode that showed a significant negative effect as determined by





**Figure 4.** Latency and amplitude plots. Topographical representation of the (A) onset and (B) peak latencies (in milliseconds from critical word onset) on N400 electrodes for the fully incongruent (FI) condition (B and C, left panels) and latency differences between the initially congruent with the I2 condition (ICL2; A and B, middle panels) and the initially congruent with the I1 condition (ICL1; A and B, right panels) compared with the FI condition. Red color denotes latency difference greater than one bin size (30 msec; A) or significant latency difference (B;  $p$ -corrected < .05). Blue color denotes latency difference smaller than one bin size (A only). “X” denotes electrode that was excluded from the latency analysis. Bar charts (C) of the average N400 amplitude (left), N400 peak latency (middle), and N400 onset latency (right) in the FC (light gray bar; amplitude plot only), FI (white bars), ICL1 (black bars), and ICL2 (dark gray bars) conditions.

the cluster-randomization analysis. After Bonferroni correction, 9 of 18 electrodes showed a significant delay ( $p$ -corrected  $< .05$ ; Figure 4B).

### **Initially Congruent with the L1**

No significant clusters were found in the comparison of the ICL1 condition with the FI condition. Relative to the FC condition, there was a significant negative cluster starting at 368 msec ( $p < .001$ , cluster size = 6,900 data points) and lasting until 710 msec. Neither the onset (Figure 4A and C) nor the peak latency (Figure 4B and C) of the negativity in the 300- to 800-msec time window differed from the corresponding negativity in the FI condition.

## **DISCUSSION**

The present study investigated whether concurrently active intralingual and interlingual lexical candidates are considered for semantic integration in nonnative speech comprehension. Highly proficient, late onset Dutch–English bilinguals listened to sentences in English that ended in a word that was semantically congruent (FC condition), semantically incongruent (FI condition), semantically incongruent but initially overlapping with the most probable sentence completion (ICL2 condition), or semantically incongruent but initially overlapping with the L1 translation equivalent of the most probable sentence completion (ICL1 condition). We explicitly chose an all L2 experimental setting to avoid unintentionally inducing the effects of interest. Our findings provide evidence that, under these circumstances, intralingual but not interlingual lexical candidates are considered for semantic integration in nonnative speech comprehension. Possible effects of bilingual language proficiency and linguistic and nonlinguistic context will be discussed below.

### **Semantic Integration in Nonnative Listening**

As expected, we observed a significant negativity between 300 and 800 msec following critical word onset, in each of the semantically incongruent conditions compared with the fully congruent condition, consistent with an N400 effect (Kutas & Hillyard, 1980). The scalp topography of the N400 effect is comparable to earlier findings from monolingual studies of speech processing (e.g., Van den Brink et al., 2001); however, the latency of the N400 may be slightly longer in our study. As far as we know, the only studies to report peak latency measures of the N400 for speech processing in monolingual English speakers, with the sentence-final word as the critical word, are those of Connolly and Phillips (1994) and Connolly et al. (1992). Whereas the N400 in their phoneme mismatch–semantic mismatch condition peaked around 420 msec, we found the average peak latency of the N400 in the fully

incongruent condition to be approximately 490 msec. Although our study did not include a monolingual control condition, which would allow for a direct comparison of N400 latencies in native and nonnative listening, we note that this apparent delay is consistent with earlier findings of delayed N400s in nonnative written (Weber-Fox & Neville, 1996; Ardal, Donald, Meuter, Muldrew, Luce, 1990) and spoken language comprehension (Hahne, 2001).

### **Intralingual Competition**

We hypothesized that if intralingual lexical candidates are considered for semantic integration in L2 listening, the peak and/or onset latency of the N400 effect would be later for initially congruent words than for fully incongruent words. This would reflect the fact that during the initial phonemes, the congruent item is still in the cohort; thus, semantic integration at this stage would treat the cohort as congruent with the sentence. Indeed, initial phonemic overlap with the most probable sentence continuation delayed both the peak latency and the onset latency of the N400 by nearly 70 msec compared with the semantically fully incongruent condition (Figure 4). Similar results were obtained in native speech comprehension studies by Van den Brink et al. (2001) and Van Petten et al. (1999) using almost the same paradigm. Van den Brink et al. argue that this effect is driven by an N200 component that is present for fully incongruent words but absent for initially congruent words. In the present study, however, we found no evidence to suggest that this early negative effect is functionally and/or physiologically distinct from the N400 effect (see also Diaz & Swaab, 2007). A N400 peak latency delay to initially congruent words compared with fully incongruent words has also been reported earlier by Connolly and Phillips (1994) for native language listening. Visual inspection of their waveforms suggests that the onset of the N400 may also have been delayed for initially congruent words, although this is not reported by the authors.

Our results thus replicate peak and onset latency delays of the N400 to initially congruent words relative to fully incongruent words for nonnative listening. This finding suggests that nonnative listeners process speech in the same cascaded manner as do native listeners. They treat the initial phonemes of these words as congruent with the sentence context and only later detect the semantic incongruity. As the semantic assessment of word initial phonemes is contingent upon starting semantic integration, it follows that semantic integration must have started before lexical selection has occurred (for a similar view, see Van den Brink et al., 2006). That is, listeners start semantic integration whereas multiple lexical candidates are consistent with the input. This finding therefore not only shows multiple lexical activation but also cascaded lexical selection and semantic integration in L2 speech comprehension. We have thus established that intralingual competitors are available for

semantic integration in nonnative speech comprehension, leaving open the question whether the same holds for interlingual competitors.

### Interlingual Competition

We did not find a delay in either the peak or the onset latencies of the N400 in the initially congruent with the L1 condition relative to the fully incongruent condition (Figure 4). Thus, it seems that nonnative listeners do not treat initial overlap with the translation of the most likely sentence continuation as though it were initially congruent with the sentence context. This finding could mean one of two things: either (1) no L1 lexical candidates were elicited or (2) elicited L1 lexical candidates are not available for semantic integration. We will explore both of these accounts below.

A potential absence of L1 lexical activation could be explained in a relatively trivial manner; that is, that the degree of interlingual phonemic overlap was simply insufficient to elicit cross-linguistic lexical candidates. Although such an explanation cannot be completely discounted based on these data, we note that previous studies that found cross-linguistic activation did so despite using stimulus materials with nonidentical phonemic correspondences between languages. Indeed, the present study used stimulus materials with similar phonemic correspondences to Weber and Cutler (2004). Thus, it is improbable that mere phonemic mismatch could completely account for our findings. However, activation of cross-linguistic competitors may have been influenced by effects of both linguistic and nonlinguistic context.

To date, studies that demonstrated cross-linguistic lexical activation in auditory word recognition have presented critical items in isolation (Schulpen et al., 2003) or in imperatives such as “Pick up the stamp” (Weber & Cutler, 2004; Spivey & Marian, 1999), which did not vary over the course of the experiment. In our study, critical words were presented in the final position of semantically rich sentence contexts. The presence of such a sentence context may have influenced the degree of cross-linguistic activation.

Modulatory effects of sentence context on word recognition are not unique to bilingual language comprehension. In a monolingual study, Zwitserlood (1989) showed that concurrently activated lexical candidates (e.g., *kapitein* “captain,” *kapitaal* “capital”) remain in competition for longer when embedded in low-constraint sentences than when embedded in highly constraining sentences. Furthermore, numerous studies have shown that sentence context can modulate the relative availability of the dominant and subordinate meanings of intralingual homophones (e.g., Tabossi, 1988) or increase the salience of particular semantic features (e.g., Moss & Marslen-Wilson, 1993). The most thorough investigation of the role of sentence context on the degree of cross-linguistic activation in L2 reading to date comes from

Duyck, Van Assche, Drieghe, and Hartsuiker (2007). They exploited the well-known facilitatory effect of between language cognates and near cognates on bilingual word recognition. In isolation and for sentences presented word by word, the authors observed cross-linguistic activation for both cognates and near cognates; however, the near-cognate effect disappeared when the full sentence was presented, whereas the cognate effect remained. The authors thus concluded that the presence of a sentence context “may influence, but does not nullify” cross-linguistic lexical activation. Our study used critical items that had considerably less cross-linguistic overlap than the cognates in Duyck et al. (2007). It is therefore not implausible that cross-linguistic lexical candidates are only available given enough bottom-up support (i.e., enough phonological overlap).

That nonlinguistic context can also influence cross-linguistic activation is nicely demonstrated by Elston-Guettler, Gunter, and Kotz (2005). They found cross-linguistic homograph priming in L2 sentence comprehension but only in participants that had previously been exposed to an L1 narrated silent film and only in the first experimental block. This led them to posit that bilinguals restrict their lexical search by gradually “zooming in” to the language at hand. This may indicate that cross-linguistic activation may only occur in situations where the salience of the non-target language is in some way enhanced. In our study, we took care not to cue our participants to the fact that their L1 was under investigation by addressing them in their L2 for the duration of the experiment, thereby arguably decreasing the chances of finding cross-linguistic activation. It should be noted, however, that the studies by Weber and Cutler (2004) and Marian and Spivey (2003a, 2003b; Marian et al., 2003) used similar measures but still found cross-linguistic activation.

Effects of both the linguistic and the nonlinguistic context can be accounted for by the BIA+ (Dijkstra & Van Heuven, 2002) model of bilingual language comprehension. This model was primarily intended to be applied within the domain of visual word recognition; however, many of the tenets of BIA+ may still hold for speech comprehension. The model assumes that the bilingual language comprehension system is fundamentally nonselective in nature, allowing for cross-linguistic activation of lexical candidates. Such cross-linguistic activation is thought to be unencumbered by top-down influences of linguistic or nonlinguistic context. These factors come into play at a task schema level and thus may influence postlexical selection and/or semantic integration processes.

Our choice of participants may also have influenced the likelihood of finding interlingual competition. All our participants were highly proficient speakers of English, who learned English at high school, and continued using it at university level. The Revised Hierarchical Model of the bilingual lexicon (Kroll & Stewart, 1994) assumes that

lexical access is less reliant on the L1 in highly proficient than in less proficient bilinguals. Indeed, Elston-Guettler and Gunter (2008) and Elston-Guettler, Paulmann, and Kotz (2005) showed that highly proficient bilinguals are more able to “zoom in” to the target language compared with less proficient bilinguals. We cannot fully exclude that we would have found interlingual competition with a less proficient group of participants. However, we note that our participants were drawn from the same population as was used in the studies by Weber and Cutler (2004) and Schulpen et al. (2003), both of which show cross-linguistic lexical activation.

By choosing Dutch and English as the language pair under investigation, we may also have affected our chances of finding non-target language competition due to the fact that these two languages are closely related and share a high proportion of cognates and near cognates. Note, however, that this property of the two languages might rather lead one to expect increased competition because language membership might be relatively more difficult to assess for Dutch–English bilinguals than for speakers of two unrelated languages.

This consideration becomes relevant because an alternative interpretation of our findings could be that cross-linguistic lexical candidates are active but are simply not considered for semantic integration. Because our study focuses on the N400 effect, our data preclude semantic integration of non-target language items based on initial phonemic overlap alone; however, these data do not exclude competition by activated L1 lexical candidates. Various studies have shown bilinguals to be sensitive to fine-grained phonetic information in the speech signal that enables them to accurately judge the language membership of incoming words based on very little input (Ju & Luce, 2004; Pallier, Colomé, & Sebastián-Gallés, 2001; Li, 1996; Grosjean, 1988). The early availability of language membership information may be sufficient to exclude spuriously activated cross-linguistic lexical candidates from further semantic processing. To reconcile such an interpretation with previous findings, one would have to argue that cross-linguistically elicited lexical candidates can nonetheless be active to such a degree that they can cause orthographic priming effects (e.g., Schulpen et al., 2003) and influence the visual search for a referent in eye-tracking paradigms (e.g., Weber & Cutler, 2004; Spivey & Marian, 1999). Further studies will need to be conducted to disentangle the effects of sentence context, non-linguistic context, language proficiency, and the degree of interlingual phonological overlap on cross-linguistic lexical activation.

## Conclusions

Our findings may represent mixed blessings for the proficient nonnative listener. On the one hand, we show that nonnative listeners are capable of semantically integrating words in speech before a unique lexical candidate

is identified. This is encouraging as subjectively reported lowered confidence in nonnative word identification does not cause the L2 listener to adopt a more cautious approach to word recognition, such as delaying semantic integration until words can be positively identified by the input. Our findings further show that, in an all L2 context, non-target language candidates are not considered for semantic integration-based initial phonemic overlap alone. Although this is good news for immersed nonnative listeners, it is at least conceivable that it might lead to a delay in recognizing non-target language words that actually do appear, such as code switches. This implication constitutes an intriguing question for future studies.

## Acknowledgments

The authors thank Daniëlle van den Brink and Robert Oostenveld and three anonymous reviewers for their insightful comments and suggestions. They also express their gratitude to Michel Bex, Vasiliki Folia, Jana Hanulova, Lilla Magyari, Stephan Miedl, Anouk Peijnenborgh, Willemijn Schot, and Kirsten Weber for assisting with the EEG electrode application.

Reprint requests should be sent to Ian FitzPatrick, Max Planck Institute for Psycholinguistics, P.O. Box 310, NL-6500 AH Nijmegen, The Netherlands, or via e-mail: Ian.FitzPatrick@mpi.nl.

## REFERENCES

- Allan, D. (1992). *Oxford Placement Test*. Oxford: Oxford University Press.
- Ardal, S., Donald, M. W., Meuter, R., Muldrew, S., & Luce, M. (1990). Brain responses to semantic incongruity in bilinguals. *Brain and Language*, *39*, 187–205.
- Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database*. Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.
- Broersma, M. E. (2002). Comprehension of non-native speech: Inaccurate phoneme processing and activation of lexical competitors. *Proceedings of the 7th International Conference on Spoken Language Processing*. Available at [www.isca-speech.org/archive/icslp02](http://www.isca-speech.org/archive/icslp02).
- Broersma, M. E. (2005). *Phonetic and lexical processing in a second language*. Unpublished doctoral dissertation, Radboud University Nijmegen, Nijmegen.
- Brown, C. M., & Hagoort, P. (1993). The processing nature of the N400—Evidence from masked priming. *Journal of Cognitive Neuroscience*, *5*, 34–44.
- Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, *6*, 256–266.
- Connolly, J. F., Phillips, N. A., Stewart, S. H., & Brake, W. G. (1992). Event-related potential sensitivity to acoustic and semantic properties of terminal words in sentences. *Brain and Language*, *43*, 1–18.
- Connolly, J. F., Stewart, S. H., & Phillips, N. A. (1990). The effects of processing requirements on neurophysiological responses to spoken sentences. *Brain and Language*, *39*, 302–318.
- Cutler, A. (2005). The lexical statistics of word recognition problems caused by L2 phonetic confusion. *Proceedings of the 9th European Conference on Speech Communication and Technology*. Available at [www.isca-speech.org/archive/interspeech\\_2005](http://www.isca-speech.org/archive/interspeech_2005).

- Diaz, M. T., & Swaab, T. Y. (2007). Electrophysiological differentiation of phonological and semantic integration in word and sentence contexts. *Brain Research*, *1146*, 85–100.
- Dijkstra, T., & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189–225). Mahwah, NJ: Erlbaum.
- Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*, 175–197.
- Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 663–679.
- Elston-Guettler, K. E., & Gunter, T. C. (2008). Fine-tuned: Phonology and semantics affect first- to second-language zooming in. *Journal of Cognitive Neuroscience*, *21*, 180–196.
- Elston-Guettler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Brain Research: Cognitive Brain Research*, *25*, 57–70.
- Elston-Guettler, K. E., Paulmann, S., & Kotz, S. A. (2005). Who's in control? Proficiency and L1 influence on L2 processing. *Journal of Cognitive Neuroscience*, *17*, 1593–1610.
- Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, *28*, 501–518.
- Grainger, J., & Dijkstra, T. (1992). On the representation and use of language information in bilinguals. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 207–220). Amsterdam: Elsevier Science Publishers.
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.
- Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Cambridge, MA: Harvard University Press.
- Grosjean, F. (1988). Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes*, *3*, 233–274.
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*, 3–15.
- Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech: Semantic ERP effects. *Neuropsychologia*, *38*, 1518–1530.
- Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, *30*, 251–266.
- Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, *4*, 123–141.
- Holcomb, P. J., & Neville, H. J. (1991). Natural speech processing - An analysis using event-related brain potentials. *Psychobiology*, *19*, 286–300.
- Ju, M., & Luce, P. A. (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science*, *15*, 314–318.
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences—Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.
- Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*, 161–163.
- Lemhöfer, K., Dijkstra, T., & Michel, M. C. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes*, *19*, 585–611.
- Li, P. (1996). Word recognition of code-switched words by Chinese-English bilinguals. *Journal of Memory and Language*, *35*, 757–774.
- Marian, V., & Spivey, M. J. (2003a). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics*, *24*, 173–193.
- Marian, V., & Spivey, M. J. (2003b). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, *6*, 97–115.
- Marian, V., Spivey, M. J., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and Language*, *86*, 70–82.
- Maris, E. (2004). Randomization tests for ERP topographies and whole spatiotemporal data matrices. *Psychophysiology*, *41*, 142–151.
- Marslen-Wilson, W. D., & Tyler, L. K. (1980). Temporal structure of spoken language understanding. *Cognition*, *8*, 1–71.
- McCallum, W. C., Farmer, S. F., & Pocock, P. V. (1984). The effects of physical and semantic incongruities on auditory event-related potentials. *Electroencephalography and Clinical Neurophysiology*, *59*(Suppl.), 477–488.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.
- McQueen, J. M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 621–638.
- Meara, P. M. (1996). *English vocabulary tests: 10k*. Swansea, UK: Centre for Applied Language Studies.
- Moreno, E. M., Rodriguez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics*, *21*, 477–508.
- Moss, H. E., & Marslen-Wilson, W. D. (1993). Access to word meanings during spoken language comprehension: Effects of sentential semantic context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1254–1276.
- Newman, R. L., Connolly, J. F., Service, E., & McIvor, K. (2003). Influence of phonological expectations during a phoneme deletion task: Evidence from event-related brain potentials. *Psychophysiology*, *40*, 640–647.
- Norris, D. (1994). Shortlist—A connectionist model of continuous speech recognition. *Cognition*, *52*, 189–234.
- Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1209–1228.
- O'Rourke, T. B., & Holcomb, P. J. (2002). Electrophysiological evidence for the efficiency of spoken word processing. *Biological Psychology*, *60*, 121–150.
- Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science*, *12*, 445–449.
- Praamstra, P., Meyer, A. S., & Levelt, W. J. M. (1994). Neurophysiological manifestations of phonological processing—Latency variation of a negative ERP component time-locked to phonological mismatch. *Journal of Cognitive Neuroscience*, *6*, 204–219.
- Rueschemeyer, S. A., Nojack, A., & Limbach, M. (2008). A mouse with a roof? Effects of phonological neighbors on processing of words in sentences in a non-native language. *Brain and Language*, *104*, 132–144.

- Scarborough, D. L., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. *Journal of Verbal Learning and Verbal Behavior*, *23*, 84–99.
- Schouten, M. E. H. (1975). *Native-language interference in the perception of second-language vowels: An investigation of certain aspects of the acquisition of a second language*. Unpublished doctoral dissertation, Utrecht University, The Netherlands.
- Schulpen, B., Dijkstra, T., Schriefers, H. J., & 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., 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.
- Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode—The effect on lexical access. *Memory and Cognition*, *12*, 380–386.
- Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language*, *45*, 412–432.
- Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10*, 281–284.
- Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language*, *27*, 324–340.
- Takashima, A., Jensen, O., Oostenveld, R., Maris, E., van de Coevering, M., & Fernández, G. (2006). Successful declarative memory formation is associated with ongoing activity during encoding in a distributed neocortical network related to working memory: A magnetoencephalography study. *Neuroscience*, *139*, 291–297.
- Tuladhar, A. M., Ter Huurne, N., Schoffelen, J. M., Maris, E., Oostenveld, R., & Jensen, O. (2007). Parieto-occipital sources account for the increase in alpha activity with working memory load. *Human Brain Mapping*, *28*, 785–792.
- Van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience*, *13*, 967–985.
- Van den Brink, D., Brown, C. M., & Hagoort, P. (2006). The cascaded nature of lexical selection and integration in auditory sentence processing. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 364–372.
- Van den Brink, D., & Hagoort, P. (2004). The influence of semantic and syntactic context constraints on lexical selection and integration in spoken-word comprehension as revealed by ERPs. *Journal of Cognitive Neuroscience*, *16*, 1068–1084.
- Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, *39*, 458–483.
- Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 394–417.
- Verhoeven, L., & Vermeer, A. (1985). Ethnic group differences in children's oral proficiency of Dutch. In G. Estra & T. Vallen (Eds.), *Ethnic minorities and Dutch as a second language* (pp. 105–132). Dordrecht, The Netherlands: Foris Publications.
- Vermeer, A. (1992). Exploring the second language learner lexicon. In L. Verhoeven & J. H. A. L. de Jong (Eds.), *The construct of language proficiency: Applications of psychological models to language assessment* (pp. 147–162). Amsterdam: John Benjamins.
- Vroomen, J., & De Gelder, B. (1995). Metrical segmentation and lexical inhibition in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 98–108.
- Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*, 1–25.
- Weber-Fox, C. M., & Neville, H. J. (1996). maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, *8*, 231–256.
- Zwitserslood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, *32*, 25–64.