



## Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming

Kim Verhoef\*, Ardi Roelofs, Dorothee J. Chwilla

*Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands*

### ARTICLE INFO

#### Article history:

Received 23 March 2007

Revised 12 October 2008

Accepted 23 October 2008

#### Keywords:

Bilingual

Control

ERP

Inhibition

N2

Speech production

### ABSTRACT

How are bilinguals able to switch from one language to another? The prevailing inhibition hypothesis takes larger reaction-time (RT) costs for switching to the first language (L1) than to the second language (L2) as evidence for suppression of the non-target language. Switch cost asymmetries can alternatively be explained by an L1-repeat-benefit, assuming selective absence of language competition in repeating L1. To test the latter hypothesis, RTs and event-related brain potentials (ERPs) were recorded when unbalanced Dutch–English bilinguals switched between overt picture naming in L1 and L2. Preparation time (short versus long cue–stimulus intervals) modulated the degree to which top-down inhibitory control biased language competition, as indexed by the N2 component of the ERP. RT switch costs were asymmetrical on short intervals and symmetrical on long intervals. Preparation interval effects were observed in all conditions except for L1 repeat trials. This pattern was present both for RT and N2 data, thereby supporting the L1-repeat-benefit hypothesis. These results suggest that inhibition is not necessary, but can modulate the efficiency of language switching.

© 2008 Elsevier B.V. All rights reserved.

### 1. Introduction

One of the most remarkable abilities of bilinguals is that of separating their two languages during speech production, especially when they switch between languages. Although bilinguals can choose from at least two response alternatives for any given concept, they are able to restrict their speech to one language only, rarely making cross-language intrusion errors (Poulish, 1999; Poulish & Bongarts, 1994). Even bilingually unbalanced individuals, whose native language (L1) is much stronger than their non-native language (L2), can speak one language without being much hindered by the other language. Especially when speaking their weaker L2, unbalanced bilinguals seem to have to overcome interference from the relatively automatic responses in their L1, see Kroll, Bobb, and Wodniecka

(2006). Bilingual language use is therefore an instance of very powerful attentional control in a naturalistic situation. Better understanding of how bilinguals pull off this impressive feat will be informative not only regarding bilingual language performance, but efficient attentional control in general.

A prominent account of bilingual performance is that bilinguals can speak one language by inhibiting the other (Green, 1998). A common paradigm used to study how bilinguals control their two languages is the language switching paradigm (e.g., Meuter & Allport, 1999) in which the response language of two subsequent trials can be the same (repeat trials) or different (switch trials). Typically, response latencies are longer on switch trials than repeat trials, called the switch cost. This switch cost can be reduced with sufficient preparation time (about 600 ms in task switching), but a residual switch cost usually remains (Rogers & Monsell, 1995).

Evidence from language switching has been taken as support for the inhibition hypothesis. Meuter and Allport (1999) asked unbalanced bilinguals to switch between

\* Corresponding author. Address: Department of Anatomy and Neuroscience, Vrije Universiteit medical centre (VUmc), P.O. Box 7057, 1007MB Amsterdam, The Netherlands.

E-mail address: [k.verhoef@vumc.nl](mailto:k.verhoef@vumc.nl) (K. Verhoef).

naming digits in L1 and L2. The response language was indicated by the background color of the screen (i.e., if blue name in L1, if red name in L2). Switch costs were larger when switching to the dominant L1 from the weaker L2 than vice versa, called asymmetrical switch costs. Meuter and Allport (1999) took this as support for the inhibition hypothesis, which explains the switch cost asymmetry in terms of task-set inertia (Allport, Styles, & Hsieh, 1994). That is, activation of the previously relevant task-set or inhibition of the previously irrelevant task-set carries over into the current trial. More specifically for the bilingual case, the inhibition hypothesis maintains that while speaking the target language, the non-target language needs to be inhibited or suppressed. The inhibition persists into the next trial. The amount of inhibition is dependent on the relative dominance of the two languages. Thus, when speaking in the weak L2, the more dominant L1 needs to be inhibited more than the other way around. Therefore, the amount of inhibition that needs to be overcome on an L1 switch trial (from L2 to L1) is larger than that on an L2 switch trial (from L1 to L2), resulting in asymmetrical switch costs.

Converging evidence for the inhibition hypothesis comes from an electrophysiological study by Jackson, Swainson, Cunnington, and Jackson (2001). In a color-cued digit naming task, they measured event-related potentials (ERPs) in addition to naming latencies. To prevent contamination of the EEG signal by muscle EMG associated with jaw and head movement during overt speech, naming latencies were measured from one subset of trials and EEG was recorded from a different subset of trials in which naming was delayed. In particular, responses were given after a go-signal. In contrast to Meuter and Allport (1999), Jackson et al. (2001) used the alternating runs paradigm (Monsell, 2003) in which languages are switched in a predictable fashion (e.g., every second trial), allowing bilinguals to prepare the response language in advance. The participant pool was heterogeneous, consisting of a mixture of balanced and unbalanced bilinguals that spoke English as a first language and one of five other languages as a second language. Again, naming latencies showed asymmetrical switch costs: larger switch costs for L1 than for L2. Especially on switch trials, naming is usually faster in L2 than in L1, called the “paradoxical” language effect (Christoffels, Firk, & Schiller, 2007; Costa & Santesteban, 2004; Kroll et al., 2006). The absence of such a paradoxical language effect for switch trials, and even faster responses for L1 than for L2 on repeat trials, led Jackson et al. (2001) to conclude that there is an asymmetrical repeat benefit, instead of an asymmetrical switch cost. The ERPs showed a frontocentral negativity peaking around 320 ms, which was taken to reflect an N2. Independent component analysis suggested that the N2 elicited by Jackson and colleague’s switch paradigm originates from the same neural source as the nogo N2 that is supposed to reflect response inhibition (e.g., Kopp, Mattler, Goertz, & Rist, 1996; Schmitt, Munte, & Kutas, 2000). The N2 amplitude was more negative going for L2 switch than repeat trials, an effect that was absent for L1. The ERP data were interpreted as a confirmation of the inhibition hypothesis: A language switch is brought about by inhibition of the non-target lan-

guage and more inhibition is needed to suppress L1 on L2 switch trials than to suppress L2 on L1 switch trials.

In this line of reasoning, the functional significance of the N2 is generalized from response inhibition to inhibition of a language as a whole. However, from this perspective, it is unclear why there is no N2 effect for L1 at all, since one would expect that on switch trials L2 is still active from the previous trial and needs to be inhibited. Furthermore, it cannot be excluded that the data were confounded with response inhibition needed to withhold the response in the delayed naming paradigm. In a task-switch study, Swainson et al. (2003) only found N2 effects for delayed response trials, not for trials in which participants were free to respond. Relevant for the present article is that the N2 component is not only elicited by tasks requiring inhibition of manual responses (e.g., Kopp, Rist, & Mattler, 1996; Schmitt, Schiltz, Zaake, Kutas, & Munte, 2001) but also by tasks requiring inhibition of vocal responses. Another important point for the present study is that the N2 not only reflects response inhibition, but also inhibition of languages as a whole.

Christoffels et al. (2007) investigated the effects of blocked picture naming (one language only) versus mixed picture naming (switching between L1 [German] and a weaker foreign language [Dutch]) using ERPs. Like in the studies of Meuter and Allport (1999) and Jackson and colleagues (2001) the response language was indicated by the color of the pictures. Christoffels et al. (2007) found that mixed language context affected especially the L1 by slowing down latencies and modulating the ERPs in the time window of the N2. Unexpectedly, however, they found more negative amplitudes to the non-switch than the switch trials, which they explain in terms of the unpredictability of the language switch in their study as opposed to the study of Jackson et al. (2001). This study suggests that inhibition plays some role in bilingual language control, but it is unclear exactly what role.

In a functional magnetic resonance imaging (fMRI) study, Wang, Xue, Chen, Xue, and Dong (2007) had Chinese learners of English switch between picture naming in L1 and in L2. Congruent with the results of Jackson et al. (2001), they found that switching to L2, but not switching to L1, activated brain regions related to attentional control (i.e., bilateral frontal cortices and left anterior cingulate cortex). From this they conclude that inhibition of the non-target language is especially important in switching to L2. Just like in the study of Jackson, it is unclear why there is no effect for L1 at all. From the inhibition hypothesis, one would expect that on L1 switch trials, L2 is still active from the previous trial and needs to be inhibited.

Costa and Santesteban (2004) replicated the switch cost asymmetry in response times for unbalanced bilinguals. In contrast to Meuter and Allport (1999), Costa and colleagues used a picture naming task. Once more, switch costs were larger for L1 than for L2. For balanced bilinguals, Green’s inhibition hypothesis predicts symmetrical switching costs since the relative dominance of the two languages, and therefore the amount of inhibition that needs to be overcome on a switch trial, is equal for both languages. The predicted symmetrical switch costs were observed by Costa and Santesteban (2004). Surprisingly, the

paradoxical language effect, slower responses in L1 than in L2, was present for both balanced and unbalanced bilinguals.

When balanced bilinguals switch between their L1 and a much weaker L3, the inhibition hypothesis predicts asymmetrical switch costs. More specifically, L1 needs to be suppressed more when speaking in L3 than vice versa resulting in relatively slow responses on L1 switch trials due to the large amount of inhibition that needs to be overcome. In contrast to this prediction, [Costa and Santesteban \(2004\)](#) found symmetrical switch costs when balanced bilinguals switched between naming in L1 and the much weaker L3. This finding led [Costa and Santesteban \(2004\)](#) to conclude that language selection is accomplished by different mechanisms in proficient and non-proficient bilinguals. They suggest that while unbalanced bilinguals use inhibition of the non-target language to speak the target language, balanced bilinguals do not need to recruit inhibition since they have developed a mechanism allowing language-specific lexical access. Note, however that, although the paradoxical L1 speed disadvantage for unbalanced bilinguals can readily be explained in terms of relative inhibition, under the language-specific selection hypothesis it is not clear why balanced bilinguals also show this effect. According to [Costa and Santesteban \(2004\)](#), the effect stems from a heightened L1 selection threshold that is strategically applied to bias selection in favor of L2.

### 1.1. *The L1-repeat-benefit hypothesis*

Do we really need to assume different proficiency dependent mechanisms that allow speaking one language rather than another? According to the inhibition hypothesis, the switch cost asymmetry for unbalanced bilinguals is due to disproportionately slow responding on L1 switch trials. On L1 switch trials, the relatively strong inhibition that was needed to speak L2 on the previous trial needs to be overcome. However, from this perspective there is no obvious reason why there should be an asymmetry. The benefit of repeating L2 might be just as large as the switch cost for L1 ([Yeung & Monsell, 2003](#)). When the L1 inhibition remaining from the previous trial has an adverse effect on L1 switch trials, it should be beneficial on L2 repeat trials, resulting in symmetrical switch costs.

An alternative way of explaining the larger switch cost for L1 compared to L2 is to assume that L1 repeat trials are disproportionately fast. When unbalanced bilinguals are well prepared to speak their L1, as is the case on repeat trials, the task-set of their weaker L2 is not activated and therefore L2 responses are not considered for selection (cf. [Kroll et al., 2006](#)). It should be noted that this situation is unique for L1 repeat trials in unbalanced bilinguals. On L1 switch trials, the L2 task-set is still active and therefore L2 responses compete for selection. Also, when naming in L2, unbalanced bilinguals suffer from interference from the stronger L1 task-set on both switch and repeat trials. Thus, in all conditions except L1 repeat, the non-target language competes for selection, resulting in relatively fast L1 repeat responses and thus asymmetrical switch costs. This will be referred to as the L1-repeat-benefit hypothesis. Only when a language-set is weakly established, when

the language is weaker to begin with or when the task-set of the other language is still active (as is the case on switch trials), there is language competition, slowing down responses.

In balanced bilinguals, both language-sets are well established and therefore both languages compete for selection, even on repeat trials. Therefore, balanced bilinguals do not show the L1-repeat-benefit. Even when switching between L1 and L3, L2 responses must be prevented from selection. Thus, the L1-repeat-benefit hypothesis explains why [Costa and Santesteban \(2004\)](#) observed asymmetrical switch costs for unbalanced bilinguals and symmetrical switch costs for balanced bilinguals, without postulating different, proficiency dependent mechanisms for speaking one language rather than another.

Language switching in unbalanced bilinguals is similar to switching between tasks of unequal strength. The idea that repeat trials in the dominant language are disproportionately fast due to a selective absence of task-set competition receives support from a task switch study of [Waszak, Hommel, and Allport \(2003\)](#). In a picture-word Stroop interference task, Waszak and colleagues manipulated pre-exposure of incongruent picture-word stimuli in the picture task, the word task or both tasks. Stimuli that were preexposed in both tasks showed the standard dominance asymmetry: larger switch costs for the stronger word reading task than for the weaker picture naming task. The asymmetry was reduced for stimuli that were only associated with the target task. Moreover, stimuli that were preexposed only in the target task were named faster than stimuli that were preexposed in both tasks. However, this was only true for switch and repeat trials in the weak picture naming task and for switch trials in the strong word naming trials, not for repeat trials in the strong word naming task. Preexposure of stimuli in only one task creates a bias for responding in that task by means of stimulus-task associations (see also [Koch and Allport \(2006\)](#)). In the study of Waszak and colleagues, the beneficial effect of biasing the stimulus-task associations was selectively absent for repeat trials in the stronger word naming task. This suggests that repeat trials in the stronger task do not suffer task-set interference in the first place (a ceiling effect). This would imply that switch cost asymmetries in tasks of unequal strength result from a selective absence of task competition for repeat trials in the stronger task, making them disproportionately fast. What is important to investigate is whether this explanation of asymmetrical switch costs generalizes to the situation of unbalanced bilinguals. This issue was addressed in the present study.

### 1.2. *The aim of the present study*

The aim of the present study is to examine the merits of the L1-repeat-benefit hypothesis of asymmetrical switch costs in switching between languages. In a language switching task, Dutch-English unbalanced bilinguals were asked to name pictures in the response language indicated by a pre-cue. Unlike previous bilingual research (e.g., [Christoffels et al., 2007](#); [Costa & Santesteban, 2004](#); [Meuter & Allport, 1999](#); [Jackson et al., 2001](#)) in which the required language was indicated by the color of the stimulus, we

used symbolic cues to indicate the target language in advance, on each trial. Instead of biasing stimulus-driven retrieval of the target task-set like Waszak and colleagues did, we allowed bilinguals to bias the target language-set by means of *endogenous* control. Koch and Allport (2006), who biased stimulus-driven retrieval of the non-target task, showed that the negative effects of stimulus-based task-set priming can be counteracted by cue-based preparation. Thus, in a standard language switch situation in which stimulus-driven retrieval of the language-set is not biased, cue-based preparation should be able to bias the target language endogenously. To manipulate the *readiness* of the target language-set, unbalanced bilinguals were randomly given short (750 ms) or long (1500 ms) periods of time to prepare the response language in advance. According to Rogers and Monsell (1995) preparation is optimal after 600 ms, thus switch costs are residual.

Although both intervals are sufficiently long for preparation of the target language, response times are predicted to be faster for long compared to short preparation intervals as was previously observed by several researchers (Coull & Nobre, 1998; Meiran, 1996). This is based on what is known about temporal attention (Coull & Nobre, 1998; Miniussi, Wilding, Coull, & Nobre, 1999; Naccache, Blandin, & Deheane, 2002). Advantages in task performance are reported for temporally predictable compared to unpredictable stimulus presentation. When there are two cue-stimulus intervals, omission of a stimulus at the short interval guarantees it would occur at the long interval. Participants can prepare to be ready at a later point in time, which results in faster responses and better task performance for long compared to short intervals (e.g., Coull & Nobre, 1998). We assume that the relative timing of the preparation intervals results in a strategic effect of preparation interval. Consequently, on long intervals more endogenous control over language selection can be exerted than on short intervals. Thus, it is not the absolute, but the relative length of the cue-stimulus interval that is responsible for the hypothesized interval effect. In this line of reasoning we propose that optimal cue-based preparation on long intervals, results in language-set readiness. Readiness in turn results in inhibition of the non-target language as soon as it is triggered by the stimulus. Based on experience due to picture familiarization, the stimulus has become associated with both languages that are triggered once the stimulus appears. We assume that the language-set competition that results from stimulus-language associations can be biased by inhibition of the non-target language once language-set readiness is established. In short, endogenous control consists of the readiness to respond in the target language that in turn allows strategic inhibition to bias stimulus-based language competition.

If this manipulation of endogenous cognitive control is successful and the L1-repeat-benefit hypothesis is correct, we predict a similar change in switch cost pattern as found by Waszak and colleagues. That is, for short preparation intervals we predict the standard asymmetrical switch cost pattern. However, for long preparation intervals when bilinguals can bias the target language by endogenous control, we predict that all conditions improve except L1 repeat which cannot improve further (due to a ceiling

effect) resulting in less asymmetrical or even symmetrical switch costs. Switch costs are assumed to arise during language or response selection processes, before response programming. This was assessed by manipulating the cognate status of the picture names. Cognate effects are assumed to arise during word-form encoding processes (i.e., response programming) in picture naming (e.g., Costa, Caramazza, & Sebastian Galles, 2000; Roelofs & Verhoef, 2006). Thus, switch costs and preparation patterns should be independent of cognate effects.

In addition to RTs, ERPs were recorded. Possibly, language competition is compensated for, or is biased, by means of strategically applied top-down inhibition of the non-target language. Language inhibition or response inhibition is assumed to be reflected by a specific ERP component, the N2 (e.g., Jackson et al., 2001; Schmitt et al., 2000, 2001). This claim is bolstered by the finding that trials on which inappropriate behavioral responses need to be inhibited reliably have been shown to elicit an N2, like nogo trials in the go-nogo paradigm, stop-trials in the stop task or incongruent trials in Stroop and flanker tasks (Kopp, Mattler, et al., 1996; Kopp, Rist, et al., 1996; Liotti, Woldorff, Perez, & Mayberg, 2000; Pliszka, Liotti, & Woldorff, 2000; Thorpe, Fize, & Marlot, 1996). The N2 has been proposed to index processes of inhibition both for manual responses as well as for vocal responses and for response inhibition as well as language inhibition (Jackson et al., 2001). In this context, the N2 is regularly reported to be right-shifted (Falkenstein, 2006; Pliszka et al., 2000; Schmajuk, Liotti, Busse, & Woldorff, 2006), with a maximum over right inferior frontal scalp regions. In accordance with this topography, sources of the N2 have been frequently localized in the right orbitofrontal cortex for nogo-tasks (Bokura, Yamaguchi, & Kobayashi, 2001; Lavric, Pizzagalli, & Forstmeier, 2004) and in the right dorsolateral prefrontal cortex for stop-signal tasks (Liotti et al., 2007). These findings are consistent with the activation in the right inferior frontal gyrus that is observed in fMRI studies using similar tasks (Aron, Fletcher, Bullmore, Sahakian, & Robbins, 2003; Aron, Monsell, Sahakian, & Robbins, 2004; Brown, Vilis, & Everling, 2008; Konishi et al., 1999). In a recent TMS study, Chambers et al. (2007) show that the right inferior frontal gyrus is crucial for inhibition. To get an indication of the amount of inhibition of the non-target language we measured ERPs (Jackson et al., 2001). Our target region for localizing inhibition N2 effects is frontocentral, including midline and right hemisphere sites.

If the preparation interval effect in the current study is due to the compensatory effect of language inhibition, the N2 amplitude is expected to be larger for long compared to short preparation intervals. If language competition is indeed absent for L1 repeat trials like the L1-repeat-benefit hypothesis predicts, not only should the interval effect be absent in the naming latencies, it should also be absent in the N2. More specifically, if there is no competition for L1 repeat trials to begin with, it cannot be compensated for. Therefore, in contrast to the N2 amplitude increase for long compared to short preparation intervals for the other conditions, the N2 amplitude for L1 repeat trials is not expected to differ between long and short preparation intervals.

The inhibition hypothesis also predicts slower responding on switch than repeat trials, possibly modulated by language (L1 versus L2) and the size of the preparation interval. Moreover, if the length of the preparation interval determines the amount of top-down inhibition that can be applied, larger N2 amplitudes are predicted for long than for short intervals. However, larger RT switch costs for switching to L1 than to L2 should go together with smaller N2 amplitudes for switching to L1 than to L2 (Jackson et al., 2001). This is because asymmetrical switch costs in RTs correspond to different needs of suppression. Under the inhibition hypothesis, more inhibition is needed to suppress L1 on L2 switch trials (reflected in larger N2 amplitudes) than to suppress L2 on L1 switch trials (reflected in smaller N2 amplitudes). Thus, the predictions of the L1-repeat-benefit and the inhibition hypotheses differ.

In the current study, it was important to allow overt naming from our participants in order to prevent confounding effects of response inhibition (N2) needed to temporarily withhold a response in the delayed naming paradigm. In the past, ERP researchers have refrained from measuring ERPs in overt production tasks. The main reason for this was to avoid a confounding of movement-related artifacts caused by the naming response with the ERP signal of interest. In line with this, Jackson et al. (2001) have recorded ERPs only to nogo trials. Of course it would be important to tap ERPs during the critical naming trials. Recent studies seem to indicate the feasibility of this approach (Christoffels et al., 2007; Liotti et al., 2000; Schmitt, Bles, Schiller, & Münte, 2002). In the present article, therefore, we recorded ERPs from overt picture naming trials. Here we tried to separate the relevant cognitive processes from movement-confounded activity as follows. Based on the language-switch literature we expect that naming latencies will vary between 500 and 1000 ms (Costa & Santesteban, 2004; Jackson et al., 2001; Meuter & Allport, 1999), while N2 effects have been reported in an earlier time window (200–400 ms) (e.g., Pfefferbaum, Ford, Weller, & Kopell, 1985). Hence, it should be possible to tease apart the relevant cognitive processes from those of the actual motor response in time. In line with the above mentioned studies (Christoffels et al., 2007; Liotti et al., 2000; Schmitt et al., 2002), we assumed that the ERPs in the early N2 window would be (more or less) free of motor artifacts, especially because naming latency distributions are typically positively skewed (e.g., Luce, 1986).

To summarize, the current study investigated the merits of the L1-repeat-benefit hypothesis. RTs were recorded while unbalanced bilinguals switched between overt picture naming in L1 and in L2. In order to bias the target language, the degree of inhibitory control was manipulated by means of short (low top-down control) and long (high top-down control) preparation intervals. The L1-repeat-benefit hypothesis predicts asymmetrical RT switch costs for short intervals and a symmetrical switch cost pattern for long intervals. The L1-repeat-benefit hypothesis predicts the N2 to be larger for long compared to short intervals for all conditions except for L1 repeat. In contrast, the inhibition hypothesis predicts that larger N2 amplitudes and smaller RT switch costs should go together.

## 2. Method

### 2.1. Participants

Nineteen students of the Radboud University of Nijmegen participated in return for either payment or course credits. All participants were right handed native Dutch speakers with normal or corrected-to-normal vision, who learned English as a second language from about the age of 11 (for self-rating measures on proficiency, see Appendix A). None of the participants had any neurological or psychological impairment or had used psychoactive medication. All participants gave informed consent. Data from two participants were eliminated, one on behavioral grounds (too many errors) and one due to excessive EEG artifacts. The final sample consisted of 17 participants, 15 female, aged from 18 to 25 years ( $M = 21.3$  years).

### 2.2. Materials, procedure and design

Stimuli were presented at the center of a black, 15-inch computer screen set to  $1024 \times 768$  pixel resolution, viewed at a distance of approximately 80 cm. A trial started with the 250 ms presentation of a cue, followed by a blank screen for a duration of 500 or 1250 ms. Then a picture stimulus appeared on the screen for 250 ms after which the screen blanked again for the response latency (triggering of a voice key), plus a latency jitter of 1500–2300 ms. If the voice key was not triggered or the participant did not respond within 3000 ms, the screen was blank for the 3000 ms timeout period plus the intertrial latency jitter. Then the next trial began, with the presentation of the next cue (see Fig. 1). Thus, the cue–stimulus interval was 750 ms (short) or 1500 ms (long), with the intertrial interval being variable.

The language cue was the Dutch or British national flag, 150 mm wide  $\times$  85 mm high and subtending a horizontal visual angle of  $10.6^\circ$  and vertical angle of  $6.1^\circ$ . Stimuli consisted of 48 black-and-white line drawings, taken from the international picture naming project (IPNP) database (Bates et al., 2003). Picture stimuli did not extend beyond an invisible square of 80 mm wide  $\times$  80 mm high and subtended a maximal visual angle of  $5.7^\circ$  horizontally and vertically. Half of the picture names were Dutch–English cognates (e.g., Dutch *appel*, English *apple*), the other half were non-cognates (e.g., Dutch *wortel*, English *carrot*). See Appendix B for further description. There were two types of trial sequences: switch and repeat. In switch trials the response language of the current trial was different from the response language of the previous trial (L1, L2 or L2, L1), while in repeat trials the response language of (at least) two subsequent trials was the same (L1, L1 or L2, L2). The trials were randomized such that the proportion of L1 and L2 switch trials was equal to the L1 and L2 repeat trials.

After signing the consent form, participants completed a bilingual proficiency questionnaire and were familiarized with the Dutch and English picture names while placement of the electrode cap took place. Thereafter, participants were seated in a soundproof chamber in front of the com-

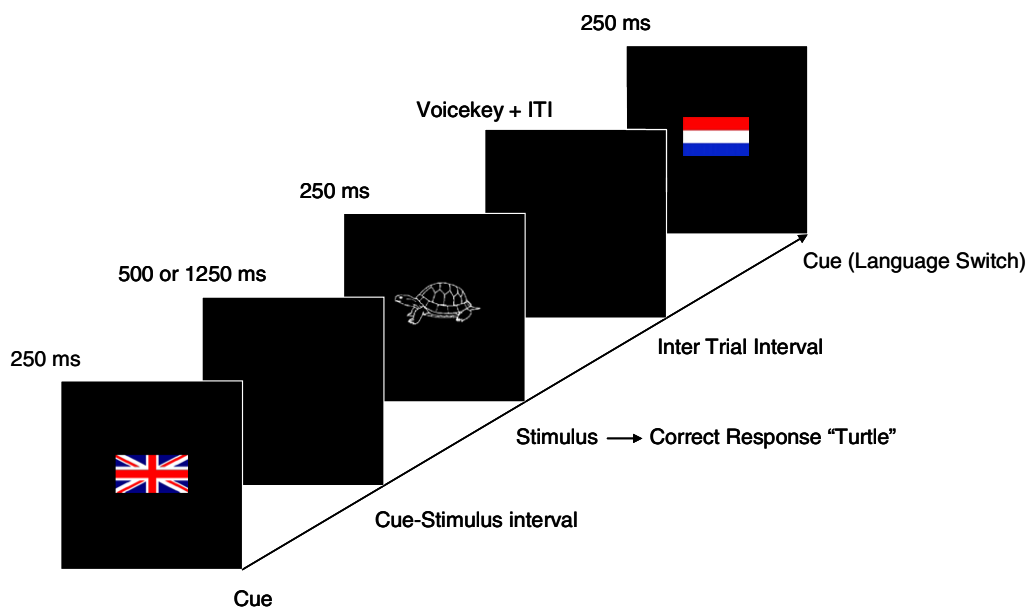


Fig. 1. Example of a trial sequence used in the experiment.

puter screen and were instructed to name the pictures as quickly and accurately as possible in the language indicated by the cue. They were further asked to minimize blinking until after picture naming. Naming latencies were registered with a 1 ms accurate voice key (1000 Hz). Cue-stimulus pairs were randomly presented in 16 blocks of 96 trials, with a total of 1536 trials. Each block lasted about 8 min. For every participant a new list of pseudorandom cue-stimulus pairs was generated with the constraint that pictures were never repeated within three trials and that language repetitions occurred no more than five times in a row. In each stimulus list every picture occurred four times in all eight conditions: *language* (L1 versus L2) by *language sequence* (switch versus repeat) by *preparation interval* (short versus long). Note that the design contains 16 conditions in total with the factor *cognate status* (cognate versus non-cognate) as a between picture manipulation. Participants were offered refreshments between blocks and could decide when they were ready to go on. The total testing session for each participant, including questionnaire, instructions, picture familiarization, cap application, and breaks took approximately 4 h.

### 2.3. EEG acquisition

EEG activity was recorded from the scalp with 61 tin electrodes mounted in an elastic electrode cap. The electrodes were arranged according to the extended international 10–20 system (Jasper, 1958). All electrodes were initially referenced to the left mastoid and later offline referenced to the average of the left and the right mastoid. The electro-oculogram (EOG) was recorded bipolarly; horizontal EOG was measured by placing electrodes on the outer canthus of each eye, vertical EOG by placing electrodes on the infra-orbital and the supra-orbital of the left eye. Electrode impedance was kept below 3 k $\Omega$ . Neuroscan

amplifiers (synamps) were used to amplify the EEG and EOG signals. All signals were sampled at 250 Hz and filtered on-line using a .05–70 Hz bandpass filter with an 8 second time constant.

### 2.4. Data analyses

For each participant, naming latencies and mean EEG signals were calculated for the correct trials only. Trials that were discarded from the analyses could be classified into four categories: (a) errors in language selection (utterances that started with the inappropriate response language), (b) within language errors (responses that differed from those designated by the experimenter in all but the response language), (c) trials that could not be classified as either switch or repeat (trials at the beginning of each block and trials following language selection errors), and (d) recording failures and time-outs (naming latencies shorter than 600 ms or longer than 2000 ms).

*Naming latency and error analyses:* The experimental design included the following within-subject factors: language (L1 versus L2), language sequence (switch versus repeat), preparation interval (short versus long), and cognate status (cognate versus non-cognate). Error rates and naming latencies were submitted to repeated-measures analyses of variance (ANOVAs) for subjects ( $F_1$ ) as well as items ( $F_2$ ). For the item analysis a mixed model ANOVA was used with cognate status as a between-item factor.

*ERP analyses:* Our analysis approach was similar to that of Poulsen, Luu, Davey, and Tucker (2005). In particular, analysis of stimulus-locked ERPs was guided by two methodological issues. First, as a result of preparatory effects in the preparation interval, the period prior to stimulus onset does not constitute a neutral baseline. However, slow waves resolve to baseline after presentation of the anticipated stimulus (Rockstroh, Elbert, Canavan, Lutzenberger,

& Birbaumer, 1989). Rather than using a pre-stimulus baseline, a 100-ms pre-cue baseline was used to reference an epoch that extends from cue onset to at least 1500 ms after stimulus onset. This was to avoid post-stimulus differences created by a mid-trial change in baseline (for details see Poulsen et al. (2005)). Second, the data were filtered offline with a 1–30 Hz bandpass filter to remove slow waves build up in the preparation interval that are superimposed on the stimulus-induced signal (Poulsen et al., 2005). Specific analysis steps were as follows. After rereferencing to the mean of both mastoids, the EEG signal was filtered (bandpass 1–30 Hz) and segmented into cue-locked –100 ms to 3000 ms epochs. These epochs include a stimulus interval that extends to at least 1500 ms after stimulus onset, depending on the length of the preparation interval. The epochs from cue onset to 3000 ms thereafter were referenced to a 100 ms pre-cue baseline. Before artifact rejection, epochs were segmented again into stimulus-locked intervals of 200 ms pre-stimulus to 600 ms after stimulus onset. This results in pre-cue baseline corrected stimulus-locked epochs. The relatively short stimulus interval was chosen to prevent severe reduction of artifact-free trials as a result of speech artifacts that start from about 600 ms after stimulus onset (the fastest responses in our study). Finally, stimulus-locked epochs were averaged.

The window for quantifying N2 was based on visual inspection of the waveforms and corresponded to the time window in which maximal differences between conditions occurred. This was the time window starting 30 ms before the mean peak of the N2 up to 30 ms after the peak, which in the following will be referred to as N2 window. The mean amplitudes for the different conditions (language sequence, preparation interval, and cognate status) were entered into separate ANOVAs for L1 and L2. Overall four-way ANOVAs were carried out with repeated measures on the three experimental factors and electrode site (61 levels).

The omnibus analyses were complemented with region of interest (ROI) analyses to quantify possible N2 effects. Based on the N2 literature, the ROI analyses for the N2 included frontocentral sites, in particular, two midline and eight right hemisphere electrodes (i.e., FCz, Cz, FC2, C2, FC4, C4, FC6, C6, FT8 and T8; see Fig. 2).

When appropriate, the estimated Greenhouse–Geisser coefficient  $\epsilon$  was used to correct for violations of the sphericity assumption (Geisser & Greenhouse, 1958). All reported  $p$ -values are based on corrected degrees of freedom, but to aid the reader in interpreting our statistical design, the stated degrees of freedom are uncorrected. Only comparisons significant at an alpha level of .05 are reported.

### 3. Results and discussion

#### 3.1. Naming latency results

To test the predictions outlined in the introduction, we analyzed switch cost patterns and interval effects. We also analyzed cognate effects.

*Switch costs patterns:* Fig. 3 shows the mean naming latencies for L1 and L2 as a function of language sequence

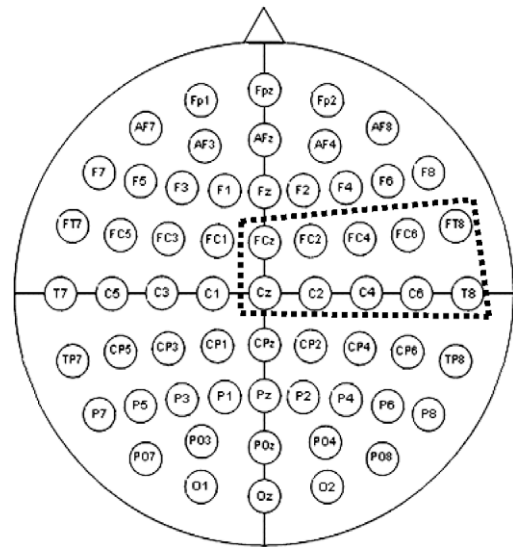
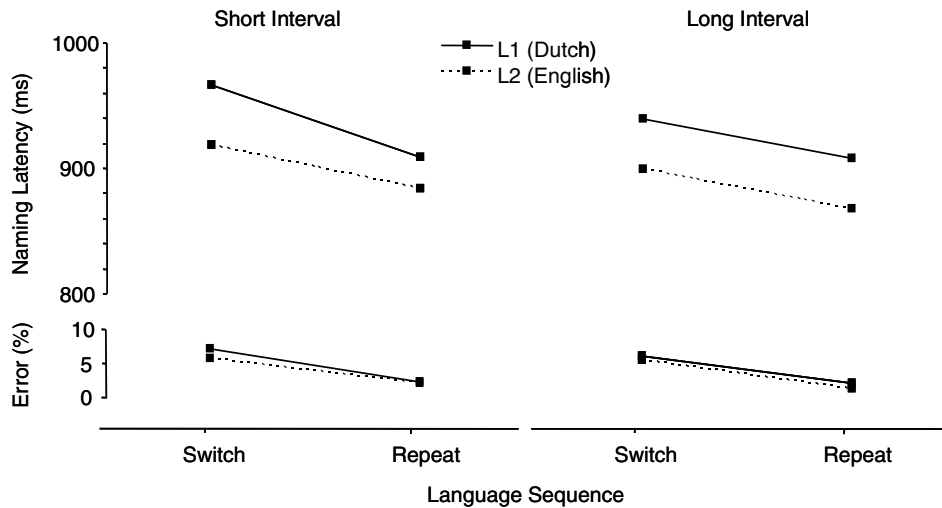


Fig. 2. Electrode positions enclosed in the frontocentral region of interest analysis, including two midline and eight sites over the right hemisphere.

separately for short and long preparation intervals. The statistical analysis yielded a main effect of language sequence,  $F_1(1,16) = 103.35$ ,  $p = .000$ ;  $F_2(1,46) = 183.15$ ,  $p = .000$ . Naming latencies on switch trials (930 ms) were longer than naming latencies on repeat trials (892 ms). There was also a main effect of language,  $F_1(1,16) = 43.58$ ,  $p = .000$ ;  $F_2(1,46) = 32.79$ ,  $p = .000$ . Naming latencies on L1 trials (930 ms) were longer than naming latencies on L2 trials (892 ms). The three-way interaction between language, language sequence, and preparation interval was significant as well,  $F_1(1,16) = 6.36$ ,  $p = .023$ ;  $F_2(1,46) = 4.26$ ,  $p = .045$ . Planned comparisons for short and long intervals were performed separately to explore whether switch cost patterns varied as a function of preparation interval. Since the interaction between language, language sequence, and preparation interval was independent of cognate status ( $F_1 < 1$ ,  $F_2 < 1$ ), naming latencies were collapsed over cognate status. For short intervals, switch costs were asymmetrical as indicated by a language by language sequence interaction  $F_1(1,16) = 10.87$ ,  $p = .005$ ;  $F_2(1,46) = 5.93$ ,  $p = .019$ . Although switch costs were larger for L1 (57 ms) than for L2 (34 ms), clear switch cost effects occurred for both languages, L1:  $F_1(1,16) = 80.99$ ,  $p = .000$ ,  $F_2(1,46) = 82.15$ ,  $p = .000$ ; L2:  $F_1(1,16) = 36.73$ ,  $p = .000$ ,  $F_2(1,46) = 49.63$ ,  $p = .000$ . There was no hint of a language by language sequence interaction for long intervals ( $F_1 < 1$ ,  $F_2 < 1$ ), indicating that switch costs were symmetrical. The magnitude of the switch costs was similar ( $\sim 31$  ms) and significant for both languages, L1:  $F_1(1,16) = 41.55$ ,  $p = .000$ ,  $F_2(1,46) = 31.69$ ,  $p = .000$ ; L2:  $F_1(1,16) = 25.14$ ,  $p = .000$ ,  $F_2(1,46) = 26.02$ ,  $p = .000$ . In sum, switch costs were asymmetrical on short intervals and symmetrical on long intervals.

*Interval effects:* Fig. 4 shows the mean naming latencies for L1 and L2 switch and repeat trials as a function of preparation interval. A main effect of preparation interval,  $F_1(1,16) = 7.86$ ,  $p = .013$ ;  $F_2(1,46) = 41.25$ ,



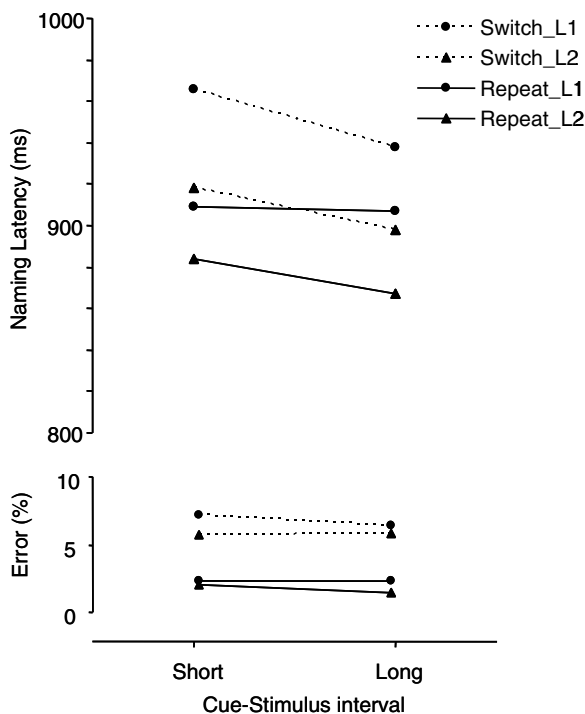
**Fig. 3.** Mean latencies and error percentages for both languages as a function of language sequence are depicted separately for short (left panels) and long (right panels) preparation intervals.

$p = .000$ , reflected that naming latencies for long intervals (903 ms) were shorter than those for short intervals (919 ms). Planned comparisons to determine the effect of preparation interval were also performed. Note that for these comparisons, we analyze the same three-way interaction as described above.

For switch trials, there was a 23-ms benefit for long over short preparation intervals,  $F_1(1,16) = 14.83$ ,  $p = .001$ ,  $F_2(1,46) = 40.09$ ,  $p = .000$ , which was independent of language ( $ps > .05$ ). A main effect of preparation interval was present for both languages, L1:  $F_1(1,16) = 16.99$ ,  $p = .001$ ,  $F_2(1,46) = 23.72$ ,  $p = .000$ ; L2:  $F_1(1,16) = 8.49$ ,  $p = .010$ ,  $F_2(1,46) = 26.30$ ,  $p = .000$ .

The interval effect for the repeat condition did approach significance,  $F_1(1,16) = 2.20$ ,  $p = .157$ ,  $F_2(1,46) = 6.84$ ,  $p = .012$ , as did the interaction with language,  $F_1(1,16) = 4.76$ ,  $p = .044$ ,  $F_2(1,46) = 3.21$ ,  $p = .080$ . Post hoc analysis showed that L2 repeat trials showed an interval effect of 17 ms in the same direction as the interval effect for switch trials,  $F_1(1,16) = 5.03$ ,  $p = .039$ ,  $F_2(1,46) = 11.19$ ,  $p = .002$ . However, there was no interval effect for the L1 repeat trials ( $F_1 < 1$ ,  $F_2 < 1$ ). In agreement with the L1-repeat-benefit hypothesis, the only condition that did not show a benefit of preparation interval is the L1 repeat condition.

**Cognate effects:** Fig. 5 shows naming latencies for L1 and L2 as a function of cognate status. There was a main effect of cognate status,  $F_1(1,16) = 132.69$ ,  $p = .000$ ;  $F_2(1,46) = 22.75$ ,  $p = .000$ , which interacted with language,  $F_1(1,16) = 64.99$ ,  $p = .000$ ;  $F_2(1,46) = 14.51$ ,  $p = .000$ . This interaction reflected that the cognate facilitation effect was larger for L1 (84 ms) than for L2 (31 ms). Nevertheless, the cognate facilitation effect was significant for both languages, L1:  $F_1(1,16) = 145.79$ ,  $p = .000$ ,  $F_2(1,46) = 29.35$ ,  $p = .000$ ; L2:  $F_1(1,16) = 42.40$ ,  $p = .000$ ,  $F_2(1,46) = 6.53$ ,  $p = .014$ . As indicated earlier, the cognate effect was independent of language sequence and interval effects.

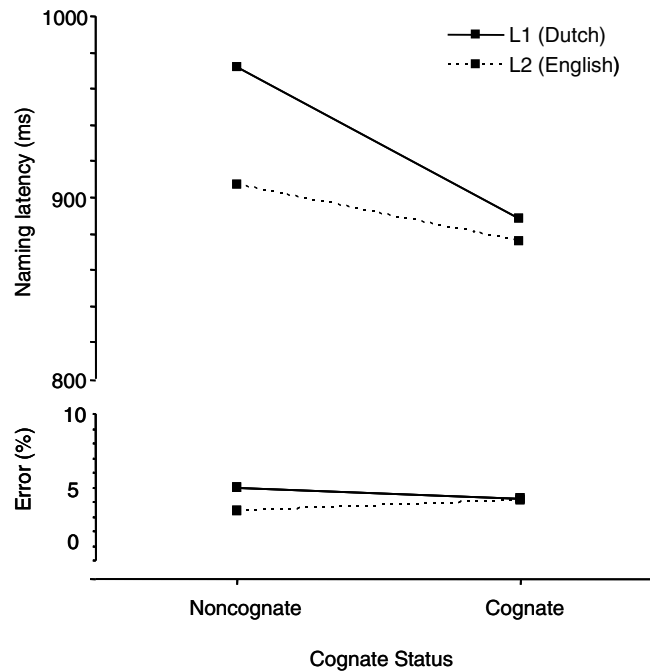


**Fig. 4.** Mean naming latencies (upper panel) and error percentages (lower panel) reflecting the effect of preparation interval as a function of language (L1 and L2) and language sequence (switch and repeat).

### 3.2. Error results

Following the exclusion criteria described above, 5.7% of all trials were discarded from the naming latency analyses. Time-out errors (1.6% smaller than 600 ms and 0.89% larger than 2000 ms) and recording errors were discarded from the error analysis. Only 4.1% of the remaining trials





**Fig. 5.** Mean naming latencies (upper panel) and error percentages (lower panel) for the two languages as a function of cognate status, representing cognate facilitation effects.

could be classified as errors of which 3.0% were language errors and 1.1% were all other errors that could occur within the designated language. In the analyses we collapsed across error types.

Switch trials (6.3%) led to more errors than repeat trials (2.0%),  $F_1(1,16) = 21.69$ ,  $p = .000$ ;  $F_2(1,46) = 396.24$ ,  $p = .000$  (see Fig. 3). Error proportions were greater for L1 (4.6%) than for L2 (3.7%),  $F_1(1,16) = 7.92$ ,  $p = .012$ ;  $F_2(1,46) = 7.95$ ,  $p = .007$  (see Figs. 3–5). There was a significant language by cognate status interaction,  $F_1(1,16) = 4.78$ ,  $p = .044$ ;  $F_2(1,46) = 5.34$ ,  $p = .025$  (see Fig. 5). In L1 less errors were made for cognates (4.2%) than for non-cognates (5.0%), while in L2 more errors were made for cognates (4.1%) than for non-cognates (3.4%). Follow-up analyses showed that this interaction was due to the difference in error percentages between the languages for non-cognates,  $F_1(1,16) = 8.09$ ,  $p = .012$ ;  $F_2(1,23) = 10.51$ ,  $p = .004$ , which was absent for cognates, ( $F_1 < 1$ ,  $F_2 < 1$ ). The differences in error percentages between cognates and non-cognates failed to reach significance for either language ( $ps > .05$ ). No other effects in the error analysis reached significance. There was no indication of a speed-accuracy trade-off.

### 3.3. ERP results

We examined the interval effect for ERPs, in particular for the N2 component to get an indication of the mechanism that allows endogenous control to compensate for stimulus-based task-set priming (Koch & Allport, 2006). Specifically, we used the N2 as an index of language inhibition, which according to the L1-repeat-benefit hypothesis should be larger for long than for short intervals for all con-

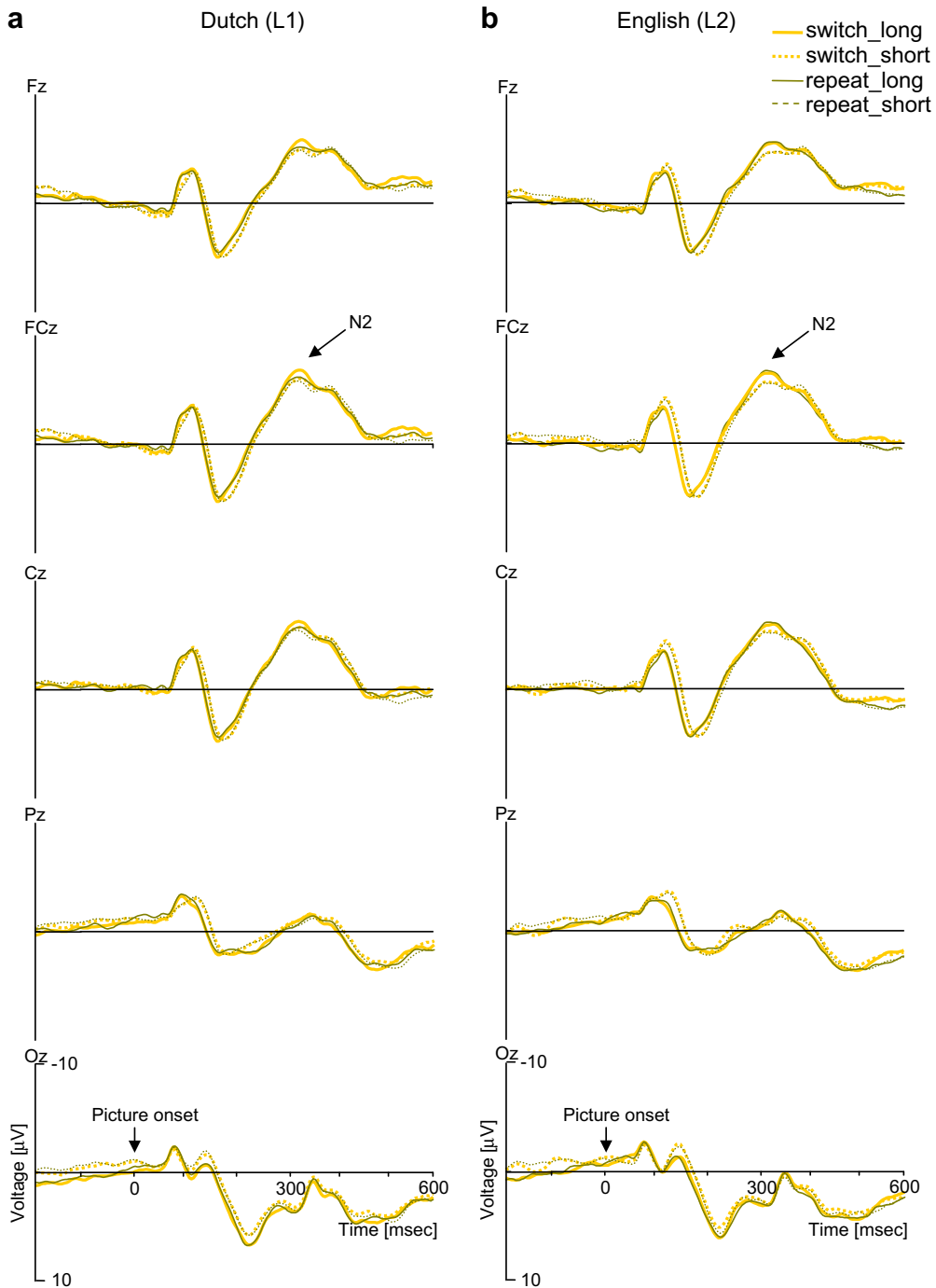
ditions except for L1 repeat trials. In contrast, the inhibition hypothesis predicts larger N2 amplitudes on switch than repeat trials and a correspondence of N2 amplitude with behavioral switch costs. Specifically, larger N2 amplitudes should be accompanied by smaller RT switch costs. Given the observed symmetrical RT switch costs for long intervals and the asymmetrical RT costs for short intervals, the inhibition hypothesis predicts a larger N2 for switching to L2 than to L1 for the short but not for the long intervals.

#### 3.3.1. Evaluating the L1-repeat-benefit hypothesis

In Figs. 6 and 7, the grand mean waveforms for the two levels of language sequence (switch-repeat) and preparation interval (long-short) are displayed separately for L1 and L2.<sup>1</sup> All conditions elicited the for visual stimuli characteristic early ERP response. That is, an N1 was followed by a P2, which at occipital sites was preceded by a P1 component.

**3.3.1.1. L1 waveforms.** Visual examination of switch trial ERPs (Fig. 7) suggests that the mean waveforms elicited by pictures that were preceded by long preparation intervals were a bit more negative going than those preceded by short preparation intervals between 300 and 360 ms. For repeat trials, the mean waveforms elicited by pictures preceded by short and long preparation intervals were

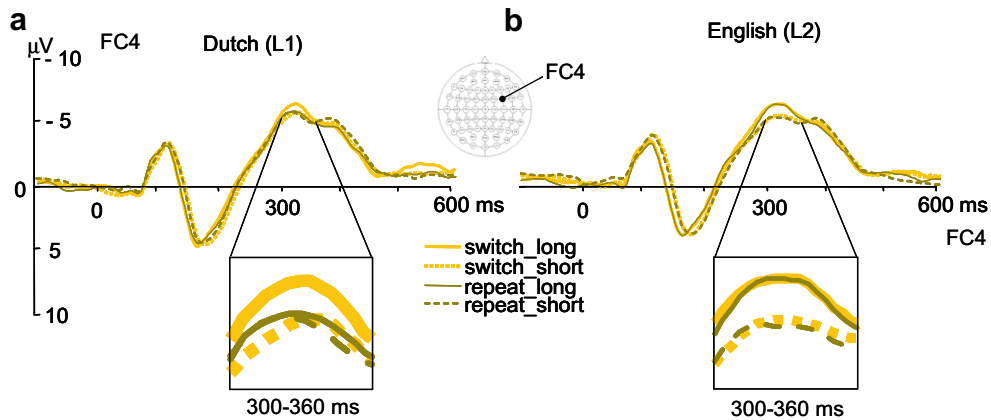
<sup>1</sup> Based on our hypotheses, analyses are performed separately for L1 and L2. The three-way interaction between language, language sequence and interval did not reach significance ( $F(1,16) = 2.62$ ,  $p = .125$ ) for the 300–360 ms time-window. Probably this reflects a power problem. In line with this, reducing the noise by using a smaller window of 25 ms around the peak yielded a significant interaction between language, language sequence and interval ( $F = 5.77$ ,  $p = .029$ ).



**Fig. 6.** Grand ERP averages for preparation interval (long-short) and language sequence (switch-repeat) are shown separately for L1 (a) and L2 (b) for a representative subset of anterior and posterior sites. Averages are time-locked to the onset of the picture and superimposed for the two levels of preparation interval and for the two levels of language sequence.

superimposed on the waveform elicited by pictures that were preceded by short preparation intervals on switch trials in the 300–360 ms time window. Visual inspection of the L1 waveforms suggests that a small N2 effect of preparation interval was present for switch trials, but not for repeat trials.

The omnibus analysis for L1 showed a preparation interval  $\times$  language sequence  $\times$  site interaction,  $F(60,960) = 3.01$ ,  $p = .037$ , which was independent of cognate status ( $F < 1$ ). The ROI analysis for the frontocentral electrodes of the midline and the right hemisphere showed a significant language sequence by interval interaction



**Fig. 7.** Grand ERP averages for preparation interval (long-short) and language sequence (switch-repeat) are time-locked to the picture and shown separately for L1 (a) and L2 (b) at the electrode site FC4.

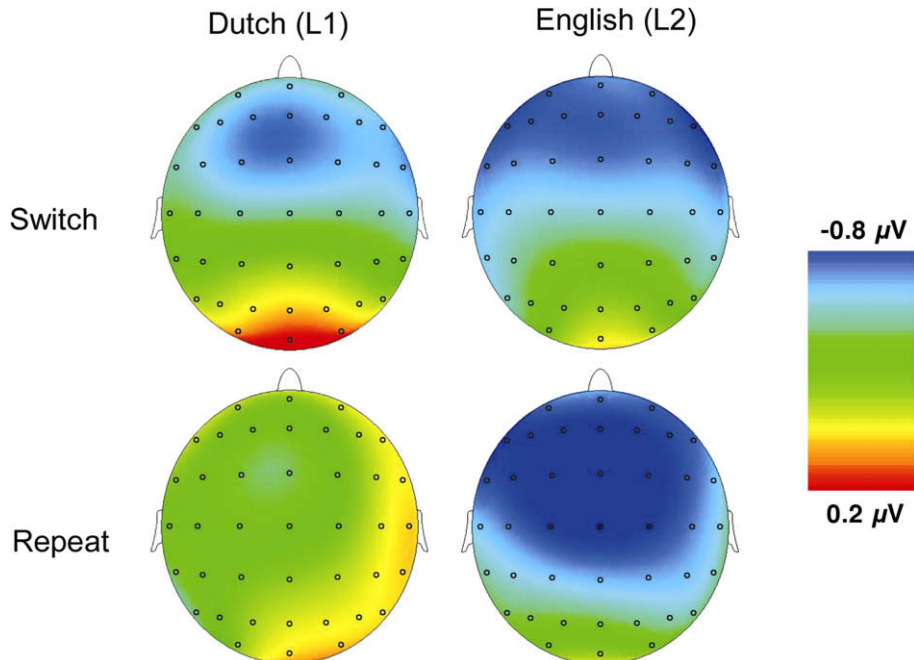
( $F(1,16) = 5.17, p = .037$ ). This two-way interaction between language sequence and interval could indicate that there are differences in N2 interval effects between switch and repeat trials. To test whether the interaction is indeed in the direction that is predicted by the L1-repeat-benefit hypothesis, separate analyses were conducted for switch and repeat trials. Since the effects of interest are independent of cognate status, in these analyses the data were collapsed over cognate status.

L1 switch trials on average show an  $0.5 \mu\text{V}$  preparation interval effect that was significant in the omnibus analysis [ $F(1,16) = 4.68, p = .046$ ] and the ROI analysis [ $F(1,16) = 5.96, p = .027$ ]. Mean N2 amplitude was more negative for long than for short intervals. The topographical differ-

ence plots in Fig. 8 show the scalp distribution of the interval effect. The effect is largest over anterior scalp sites, which is typical for the N2.

Most importantly, for the L1 repeat trials there was no indication for an interval effect in the omnibus analysis ( $F < 1$ ) or for the frontocentral ROI analysis ( $F < 1$ ). Nor was there an interaction between interval and electrode in the omnibus analysis that could have pointed to the presence of an N2 effect for L1 repeat trials. The analyses, therefore, support that no N2 effect occurred for L1 repeat trials.

**3.3.1.2. L2 waveforms.** Visual inspection of the L2 waveforms suggests that there was an N2 interval effect for both



**Fig. 8.** Scalp distribution maps obtained by interpolation from 61 sites at the peak (328 ms) in the 300–360 ms time window. Maps were computed from values resulting from the subtraction of short from long interval waves (N2 interval effects). The electrode locations are marked by black dots.

switch and repeat trials. The mean waveforms elicited by the pictures that were preceded by long preparation intervals appear to be more negative going than those preceded by short preparation intervals (see Figs. 6 and 7).

The analyses for L2 revealed main effects of interval [omnibus:  $F(1,16) = 11.34$ ,  $p = .004$ ; ROI:  $F(1,16) = 10.83$ ,  $p = .005$ ], reflecting that long intervals on average were 0.5  $\mu\text{V}$  more negative than short intervals. The effect of interval did not interact with language sequence or cognate status in either analysis (all  $F_s < 1$ ). The frontocentral distribution of the N2 interval effect for L2 switch and repeat trials is depicted in Fig. 8.

To summarize, the N2 was larger for long than short preparation intervals for all conditions except L1 repeat. These findings support the L1-repeat-benefit hypothesis.

### 3.3.2. Evaluating the inhibition hypothesis

**3.3.2.1. L1 waveforms.** No main effect of language sequence was found for L1 (omnibus and ROI analyses, all  $p_s > .05$ ). However, as described in the previous section, there was a three-way language sequence  $\times$  preparation interval  $\times$  electrode interaction in the omnibus analysis and a two-way language sequence  $\times$  preparation interval in the ROI analysis. These interactions could indicate that at least for the frontocentral ROI (midline and right hemisphere sites) the effect of language sequence was different for short than for long preparation intervals. To test whether the interaction indeed reflects a larger N2 switch-repeat effect for short than long intervals as the inhibition hypothesis predicts, separate analyses were performed for both levels of interval. In these analyses the data were collapsed over cognate status, since the effects of interest were not modulated by this factor.

For L1 long interval trials, the main effect of language sequence did not reach significance (omnibus and ROI analyses, all  $p_s > .05$ ). In the omnibus analysis, interactions with the factor language sequence also failed to reach significance ( $p_s > .05$ ). So there was no indication for the presence of an N2 language sequence effect for L1 long interval trials.

Also for L1 short interval trials, the main effect of language sequence failed to reach significance in the omnibus and ROI analyses ( $p_s > .05$ ). In addition, no interactions with the factor language sequence were found (all  $p_s > .05$ ), further supporting the absence of an N2 language sequence effect.

In short, the language sequence by interval interaction for L1 trials does not reflect a larger N2 switch-repeat effect for short than long intervals as predicted by the inhibition hypothesis.

**3.3.2.2. L2 waveforms.** For L2, no main effect of language sequence was found in the omnibus or the ROI analyses ( $p_s > .05$ ). Nor was there an interaction of language sequence with electrode in these analyses ( $p_s > .05$ ). Hence, we did not find any evidence that pointed to larger N2 effects for switch than repeat trials.

To sum up, the analyses of the factor language sequence for L1 and L2 failed to show larger N2 effects for switch compared to repeat trials. Thus, the predictions derived from the inhibition hypothesis are not supported by the present data.

## 4. General discussion

The prevailing inhibition hypothesis takes larger switch costs for L1 than for L2 as evidence for suppression of the non-target language. However, we argued that switch cost asymmetries can alternatively be explained by an L1-repeat-benefit hypothesis, assuming selective absence of language competition for L1 repeat trials. To test the latter hypothesis, RTs and ERPs were recorded when unbalanced bilinguals switched between overt picture naming in L1 and L2. Preparation time (short versus long cue-stimulus intervals) modulated the degree to which temporal attention influenced top-down language inhibition as indexed by the N2 component of the ERP.

In our study, RT switch costs were asymmetrical on short intervals and symmetrical on long intervals. Preparation interval effects were reflected in the RTs and N2 amplitude. Preparation interval effects were selectively absent for L1 repeat trials in both RT and N2 data, supporting the L1-repeat-benefit hypothesis. These results suggest that inhibition is not necessary, but can modulate the efficiency of language selection. Moreover, we replicated the paradoxical L1 disadvantage in reaction times. Switch costs and preparation patterns were independent of cognate status, which suggests that the effects arose in language selection, before response programming.

Our results have implications for the mechanism(s) underlying switch costs and effects of preparation interval and theories of language selection. These topics are discussed in separate sections.

### 4.1. What do RT switch cost patterns and interval effects reflect?

Our results suggest that switch cost patterns are not proficiency dependent, unlike what Costa and Santesteban (2004) suggested. Based on their findings that unbalanced show asymmetrical switch costs and balanced bilinguals show symmetrical switch costs, Costa and Santesteban (2004) claim that mechanisms of language selection are proficiency dependent as reflected by the different switch cost patterns. In the present study, both asymmetrical and symmetrical switch cost patterns were observed as a function of preparation interval in a single population of unbalanced bilinguals. As predicted, short preparation intervals elicited asymmetrical switch costs and long preparation intervals elicited symmetrical switch costs. We hypothesized that little or no language inhibition is exerted on short intervals, while it is exerted on long intervals to compensate for language competition that was evoked by the stimulus. This manipulation appeared to be successful and therefore we take the pattern of interval effects to indicate effects of endogenous control or readiness.<sup>2</sup> The interval manipulation modulates switch cost

<sup>2</sup> Note that based on the principles of temporal attention, the interval manipulation is dependent on relative timing instead of absolute timing. A similar bilingual language switch study performed in our lab using only short preparation intervals showed symmetrical switch costs (Verhoef, 2008). This indicates that short intervals of 750 ms are long enough to optimally prepare a language switch.

patterns (short interval: asymmetrical switch costs, long interval: symmetrical switch costs). This provides evidence for the idea that the larger switch cost for L1 compared to L2 in unbalanced bilinguals is due to a selective absence of language competition on L1 repeat trials.

According to the inhibition hypothesis of language selection, asymmetrical switch costs in unbalanced bilinguals are the result of a relatively large amount of inhibition that needs to be overcome on L1 switch trials. From this it can be predicted that the benefit of increasing preparation time (long compared to short intervals) would be disproportionately large for L1 switch trials. Our findings are in clear disagreement with this prediction. The preparation interval effect for switch trials was equal for both languages. In contrast, the preparation interval effect for repeat trials was language dependent. Although the preparation interval effect for L2 repeat trials was similar as that for switch trials, it was completely absent for the L1 repeat condition, supporting the L1-repeat-benefit hypothesis of asymmetrical switch costs. That is, readiness and in turn language inhibition seems to effectively overrule stimulus-driven language competition for all conditions except L1 repeat in which competition seems to be absent to begin with.

The claim that L1 repeat performance is at ceiling seems to be contradictory with the observation that this condition is not the fastest (see Fig. 3). However, this is because there seems to exist an experiment-wide bias against L1 or biasing of L2, resulting in the paradoxical language effect (Christoffels et al., 2007; Costa & Santesteban, 2004; Kroll et al., 2006). L1 repeat performance is at ceiling within the boundaries of this bias.

In a recent study, Costa, Santesteban, and Ivanova (2006) replicated the symmetrical RT switch cost for balanced bilinguals, except when they switched to their L4 or a newly learned language. However, in the latter situation, there was no paradoxical language effect, but the L1 was some 200 ms *faster* than the L4 or newly learned language, suggesting that this is an exceptional situation.

#### 4.2. What do the N2 effects reflect?

Previously, researchers refrained from measuring ERPs during overt speech to prevent contamination of the ERP signals of interest with movement-related artefacts. However, *nogo* trials are known to modulate the N2 amplitude (e.g., Pfefferbaum et al., 1985), which is problematic if the N2 is the ERP signal of interest. In our study, the ERP signature of the relevant cognitive processes was temporally separated from movement-related activity by hundreds of milliseconds. The N2 peaked at about 330 ms, while the minimum RT allowed in our analyses was 600 ms. Given this difference in the timing of the behavioral and ERP effects, we can confidently assume that the N2 in the present study was not confounded by movement artifacts.

The ERP data provide further evidence for the idea that readiness induced by endogenous control resolves task-set or language competition that resulted from interference elicited by the stimulus. Additionally, the ERP data suggest that the neural mechanism responsible for the preparation

benefit (of long compared to short intervals) is language inhibition as measured by the N2. However, the absence of an N2 amplitude increase for switch compared to repeat trials excludes an explanation in terms of reactive inhibition as proposed by Green (1998). According to Green (1998) language competition is resolved at the level of lexical candidates, where non-target lemmas are inhibited after they get activated (reactive inhibition). From the inhibition hypothesis, a larger N2 switch-repeat effect was predicted to be present for L2 compared to L1 reflecting the amount of inhibition of the non-target language. However, for short intervals on which an asymmetrical RT switch cost was observed, no N2 effect was present at all. Thus, our data do not support the inhibition hypothesis of language selection.

However, the N2 amplitude increase for long compared to short preparation intervals is consistent with the hypothesis that language inhibition can be strategically applied to compensate for language competition elicited by the stimulus. The N2 data converge with the naming latency data in that the N2 preparation interval effect was selectively absent for L1 repeat trials. This pattern suggests that in this condition the stimulus-driven induction of language competition was absent to begin with.

In this study, we took the N2 as reflection of inhibition. A different view is that the N2 reflects the degree of response conflict monitored by the anterior cingulate cortex (Nieuwenhuis, Yeung, Van Den Wildenberg, & Ridderinkhof, 2003; van Veen & Carter, 2002a, 2002b, but see Aarts, Roelofs, & van Turenout, 2008; Roelofs, van Turenout, & Coles, 2006). In this view more conflict is assumed to be present on switch than on repeat trials. Thus, based on the conflict monitoring account, the N2 is predicted to be larger for switch than repeat. In contrast to this prediction, the N2 did not seem to be sensitive to language sequence. With regard to the interval manipulation, the conflict monitoring hypothesis would predict the degree of conflict to be constant or to diminish with an increase in preparation time. Thus, the N2 should be similar between interval conditions or should be reduced on long compared to short intervals. We observed the opposite; the N2 is larger on long compared to short intervals. Therefore, we conclude that the N2 pattern observed in this study is more in agreement with the inhibition account than with the conflict account of the N2.

To summarize, our data do not support the idea that switch cost patterns reflect different mechanisms of language selection. The present data for unbalanced bilinguals support the notion that switch costs are a result of the interaction between readiness, language inhibition and stimulus-cued language conflict. Effects of preparation interval in both the behavioral and the ERP data are consistent with the idea that switch cost asymmetries reflect an L1-repeat-benefit instead of inhibition of the non-target language.

#### 4.3. What are the implications for theories of language selection?

Our data show that switch cost asymmetries are not dependent on language inhibition. From this we conclude

that inhibition might not be necessary for language selection (for a similar conclusion see Finkbeiner, Almeida, Janssen, and Caramazza (2006) and Phillip, Gade, and Koch (in press)). However, our data suggest that readiness can facilitate selection by resolving language competition by means of language inhibition. Thus, although inhibition is not a prerequisite for language selection, it can play a modulating role. We explained the switch cost asymmetry in terms of the L1-repeat-benefit hypothesis, which holds that language competition is absent in the special case of L1 repeat. In all other cases, L1 switch and L2 switch and repeat, language selection appears to be influenced by the relatively strong activation of the non-target language. Just like in monolingual speech, more conflict results in slower responses.

## 5. Summary and conclusions

To test the L1-repeat-benefit hypothesis of asymmetrical switch costs, we obtained RTs and ERPs when unbalanced bilinguals switched between overt picture naming in L1 and L2. Preparation time (long compared to short intervals) modulated the degree to which readiness and inhibitory control biased language competition (the target language is favored by inhibition of the non-target language), as indexed by the N2 component of the ERP. RT switch costs were asymmetrical on short preparation intervals and symmetrical on long preparation intervals. Preparation interval effects were selectively absent for L1 repeat trials in both RT and N2 data, supporting the L1-repeat-benefit hypothesis. These results suggest that inhibition is not necessary, but can strategically modulate the efficiency of language selection.

## Acknowledgements

We like to thank Judith Kroll and Lee Osterhout for helpful discussion and three anonymous reviewers for their constructive comments. Furthermore, we like to thank Jos Wittebrood, Hubert Voogd, Pascal de Water and Gerard van Oijen from the ERG group of the Donders Centre for Cognition for their technical assistance. The preparation of this article was supported by a VICI grant from the Netherlands Organization for Scientific Research (NWO) to Ardi Roelofs.

## Appendix A. Self-assessed proficiency for participants of this study

This appendix describes self-assessed proficiency scores, L2 use and age of onset information for all participants. A self-rating questionnaire was used to obtain proficiency scores. Participants needed to indicate how well their English (L2) skills (reading, writing, listening, and speaking) were compared to Dutch (L1). The scores are on a five point scale, in which 1 represents that L2 skills were just as good as L1 skills and 5 represents that L2 skills were much worse than L1 skills. On average, participants rated their proficiency for L2 compared to L1 as 2.75 (SD = .95). Scores for L2 use were also measured at a five

point scale, where 1 represents less than 1 h per week and 5 represents more than 10 h per week. Participants average L2 use score was 1.59 (SD = 1.05). Age of onset refers to the age at which participants started learning the L2; their mean age in years was 11.45 with a standard deviation of 1.39.

## Appendix B. Materials

Belgian Dutch as well as American English picture naming norms (Bates et al., 2003; Severens, Van Lommel, Ratinckx, & Hartsuiker, 2005) were used to select pictures with high name agreement in both languages (total mean = 95.5%). Dutch and English picture names for cognates and non-cognates were matched as closely as possible on number of syllables, number of phonemes and on lemma log frequencies that were obtained from the written sources of the CELEX database. Onsets of the Dutch and English picture names were matched on voicing, such that possible differences between Dutch and English naming latencies could not be explained by differences in voice-key sensitivity (e.g., /f/ and /s/ will have a later voice-key onset than /p/ and /t/).

	Non-cognates		Cognates	
	Dutch (L1) name	English (L2) name	Dutch (L1) name	English (L2) name
1	mes	knife	appel	apple
2	doos	box	arm	arm
3	bloem	flower	ballon	balloon
4	boom	tree	banaan	banana
5	bord	plate	brood	bread
6	muur	wall	dolfijn	dolphin
7	mier	ant	hamer	hammer
8	tandarts	dentist	pinguin	penguin
9	trap	stairs	kanon	cannon
10	paraplu	umbrella	kasteel	castle
11	paard	horse	fontein	fountain
12	strijkijzer	iron	maan	moon
13	spiegel	mirror	muis	mouse
14	schildpad	turtle	neus	nose
15	sleutel	key	olifant	elephant
16	vleermuis	bat	tijger	tiger
17	wortel	carrot	slang	snake
18	aap	monkey	kruis	cross
19	ananas	pineapple	spin	spider
20	aardbei	strawberry	ster	star
21	horloge	watch	tafel	table
22	kaars	candle	telefoon	telephone
23	kikker	frog	vaas	vase
24	knoop	button	zebra	zebra

## References

- Aarts, E., Roelofs, A., & van Turenout, M. (2008). Anticipatory activity in anterior cingulate cortex can be independent of conflict and error likelihood. *Journal of Neuroscience*, 28(18), 4671–4678.
- Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and Performance 15: Conscious and*

- nonconscious information processing (pp. 421–452). Cambridge, MA: The MIT Press.
- Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience*, 6(2), 115–116.
- Aron, A. R., Monsell, S., Sahakian, B. J., & Robbins, T. W. (2004). A componential analysis of task-switching deficits associated with lesions of left and right frontal cortex. *Brain*, 127, 1561–1573.
- Bates, E., D'Amico, S., Jacobsen, T., Szekeley, A., Andonova, E., Devescovi, A., et al (2003). Timed picture naming in seven languages. *Psychonomic Bulletin and Review*, 10(2), 344–380.
- Bokura, H., Yamaguchi, S., & Kobayashi, S. (2001). Electrophysiological correlates for response inhibition in a Go/NoGo task. *Clinical Neurophysiology*, 112, 2224–2232.
- Brown, M. R., Vilis, T., & Everling, S. (2008). Isolation of saccade inhibition processes: rapid event-related fMRI of saccades and nogo trials. *Neuroimage*, 39(2), 793–804.
- Chambers, C. D., Bellgrove, M. A., Gould, I. C., English, T., Garavan, H., McNaught, E., et al (2007). Dissociable mechanisms of cognitive control in prefrontal and premotor cortex. *Journal of Neurophysiology*, 98, 3638–3647.
- Christoffels, I. K., Firk, C., & Schiller, N. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, 1147, 192–208.
- Costa, A., Caramazza, A., & Sebastian Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26(5), 1283–1296.
- Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, 50, 491–511.
- Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 32(5), 1057–1074.
- Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: The neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *The Journal of Neuroscience*, 18(18), 7426–7435.
- Falkenstein, M. (2006). Inhibition, conflict and the Nogo-N2. *Clinical Neurophysiology*, 117, 1638–1640.
- Finkbeiner, M., Almeida, J., Janssen, N., & Caramazza, A. (2006). Lexical selection in bilingual speech production does not involve language suppression. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 32(5), 1075–1089.
- Geisser, S., & Greenhouse, S. W. (1958). An extension of boxes results on the use of the *F* distribution in multivariate-analysis. *Annals of Mathematical Statistics*, 29(3), 885–891.
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67–81.
- Jackson, G. M., Swainson, R., Cunnington, R., & Jackson, S. R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition*, 4(2), 169–178.
- Jasper, H. H. (1958). The ten–twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, 10(1), 371–375.
- Koch, I., & Allport, A. (2006). Cue-based preparation and stimulus-based priming of tasks in task switching. *Memory and Cognition*, 34(2), 433–444.
- Konishi, S., Nakajim, K., Uchida, I., Kiyo, H., Kameyama, M., & Miyashita, Y. (1999). Common inhibitory mechanism in human inferior prefrontal cortex revealed by event-related functional MRI. *Brain*, 122, 981–991.
- Kopp, B., Mattler, U., Goertz, R., & Rist, F. (1996). N2, P3 and the lateralized readiness potential in a nogo task involving selective response priming. *Electroencephalography and Clinical Neurophysiology*, 99, 19–27.
- Kopp, B., Rist, F., & Mattler, U. (1996). N200 in the flanker task as a neurobehavioral tool for investigating executive control. *Psychophysiology*, 33, 282–294.
- Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, 9(2), 119–135.
- Lavric, A., Pizzagalli, D. A., & Forstmeier, S. (2004). When 'go' and 'nogo' are equally frequent: ERP components and cortical tomography. *European Journal of Neuroscience*, 20(9), 2483–2488.
- Liotti, M., Pliszka, S. R., Perez, R., 3rd, Luus, B., Glahn, D., & Semrud-Clikeman, M. (2007). Electrophysiological correlates of response inhibition in children and adolescents with ADHD: Influence of gender, age, and previous treatment history. *Psychophysiology*, 44(6), 936–948.
- Liotti, M., Woldorff, M. G., Perez, R., III, & Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color–word interference effect. *Neuropsychologia*, 38, 701–711.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.
- Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22(6), 1423–1442.
- Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, 40, 25–40.
- Miniussi, C., Wilding, E. L., Coull, J. T., & Nobre, A. C. (1999). Orienting attention in time: Modulation of brain potentials. *Brain*, 122, 1507–1518.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140.
- Naccache, L., Blandin, E., & Deheane, S. (2002). Unconscious masked priming depends on temporal attention. *Psychological Science*, 13(5), 416–424.
- Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., & Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: Effects of response conflict and trial type frequency. *Cognitive, Affective, & Behavioral Neuroscience*, 3(1), 17–26.
- Pfefferbaum, A., Ford, J. M., Weller, B. J., & Kopell, B. S. (1985). ERPs to response production and inhibition. *Electroencephalography and Clinical Neurophysiology*, 60, 423–434.
- Phillip, A. M., Gade, M., & Koch, I. (in press). Inhibitory processes in language switching: Evidence from switching language-defined response sets. *European Journal of Cognitive Psychology*.
- Pliszka, S. R., Liotti, M., & Woldorff, M. G. (2000). Inhibitory control in children with attention-deficit/hyperactivity disorder: Event-related potentials identify the processing component and timing of an impaired right frontal response inhibition mechanism. *Biological Psychiatry*, 48, 238–246.
- Poulisse, N. (1999). *Slips of the tongue: Speech errors in first and second language production*. Amsterdam, Philadelphia: John Benjamins.
- Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, 15, 36–57.
- Poulsen, C., Luu, P., Davey, C., & Tucker, D. M. (2005). Dynamics of task sets: Evidence from dense-array event-related potentials. *Cognitive Brain Research*, 24(1), 133–154.
- Rockstroh, B., Elbert, T., Canavan, A., Lutzenberger, W., & Birbaumer, N. (1989). *Slow cortical potentials and behaviour*. Baltimore: Urban and Schwarzenberg.
- Roelofs, A., van Turenout, M., & Coles, M. G. (2006). Anterior cingulate cortex activity can be independent of response conflict in Stroop-like tasks. *PNAS*, 103(37), 13884–13889.
- Roelofs, A., & Verhoef, K. (2006). Modeling the control of phonological encoding in bilingual speakers. *Bilingualism: Language and Cognition*, 9(2), 167–176.
- Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124(2), 207–231.
- Schmajuk, M., Liotti, M., Busse, L., & Woldorff, M. G. (2006). Electrophysiological activity underlying inhibitory control processes in normal adults. *Neuropsychologia*, 44, 384–395.
- Schmitt, B. M., Bles, M., Schiller, N. O., & Münte, T. F. (2002). Overt naming in a picture–word interference task analysed with event-related potentials. *Journal of Cognitive Neuroscience (Supplement)*, 80.
- Schmitt, B. M., Munte, T. F., & Kutas, M. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology*, 37, 473–484.
- Schmitt, B. M., Schiltz, K., Zaake, W., Kutas, M., & Munte, T. F. (2001). An electrophysiological analysis of the time course of conceptual and syntactic encoding during tacit picture naming. *Journal of Cognitive Neuroscience*, 13(4), 510–522.
- Severens, E., Van Lommel, S., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, 119(2), 159–187.
- Swainson, R., Cunnington, R., Jackson, G. M., Rorden, C., Peters, A. M., Morris, P. G., et al (2003). Cognitive control mechanisms revealed by ERP and fMRI: Evidence from repeated task-switching. *Journal of Cognitive Neuroscience*, 15(6), 785–799.

- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522.
- van Veen, V., & Carter, C. S. (2002a). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, 14(4), 593–602.
- van Veen, V., & Carter, C. S. (2002b). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology & Behavior*, 77(4–5), 477–482.
- Verhoef, K. M. W. (2008). *Electrophysiology of language switching in bilingual speakers*. Unpublished doctoral dissertation. Radboud Universiteit Nijmegen.
- Wang, Y., Xue, G., Chen, C., Xue, F., & Dong, Q. (2007). Neural bases of asymmetric language switching in second-language learners: An ER-fMRI study. *Neuroimage*, 35, 862–870.
- Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus–task bindings in task-shift costs. *Cognitive Psychology*, 46, 361–413.
- Yeung, N., & Monsell, S. (2003). Switching between tasks of unequal familiarity: The role of stimulus–attribute and response-set selection. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2), 455–469.