Attentional inhibition in bilingual naming performance: evidence from delta-plot analyses

Ardi Roelofs*, Vitória Piai and Gabriela Garrido Rodriguez

Donders Centre for Cognition, Radboud University Nijmegen, Nijmegen, Netherlands

It has been argued that inhibition is a mechanism of attentional control in bilingual language performance. Evidence suggests that effects of inhibition are largest in the tail of a response time (RT) distribution in non-linguistic and monolingual performance domains. We examined this for bilingual performance by conducting delta-plot analyses of naming RTs. Dutch–English bilingual speakers named pictures using English while trying to ignore superimposed neutral Xs or Dutch distractor words that were semantically related, unrelated, or translations. The mean RTs revealed semantic, translation, and lexicality effects. The delta plots leveled off with increasing RT, more so when the mean distractor effect was smaller as compared with larger. This suggests that the influence of inhibition is largest toward the distribution tail, corresponding to what is observed in other performance domains. Moreover, the delta plots suggested that more inhibition was applied by high- than low-proficiency individuals in the unrelated than the other distractor conditions. These results support the view that inhibition is a domain-general mechanism that may be optionally engaged depending on the prevailing circumstances.

Keywords: attention, bilingualism, delta plots, inhibition, naming, response times

INTRODUCTION

Attentional control includes the ability to formulate goals and plans of action and to follow these while facing distraction. This ability is critical to normal human functioning and it is a hallmark of general intelligence (e.g., Wundt, 1904; Duncan, 2010). Attentional control plays a central role in human performance generally and language performance specifically (e.g., Roelofs, 2003, 2008). Bilingual language performance is an instance of powerful attentional control in a naturalistic situation. Although bilingual speakers can usually choose from at least two words for any given concept (i.e., one in each language), they are able to restrict their utterances to one language only. Even non-balanced bilinguals, whose first language (L1) is stronger than their second language (L2), can speak one language without apparently being much hampered by the other language. However, little is known about the mechanisms of attentional control in bilingual performance. The aim of the present article is to illuminate properties of these mechanisms. A better understanding of how bilinguals achieve attentional control over their languages will be informative not only regarding bilingual language performance, but also regarding efficient attentional control in general.

INHIBITION IN BILINGUAL PERFORMANCE

A prominent account of attentional control in bilingual language performance holds that inhibition is involved (Green, 1998). The inhibition is attentional, because it concerns a top-down goal-dependent modulation of language processes rather than a type of inhibition that is evoked bottom up by language perception (such as lateral inhibition present in several models of word recognition), see Aron (2007) for an extensive discussion. The notion of attentional inhibition has a long history in psychology. In the early days of experimental psychology, Wundt (the founder of modern scientific psychology and psycholinguistics) assumed that an attentional

inhibition mechanism, which he located in the frontal lobes of the human brain, modulates activity in a language network, assumed to be centered around perisylvian brain areas (e.g., Wundt, 1904). In contemporary psychology, inhibition has been proposed as a compulsory (Abutalebi and Green, 2007; Kroll et al., 2008) or an optional mechanism (Verhoef et al., 2009) of attentional control in bilingual language performance, although others argue against the assumption of inhibition (Finkbeiner et al., 2006).

The compulsory inhibition hypothesis holds that attentional control over the languages of a bilingual speaker is achieved by inhibiting the irrelevant language. Green (1998) assumed that inhibition in bilingual performance is reactive, that is, evoked in response to lexical activation. Consequently, the amount of inhibition that is applied depends on the magnitude of lexical activation in the non-target language. According to the hypothesis that inhibition is an option, inhibition is not the mechanism that achieves attentional control over the languages, but inhibition may be optionally engaged to increase the speed and accuracy of bilingual performance, depending on the prevailing circumstances (Verhoef et al., 2009).

Regardless of whether inhibition is compulsory or optional, little is known about the nature of the inhibitory mechanism. One possibility is that inhibition is a domain-general mechanism that is shared between linguistic and non-linguistic performance, as proposed by Abutalebi and Green (2007). If so, inhibition in bilingual language performance should share critical properties, such as its dynamics, with inhibition in other performance domains.

DYNAMICS OF INHIBITION IN SIMON, ERIKSEN, AND STROOP TASK PERFORMANCE

As concerns the dynamics of inhibition in non-linguistic domains, Ridderinkhof and colleagues (e.g., Ridderinkhof, 2002; Ridderinkhof et al., 2004, 2005) argued that attentional inhibition
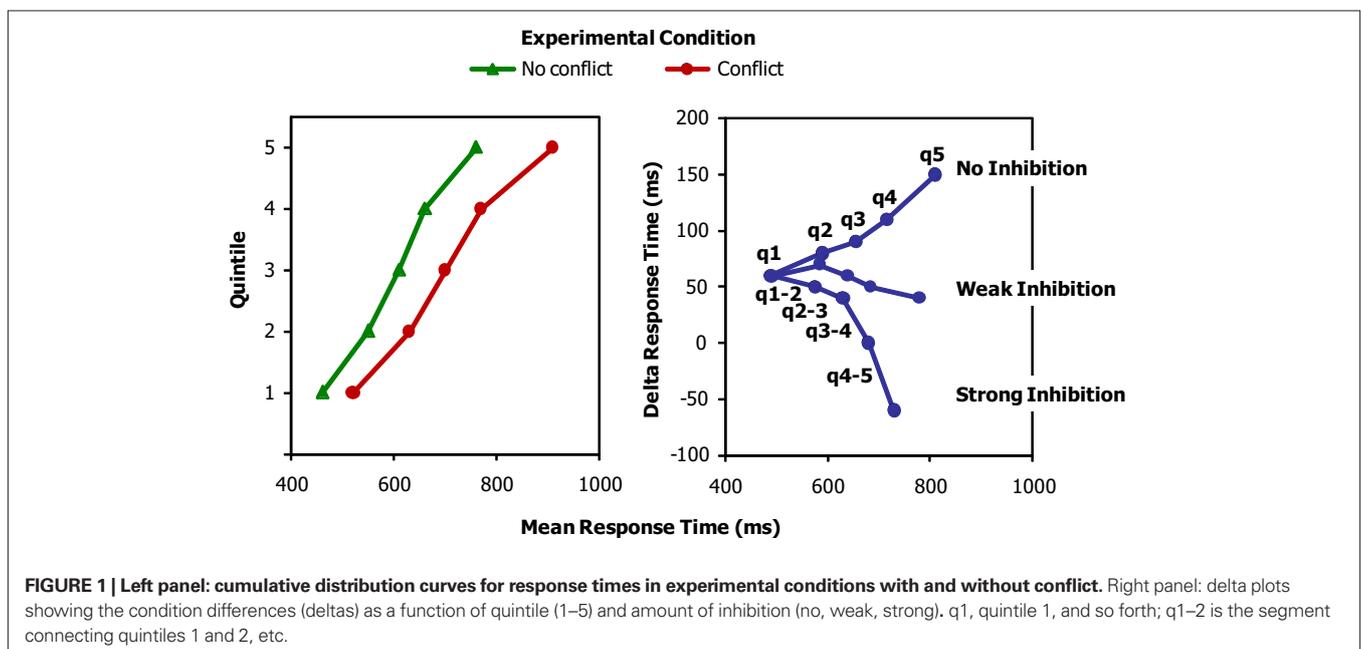
takes time to build up. Typically, in a situation with targets and distractors, interference from the distractors increases monotonically with response time (RT). Inhibition can decrease interference, but because inhibition is reactive and builds up slowly, its effects will be stronger for longer RTs. Consequently, effects of differential inhibition should be largest toward the tail of an RT distribution.

To assess the effect of differential inhibition, Ridderinkhof et al. (2004, 2005) constructed delta plots, which map out distractor condition differences as a function of RT. Delta plots prototypically have a positive slope, reflecting that the effect of an experimental factor tends to increase as a function of RT (cf. Luce, 1986). That is, the magnitude of factor effects tends to be larger for long than short RTs. However, if conditions differ in the amount of inhibition that is applied, a difference in RT between conditions should not increase linearly as a function of RT, but instead level off and become reduced for slow responses. If larger inhibition results in a more pronounced reduction of condition effects in slow responses, then the leveling off of the delta plot should be more pronounced in conditions that involve more inhibition. That is, the flattening of the delta plot should be more pronounced in experimental conditions that require more stringent inhibition compared to conditions that are less demanding.

Figure 1 illustrates the delta-plot logic. Suppose an experiment includes two conditions, one with conflict (red line) and one without conflict (green line). The left-hand panel of **Figure 1** illustrates the cumulative distributions of the two conditions by plotting the mean RT for each condition per quintile. The horizontal difference between the two distributions represents the difference in interference between the conditions. Typically, interference increases with RT. That is, the difference between the two curves becomes larger for higher quintiles. The right-hand panel of **Figure 1** shows the corresponding delta plots, which plot these condition differences per quintile against the mean of the two conditions for the corresponding quintile. The upper curve shows the situation when

no inhibition is applied to resolve the conflict: The interference becomes larger with increasing quintile. The middle curve shows that with weak inhibition, the interference decreases and the difference in RT between the two conditions becomes smaller. However, because inhibition is reactive and builds up slowly, its effect is stronger for longer RTs. Interference will therefore tend to level off for the higher quintiles of the delta plot. The lower curve shows that strong inhibition will decrease the interference further, which may even yield negative slopes for the segments connecting the higher quintiles (e.g., segment q4–5 connecting quintiles 4 and 5). To conclude, differences in delta plots index differences in the amount of inhibition that is applied in experimental conditions.

Ridderinkhof and colleagues (Ridderinkhof, 2002; Ridderinkhof et al., 2005) provided evidence from RT distributional analyses of non-linguistic Simon and Eriksen task performance that delta plots leveled off more with increased inhibition (for reviews, see Proctor et al., 2011; Van den Wildenberg et al., 2011). In the Simon task, participants indicate the identity of a left or right stimulus by pressing a left or right button, whereby the left–right position of the stimulus on a trial can be congruent (e.g., left stimulus requiring a left button press) or incongruent (e.g., left stimulus requiring a right button press). RTs are typically longer on incongruent than congruent trials, called the Simon effect. Ridderinkhof (2002) observed that delta plots leveled off more in participants with relatively small Simon effects (presumed to reflect strong inhibition) than in participants with relatively large Simon effects (presumed to reflect weaker inhibition). In an arrow version of the Eriksen task, participants have to indicate the identity of a left- or right-pointing target arrow flanked by incongruent or congruent distractor arrows on each side (e.g.,>><>> or <<<<<, respectively). RTs are typically longer on incongruent than congruent trials, called the Eriksen flanker effect. Ridderinkhof et al. (2005) observed that the leveling off of the delta plot for the Eriksen flanker effect was more pronounced for normal participants than for participants

FIGURE 1 | Left panel: cumulative distribution curves for response times in experimental conditions with and without conflict. Right panel: delta plots showing the condition differences (deltas) as a function of quintile (1–5) and amount of inhibition (no, weak, strong). q1, quintile 1, and so forth; q1–2 is the segment connecting quintiles 1 and 2, etc.

diagnosed with ADHD, which were assumed to have an inhibition deficit. The results from the distributional analyses of RTs in Simon and Eriksen task performance support the claim of Ridderinkhof and colleagues that inhibition takes time to develop during a trial in non-linguistic domains.

Similar results have been obtained in monolingual Stroop task performance. In the color–word Stroop task, individuals name the ink color of printed congruent or incongruent color words (e.g., the words GREEN or RED in green color; say "green") or neutral Xs. Mean RT is typically longer on incongruent than neutral trials, and often shorter on congruent than neutral trials. In performing delta-plot analyses on Stroop color naming RTs, Bub et al. (2006) obtained evidence that younger children (7–9 years) engage in stronger inhibition than older children (9–11 years), indicating that inhibition may also be engaged in non-linguistic domains. Moreover, delta-plot analyses of RTs in a manual version of the Stroop task by Sharma et al. (2010) revealed that the presence of a passively observing confederate during task performance leads to stronger inhibition as compared with the absence of a confederate. This influence of social context was observed when the preparation interval between consecutive Stroop stimuli was long (1000 ms), but not when the interval was short (32 ms). These results suggest that Stroop task performance may engage an inhibition mechanism, whose effects build up slowly, in line with the observations of Ridderinkhof et al. (2004, 2005). The effect of social context (Sharma et al., 2010) suggests that inhibition is an optional rather than compulsory mechanism.

Forstmann et al. (2008a) observed a strong link between individual differences in delta-plot parameters for Simon task performance and activity in right inferior frontal cortex (see Aron et al., 2004, for a review of the imaging literature localizing inhibition to right inferior frontal cortex). Moreover, when individual RT distribution parameters were used to classify subgroups of good and poor inhibitors based on a median split of the slowest segment of the delta plots, it was observed that individuals with better inhibition abilities showed higher brain connectivity values for white matter tracts in right inferior frontal cortex than poorer inhibitors. These results corroborate the assumption that delta plots reflect the operation of an attentional inhibition mechanism whose influence builds up slowly and is implemented in right inferior frontal cortex (see also Forstmann et al., 2008b), in line with Wundt's (1904) suggestion.

DYNAMICS OF INHIBITION IN BILINGUAL PICTURE NAMING
Evidence that attentional inhibition builds up slowly in bilingual language performance was provided by Verhoef et al. (2009). They measured event-related electrical brain potentials (ERPs) while Dutch–English bilingual participants named pictures in their first or second language, whereby the naming language on consecutive trials could be the same or different. The target language was indicated by a cue that preceded the picture by a short (500 ms) or long (1250 ms) preparation interval. In the ERPs time-locked to the picture onset, Verhoef et al. (2009) observed a right-lateralized N2 effect, which was argued to reflect activity in right inferior frontal cortex and to index inhibition. The N2 magnitude was modulated by the size of the preparation interval (larger N2 for long than short intervals) but not by language (L1 vs. L2) or language switching. So, the experiment provides evidence for an effect of time to pre-pare on the amount of inhibition applied during the planning of the picture name. This suggests that inhibition in picture naming takes time to develop, in line with the observations of Ridderinkhof et al. (2004, 2005). Given the timing evidence and the evidence that the N2 was frontal right-lateralized, suggesting a right inferior frontal locus, there is a clear similarity with the claims concerning the time course and right inferior frontal locus of the inhibition derived from delta plots. The effect of preparation interval corresponds to what Sharma et al. (2010) observed for manual Stroop task performance. The absence of an effect of language and switching on the N2 suggests that inhibition is an optional rather than compulsory mechanism.

Whereas Verhoef et al. (2009) argued that inhibition is used depending on the prevailing circumstances, Costa and Santesteban (2004) argued that the use of inhibition depends on language proficiency. According to them, attentional control over the two languages is accomplished by different mechanisms in linguistically balanced (i.e., equal proficiency in L1 and L2) and non-balanced bilinguals. In their view, while non-balanced 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. This claim of Costa and Santesteban (2004) was based on their observation of asymmetrical RT switch costs in language switching for non-balanced bilinguals and symmetrical switch costs for balanced bilinguals in picture naming. An RT switch cost means longer RTs when the language of the current trial is different from that of the previous trial (switch trials) than when the language is the same (repeat trials). Costa and Santesteban (2004) observed that balanced Spanish–Catalan speakers had equal switch costs in picture naming for switching to L1 and to L2, whereas non-balanced bilinguals had larger switch costs for switching to their L1 than to their L2. This finding suggests that for non-balanced bilinguals, using L2 involves stronger inhibition of L1 than using L1 involves inhibition of L2. Consequently, it will take longer to overcome the previous inhibition in switching to L1 than to L2.

However, Verhoef et al. (2009) observed that both asymmetrical and symmetrical RT switch cost patterns may occur as a function of preparation interval in a single population of non-balanced Dutch–English bilinguals. In their study, short preparation intervals elicited asymmetrical switch costs and long preparation intervals elicited symmetrical switch costs. This suggests that the engagement of inhibition may counteract the negative effect of lower proficiency in L2 than L1, provided that the preparation interval is long enough for inhibition to be engaged. The engagement of inhibition during the long preparation interval was reflected in the right-lateralized N2.

OUTLINE OF THE PRESENT STUDY
We report a bilingual picture–word interference study in which we further examined the nature of inhibition in bilingual language performance. In particular, we tested the prediction derived from the claim by Ridderinkhof et al. (2004, 2005) that effects of differential inhibition should be largest in the tail of an RT distribution. In our study, Dutch–English non-balanced bilingual participants named target pictures in their second language English (L2) while trying to ignore written Dutch distractor words (L1) superimposed onto

the pictures. The target pictures and distractor words were semantically related (e.g., picture of a rabbit, say "rabbit"; distractor word HERT, *deer*), unrelated (e.g., word STOEL, *chair*), or translation equivalents (e.g., KONIJN, *rabbit*). In addition, a series of Xs was superimposed on the picture in the control condition. Previous research has demonstrated semantic interference effects (i.e., longer RTs on semantically related than unrelated trials), lexicality effects (i.e., longer RTs on unrelated than control trials), and translation effects (i.e., longer RTs on unrelated than translation trials), see Costa et al. (1999) and Hermans et al. (1998), among others.

The interference effects from distractor words suggest that there are differential needs for inhibition on the different trial types. In particular, if the amount of inhibition depends on the magnitude of lexical activation (cf. Green, 1998), more inhibition is required on semantically related than unrelated trials and on unrelated than neutral and translation trials. If inhibition is a mechanism of attentional control in bilingual performance, and inhibition is a domain-general mechanism with a dynamics as assessed by Ridderinkhof et al. (2004, 2005), then the differential need for inhibition on the different trial types should be reflected in the tail of the corresponding RT distributions. In particular, the magnitude of the semantic, lexicality, and translation effects should decrease with increasing RT, more so with smaller effect sizes (presumed to reflect strong inhibition) than with larger effect sizes (presumed to reflect weaker inhibition), following Ridderinkhof (2002). In contrast, if inhibition is not applied in the present study or has different dynamic properties than inhibition in other performance domains (e.g., in the Simon, Eriksen, and Stroop tasks), then condition differences in RT should monotonically increase with increasing RT, and delta plots should not level off differentially among trial types depending on mean distractor effect.

In addition, we compared the delta plots between high- and low-proficiency individuals. It has been argued that bilingual individuals have enhanced inhibition abilities compared to monolingual individuals, as assessed for the non-linguistic Simon task (Bialystok et al., 2004) and the Eriksen flanker task (e.g., Costa et al., 2008). The idea is that the constant need to inhibit one of the two languages (i.e., the non-target one) provides bilinguals with an enhanced ability to ignore distracting and irrelevant stimuli, not only in linguistic tasks, but also in non-linguistic ones. However, Colzato et al. (2008) found no evidence for enhanced inhibition when comparing monolingual and bilingual individuals with regard to stop signal performance, inhibition of return, and the attentional blink, which all three can be argued to tap into aspects of inhibition. This suggests that the enhanced inhibition is restricted to Stroop-like circumstances, such as present in the Simon and Eriksen tasks. According to Green (1998), lexical competitors are more highly activated for high-proficiency as compared with low-proficiency bilinguals, and so these lexical competitors require a greater degree of inhibition. This implies that the inhibition ability develops along with skill in L2. If the inhibition ability depends on the level of skill in L2 and is reflected in performance on Stroop-like tasks, one would expect enhanced inhibition for high- compared with low-proficiency individuals in the present picture–word interference experiment. In particular, the delta plots should level off more for high- than low-proficiency individuals. Furthermore, if delta plots level off differentially

depending on effect type (i.e., semantic, translation, lexicality), this would suggest that individuals engage inhibition differently depending on the prevailing circumstances.

MATERIALS AND METHODS
PARTICIPANTS
The experiment was carried out with a group of 16 students of Radboud University Nijmegen, the Netherlands. Their mean age was 23.3 years (SD = 2.86). All participants were native speakers of Dutch with normal or corrected-to-normal vision, who learned English as a second language. The students participated in return of either payment or course credits.

A 17-item self-rating questionnaire was used to obtain proficiency scores. This questionnaire was composed of three main parts. In the first section (four items), participants indicated how well their English (L2) skills (reading, speaking, writing, and listening) were compared to Dutch (L1). The scores were 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 worse than L1 skills. On average, participants rated their proficiency for L2 compared to L1 as 2.94 (SD = 0.84). Thus, the participants were bilingually non-balanced. The second section (eight items) measured participants' use of L2 in different situations. 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. On average, participants L2 use score was 1.92 (SD = 0.76). The last part of the questionnaire (five items) evaluated age of onset (which refers to the age at which participants started learning English) and number of years of L2 use. The mean onset of L2 was at 10.63 years of age (SD = 1.86) and L2 was used, on average, for 12.65 years (SD = 3.46).

MATERIALS AND DESIGN
From the picture gallery of the Max Planck Institute for Psycholinguistics (Nijmegen, the Netherlands), 32 pictured objects from eight different semantic categories (i.e., clothing, animals, transportation, buildings, weapons, service, furniture, and body parts) were selected together with their basic-level names in English and Dutch. **Table 1** lists the English target picture names and the Dutch translation equivalents that were used as distractor words. These pictures and distractors were chosen because they yielded clear semantic, lexicality, and identity effects in earlier studies (e.g., Roelofs, 2003, 2008; Roelofs and Verhoef, 2006). We tried to avoid the use of cognates. The pictures were white line drawings on a black background and they were digitized and scaled to fit into a virtual frame of 10 cm × 10 cm. The printed words were presented in white color in 36-point lowercase Arial font.

There was one independent variable concerning the picture–word stimuli, referred to as distractor type, with four levels: semantically related, unrelated, translation, and control. Each picture was combined with a distractor from the same semantic category (the semantically related condition), with a word from another semantic category (the unrelated condition), with a word that was the Dutch translation equivalent of the English picture name (the translation condition), or with a string of Xs (the control condition). The distractor conditions were created by recombining pictures and words. For example, the picture of a car (say "car") was combined with the Dutch word FIETS (*bicycle*) in the semantic condition,

Table 1 | Basic-level names of the pictures in English (the target language) and their Dutch translation equivalents.

English name	Dutch name	English name	Dutch name
car	auto	cup	beker
bicycle	fiets	plate	bord
airplane	vliegtuig	bowl	kom
truck	vrachtwagen	jug	kan
toe	teen	coat	jas
leg	been	sweater	trui
nose	neus	skirt	rok
ear	oor	dress	jurk
deer	hert	castle	kasteel
swan	zwaan	mill	molen
rabbit	konijn	factory	fabriek
turtle	schildpad	church	kerk
table	tafel	dagger	dolk
cupboard	kast	sword	zwaard
desk	bureau	rifle	geweer
chair	stoel	tomahawk	bijl

with the Dutch word STOEL (*chair*) in the unrelated condition, with the Dutch translation equivalent AUTO (*car*) in the translation condition, and the Xs in the control condition. All target pictures and distractor words occurred equally often in each distractor type condition and they were repeated three times, yielding 384 trials in total. The order of presenting the stimuli across trials was randomized for each participant.

PROCEDURE AND APPARATUS

The participants were tested individually. They were seated in front of a CRT computer monitor and a microphone connected to an electronic voice key. The distance between participant and screen was approximately 70 cm, and the distance between participant and microphone was approximately 18 cm. Before the experiment began, participants were given a booklet that contained the set of experimental pictures and their names. They were asked to go through it in order to be familiarized with the pictures and their appropriate English names. After a participant had read the instructions, a block of 32 practice trials was administered in which the experimental pictures, combined with a row of Xs, were presented once and named in English. After this, testing began. A trial started with the presentation of a picture combined with a Dutch distractor word or the Xs for 250 ms followed by a black screen that lasted 2250 ms. The pictures were presented in the center of the screen and the distractor words were presented in the center of the pictures. The vocal response latency was measured to the nearest millisecond from target stimulus onset (with a time-out of 2500 ms). The presentation of stimuli and the recording of responses were controlled by Presentation Software (Neurobehavioral Systems, Albany, CA, USA).

DATA ANALYSIS

For each trial, the experimenter coded the response for errors. Five types of incorrect responses were distinguished: wrong pronunciation of the word, wrong response word (e.g., the response word was given in Dutch instead of English), disfluency, voice key triggered by a non-speech sound, and recording failures. Incorrect responses and RTs below 100 ms were discarded from the statistical analyses of the RTs. The mean RTs were submitted to analyses of variance. The analyses were performed both by participants (F_1) and by items (F_2). Distractor type was tested within participants and within items. In addition, the responses coded as correct or incorrect were submitted to binomial logistic regression analysis with distractor type as predictor (cf. Jaeger, 2008). An alpha level of 0.05 was used for all statistical tests. Following common parlance in the literature, we refer to statistical findings with *p*-values of between 0.05 and 0.10 as "marginally significant."

To obtain the delta plots, the RTs for each participant and distractor type were rank ordered and divided in five RT quintiles (bins) of equal or near-equal size. Next, the mean RT was determined for each quintile in each distractor condition. The delta plots for the semantic, translation, and lexicality effects were obtained by computing, for each quintile, the RT difference between, respectively, the semantically related and unrelated conditions (the semantic effect), unrelated and translation conditions (the translation effect), and the unrelated and control conditions (the lexicality effect).

To assess whether more inhibition leads to smaller distractor effects, we computed the magnitude of the semantic, translation, and lexicality effects for each participant. Next, median splits were made based on distractor effect sizes, referred to as *smaller* and *larger*, and delta plots were derived for each group (smaller, larger) and type of effect (i.e., semantic, translation, and lexicality). Similarly, median splits were made on the basis of proficiency (i.e., the mean scores for the self-ratings of skill in L2 in the first section of the questionnaire), referred to as *higher* vs. *lower*, and delta plots were derived for each group (higher, lower) and type of effect (i.e., semantic, translation, and lexicality). Because of equal skill scores for some participants, the median split yielded two proficiency groups of unequal size: There were nine participants in the high-proficiency group and seven participants in the low-proficiency group. The mean proficiency scores of the high- and low-proficiency groups were 2.31 and 3.75, respectively, which differed significantly, $F(1,14) = 46.52$, $p < 0.001$.

To assess whether the delta plots were different depending on mean distractor effect (smaller, larger) or proficiency (higher, lower), analyses of variance were performed on the slopes of the delta plots as a function of mean distractor effect size and proficiency, following Ridderinkhof and colleagues (Ridderinkhof, 2002; Ridderinkhof et al., 2004, 2005), Bub et al. (2006), and Sharma et al. (2010). For these analyses, slopes were computed for the delta-plot segments connecting the data points of consecutive quintiles (q1–2, q2–3, q3–4, and q4–5). The slope of the line segment connecting quintiles 1 and 2 was defined as the delta of quintile 2 minus that of quintile 1 divided by the difference in mean RT (across the two conditions used to compute the delta value) between quintile 2 and quintile 1 (e.g., De Jong et al., 1994; Ridderinkhof, 2002; Ridderinkhof et al., 2004, 2005). In a similar manner, the slopes of the other segments were determined. To assess whether the slopes were different depending on mean distractor effect or proficiency, analyses of variance were conducted on the slopes of consecutive quintile pairs with the within-participants factor quintile pair (q1–2, q2–3, q3–4, q4–5) and the between-participants factor effect size (smaller, larger) or proficiency (lower, higher). Following earlier

research, the delta plots were constructed such that the RT values on the horizontal axis were the means of the RTs in the two conditions used to compute each delta value (De Jong et al., 1994; Ridderinkhof, 2002; Ridderinkhof et al., 2004, 2005; Bub et al., 2006; Sharma et al., 2010).

RESULTS

MEAN PERFORMANCE

Table 2 presents the mean RTs for correct trials, their SD, and mean percentages of errors for each distractor type. RTs were longer for the semantically related and unrelated conditions compared with the translation and control conditions. Interference was found for semantically related distractors relative to unrelated ones (the mean semantic interference effect was 51 ms), and for unrelated distractors relative to control distractors (the mean lexicality effect was 69 ms). Moreover, interference was found for the unrelated condition relative to the translation condition (the mean translation effect was 95 ms). More errors were made when participants had to name a picture combined with a semantically related distractor than when the picture was presented combined with any of the other distractors.

The statistical analysis of the naming RTs yielded main effects of distractor type, $F_1(3,45) = 46.80$, $p < 0.001$, $\eta_p^2 = 0.76$; $F_2(3,93) = 53.97$, $p < 0.001$, $\eta_p^2 = 0.64$. Pairwise comparisons showed significant results for the semantic interference effect,

Table 2 | Mean response time (MRT, in milliseconds), standard deviations (SD), and percentage error (PE) per distractor type.

Distractor type	MRT	SD	PE
Unrelated	860	222	5.2
Semantically related	911	245	8.6
Translation	765	211	2.9
Control	791	215	3.6

$t_1(15) = 4.71$, $p < 0.001$, $t_2(31) = 4.02$, $p < 0.001$; for the translation effect, $t_1(15) = 5.46$, $p < 0.001$, $t_2(31) = 8.19$, $p < 0.001$; and for the lexicality effect, $t_1(15) = 5.05$, $p < 0.001$, $t_2(31) = 6.90$, $p < 0.001$.

Logistic regression analyses of the error percentages revealed that the log-odds of having a correct response in the unrelated condition were 1.71 higher compared to giving a correct response in the semantically related condition (the semantic effect), β coefficient = 0.54, SE = 0.15, Wald $Z = 3.64$, $p < 0.001$. When a translation equivalent was presented as distractor, the log-odds of having a correct response in comparison to an unrelated distractor (i.e., the translation effect) were 1.85 times higher, β coefficient = 0.62, SE = 0.19, Wald $Z = 3.19$, $p = 0.001$. Finally, when a control distractor was presented, the log-odds of having a correct response were 1.47 times higher than when an unrelated distractor was presented (the lexicality effect), β coefficient = 0.39, SE = 0.18, Wald $Z = 2.14$, $p = 0.032$. These analyses show that the effects in the errors are in the same direction as the RT effects, which indicates that there is no speed–accuracy trade-off in the data.

DELTA-PLOT ANALYSES

The groups (created by median splits based on distractor effect size) differed in the magnitude of the semantic effect (16 vs. 81 ms), $F(1,14) = 26.43$, $p < 0.001$, the translation effect (48 vs. 115 ms), $F(1,14) = 13.87$, $p = 0.002$, and the lexicality effect (25 vs. 114 ms), $F(1,14) = 28.98$, $p < 0.001$. **Figure 2** gives the delta plots for the semantic, translation, and lexicality effects as a function of relative mean effect size (smaller, larger). The figure shows that the magnitude of the semantic, translation, and lexicality effects increases with increasing RT when the mean effect size is relatively large (presumed to reflect weak inhibition), whereas the magnitude of the effects levels off when the mean effect size is relatively small (presumed to reflect stronger inhibition). The fact that the delta plots for the smaller effect sizes leveled off (i.e., they are negative going) suggests that inhibition was present.
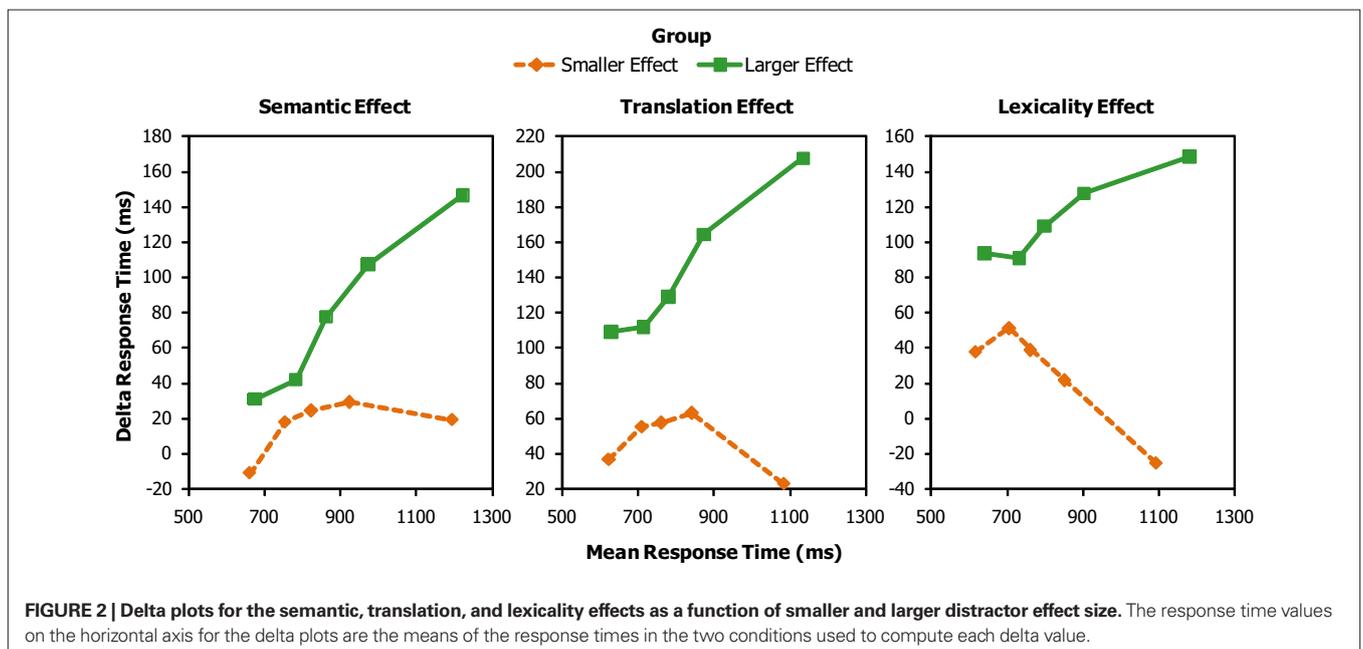

FIGURE 2 | Delta plots for the semantic, translation, and lexicality effects as a function of smaller and larger distractor effect size. The response time values on the horizontal axis for the delta plots are the means of the response times in the two conditions used to compute each delta value.

The difference in delta plots between the larger and smaller mean distractor effects was confirmed by statistical analyses of the slopes. For the semantic effect, the slopes differed between effect sizes significantly for the segment q2–3, $F(1,14) = 7.93$, $p = 0.014$, MSE = 0.056, but not for the other segments, q1–2, $F(1,14) = 1.55$, $p = 0.23$, MSE = 0.088; q3–4, $F(1,14) = 2.20$, $p = 0.16$, MSE = 0.087; q4–5, $F(1,14) = 0.68$, $p = 0.43$, MSE = 0.136. For the translation effect, the difference in slope as a function of effect size was marginally significant for the segment q4–5, $F(1,14) = 4.26$, $p = 0.058$, MSE = 0.085, but not for the other segments, q1–2, $F(1,14) = 1.18$, $p = 0.30$, MSE = 0.058; q2–3, $F(1,14) = 2.19$, $p = 0.16$, MSE = 0.061; q3–4, $F(1,14) = 2.73$, $p = 0.12$, MSE = 0.107. Finally, for the lexicality effect, the slopes differed as a function of effect size for the segment q2–3, $F(1,14) = 14.82$, $p = 0.002$, MSE = 0.053, and the segment q3–4, $F(1,14) = 5.32$, $p = 0.037$, MSE = 0.093, but not for the other segments, q1–2, $F(1,14) = 0.49$, $p = 0.50$, MSE = 0.109; q4–5, $F(1,14) = 2.59$, $p = 0.13$, MSE = 0.086.
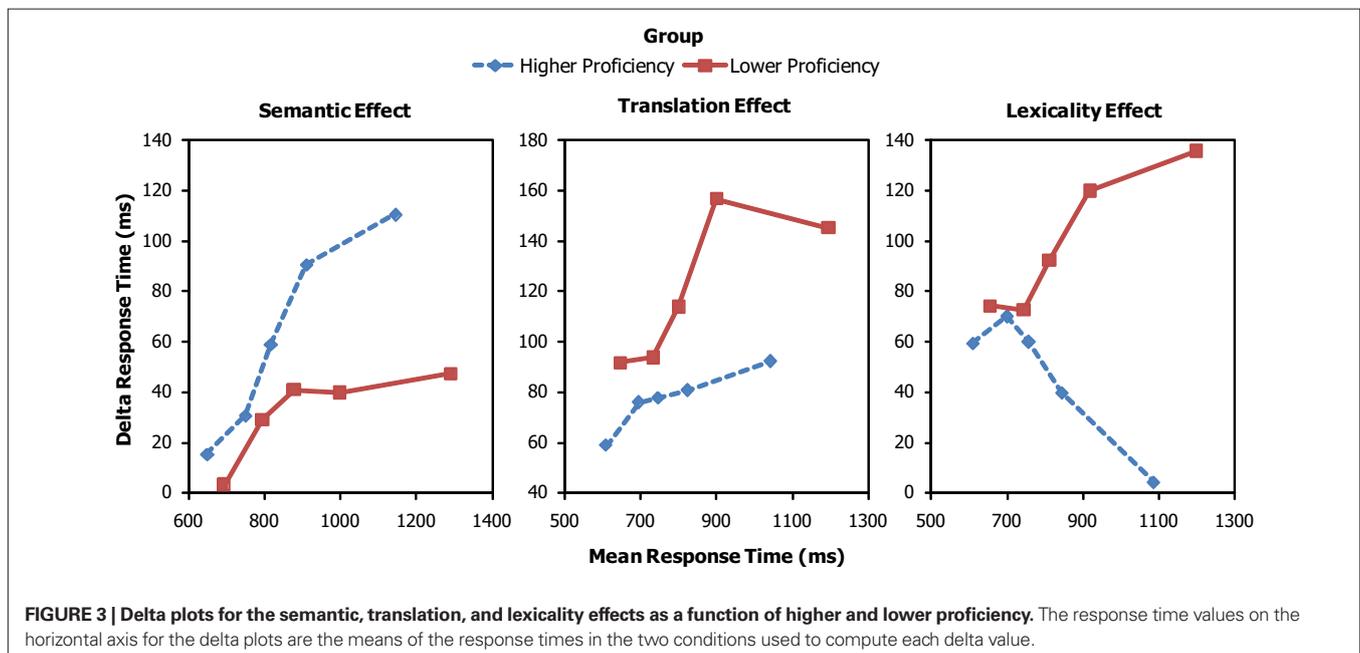
Figure 3 gives the delta plots for the semantic, translation, and lexicality effects as a function of relative proficiency (higher, lower) based on the self-ratings of the level of skill in L2. The figure shows that the magnitude of the translation and lexicality effects increases with increasing RT for the low-proficiency individuals but not or much less for the high-proficiency individuals. Moreover, the magnitude of the semantic effect increases with increasing RT for the high-proficiency individuals, but much less for the low-proficiency ones.

The difference in delta plots between the proficiency groups was confirmed by statistical analyses of the slopes. For the semantic effect, the difference in slope between groups was marginally significant for the segment q2–3, $F(1,14) = 4.28$, $p = 0.058$, MSE = 0.068, significant for segment q3–4, $F(1,14) = 8.65$, $p = 0.011$, MSE = 0.062, but not for the other segments q1–2, $F(1,14) = 0.06$, $p = 0.80$, MSE = 0.097; q4–5, $F(1,14) = 0.02$, $p = 0.88$, MSE = 0.142. For the translation effect, the difference in slope between groups was

significant for segment q2–3, $F(1,14) = 4.86$, $p = 0.045$, MSE = 0.052, and segment q3–4, $F(1,14) = 6.26$, $p = 0.025$, MSE = 0.088, but not for the other segments, q1–2, $F(1,14) = 1.42$, $p = 0.25$, MSE = 0.057; q4–5, $F(1,14) = 0.83$, $p = 0.38$, MSE = 0.105. Finally, for the lexicality effect, the slopes differed as a function of group for the segment q2–3, $F(1,14) = 12.20$, $p = 0.004$, MSE = 0.059, and the segment q3–4, $F(1,14) = 13.87$, $p = 0.002$, MSE = 0.065, but not for the other segments, q1–2, $F(1,14) = 0.37$, $p = 0.55$, MSE = 0.110; q4–5, $F(1,14) = 1.10$, $p = 0.31$, MSE = 0.094.

The inhibition effects appear ubiquitous. Yet, the observed patterns raise the question of the specificity of effects. If the differential delta-plot effects indicate effects of inhibition, then differential delta-plot effects should *not* be found when delta plots are computed for experimental factor effects that should not involve inhibition. In particular, differential delta plots should not be obtained for median splits based on factors that do not load on inhibition, such as mean RT or age (given that all participants were young). To assess the specificity of the delta-plot effects, we computed delta plots for groups based on median splits of mean RT and age.

The groups created by median splits based on naming speed differed in mean RT (766 vs. 897 ms), $F(1,14) = 17.57$, $p = 0.001$. However, the analyses of the slopes yielded no difference between the mean RT groups for the semantic effect, all ps > 0.21, the translation effect, all ps > 0.17, and the lexicality effect, all ps > 0.17. Moreover, the groups created by median splits based on age differed in mean age (21 vs. 25 years), $F(1,14) = 19.93$, $p = 0.001$. However, the analyses of the slopes yielded no difference between the age groups for the semantic effect, all ps > 0.25, the translation effect, all ps > 0.18, and the lexicality effect, all ps > 0.14. These analyses revealed that although the groups differed statistically in mean RTs and age, the delta plots for the semantic, lexicality, and translation effects did not differ. This suggests that the delta-plot effects are specific to experimental factors differing in inhibition, such as effect size and proficiency.

FIGURE 3 | Delta plots for the semantic, translation, and lexicality effects as a function of higher and lower proficiency. The response time values on the horizontal axis for the delta plots are the means of the response times in the two conditions used to compute each delta value.

DISCUSSION

Prior evidence from delta-plot analyses suggests that differential effects of attentional inhibition are largest in the tail of an RT distribution in non-linguistic and monolingual performance domains (e.g., Ridderinkhof, 2002; Ridderinkhof et al., 2004, 2005; Bub et al., 2006; Sharma et al., 2010). The reported experiment examined whether this also holds for bilingual performance by conducting delta-plot analyses of picture naming RTs. Dutch–English bilingual speakers named pictures in English while trying to ignore superimposed Dutch distractor words or neutral series of Xs. Picture name and distractor word were semantically related, unrelated, or translations. The mean RTs revealed semantic, translation, and lexicality effects. The delta pots revealed that the magnitude of the distractor effects flattened with increasing RT, more so when the mean distractor effect was smaller (presumed to reflect strong inhibition) as opposed to larger (presumed to reflect weaker inhibition). Moreover, the delta plots leveled off with increasing RT more for high- than low-proficiency individuals in the unrelated than the control and translation conditions, whereas the reverse held for the semantically related condition.

DYNAMICS OF INHIBITION

The present observation that the magnitude of the distractor effects leveled off with increasing RT, more so when the mean effect size was smaller than when it was larger, corresponds to what Ridderinkhof (2002) observed using the non-linguistic Simon task. In his study, the delta plots leveled off more in participants with relatively small Simon effects than in participants with relatively large Simon effects. According to Ridderinkhof et al. (2004, 2005), this suggests that inhibition builds up slowly during a trial in non-linguistic domains. The present evidence on the dynamics of inhibition also agrees with the evidence of Verhoef et al. (2009) that inhibition takes time to develop in bilingual language performance. In that study, it was observed that the magnitude of a right-lateralized N2 effect was modulated by the size of the preparation interval, but not by language or language switching. The effect of preparation interval suggests that inhibition builds up slowly during a trial, in line with the observations of Ridderinkhof and colleagues. However, the absence of an effect of language and switching on the N2 suggests that inhibition is an optional rather than compulsory mechanism.

The evidence for inhibition in the present experiment raises the question what exactly is inhibited in the bilingual picture–word interference task. Inhibition may concern responding to the distractor or in the irrelevant language. If language rather than distractor were inhibited, differences among distractor word conditions (i.e., semantically related, translation, unrelated) are not to be expected, unless the distractor words activate their language information to different degrees depending on the type of distractor. Thus, it is more likely that the inhibition concerns responding to the distractor word (cf. Bub et al., 2006; Sharma et al., 2010). This inhibition of distractor word processes may not be specific to bilingual performance (i.e., L2 naming), but presumably reflects the fact that distractor words have to be ignored in the picture–word interference paradigm, regardless of their language.

Another issue raised by the present findings is at what level of processing the interference effects occur. Whereas some researchers argue that the effects arise during picture name selection (Roelofs, 2003), others maintain that the effects arise during perceptual/conceptual encoding of the picture (Dell'Acqua et al., 2007; Van Maanen et al., 2009). In a recent study, we examined the time course of semantic, translation, and lexicality effects in overt picture naming by means of ERP recordings (Roelofs et al., submitted). The materials were the same as in the present study.

Predictions for the onset of the distractor effects in the ERPs were based on estimates of the timing of processing stages underlying word production provided by an influential meta-analysis by Indefrey and Levelt (2004). According to these estimations, based on an average naming latency of 600 ms, the stage of perceptual and conceptual encoding is completed around 150–200 ms after picture onset, after which lexical selection starts. As in the present study, the mean naming latencies in our bilingual EEG study were longer than 600 ms, namely around 840 ms in the control condition. Taking 840 ms as the mean naming latency, and scaling the estimates to this mean, gives us 280 ms as the end of the time window of perceptual and conceptual encoding and as the point in time at which the operation of word selection is initiated. If the effects emerge during perceptual and conceptual encoding, they should emerge in the EEG in a time window that extends at most to 280 ms post picture onset, whereas if the effects arise during lexical selection, they should emerge after 280 ms post picture onset. The ERP data revealed that the semantic, translation, and lexicality effects started to emerge 300 ms after picture onset, suggesting that they occurred during the selection of the picture name.

In the present experiment, differences in delta plots as a function of effect size and proficiency tended to occur in segments q2–3 and q3–4, but not in the other segments. The absence of a difference in the first segment q1–2 supports the assumption that inhibition takes some time to develop. Apparently, after some initial build up of inhibition over time, differences in delta-plot slopes as a function of effect size and proficiency started to arise, as reflected in the middle delta-plot segments q2–3 and q3–4. However, differences in slopes tended to be absent in the last segment q4–5. This suggests that participants did not maintain a high level of inhibition throughout a trial. One possibility is that maintaining inhibition requires effort, which participants were willing to invest for some period during a trial but not throughout a whole trial. Alternatively, it may have been impossible for the participants to keep up a high level of inhibition for a longer period (cf. De Jong et al., 1999). Either way, if inhibition diminishes toward the end of a trial, differences in delta-plot slopes as a function of effect size and proficiency will also disappear, as observed in the present experiment.

ROLE OF PROFICIENCY

It has been argued that bilingual individuals have enhanced inhibition capabilities compared to monolingual individuals (Bialystok et al., 2004; Costa et al., 2008), because of their more frequent use of inhibition. According to Green (1998), lexical competitors in the other language are more activated for high- than low-proficiency bilinguals and so require a greater degree of inhibition. As a result, the inhibition ability should improve with increased proficiency. Consequently, delta plots in bilingual picture–word interference should level off more for high- than low-proficiency individuals.

In the present study, the magnitude of the lexicality and translation effects increased with increasing RT for the low-proficiency individuals but not for the high-proficiency ones, in agreement with

the hypothesis that inhibition enhances with proficiency. Moreover, the magnitude of the semantic effect increased with increasing RT for the high-proficiency individuals, but much less for the low-proficiency ones. The present data suggest that the high-proficiency participants applied more inhibition than the low-proficiency participants on unrelated trials relative to the other trial types. This leads to a stronger decrease of the distractor effect with increasing RT for the lexicality and translation effects, and also leads to the opposite influence for the semantic effect.

As **Table 2** shows, mean RTs were longer on unrelated than translation and control trials, but shorter on unrelated than semantically related trials. Inhibition of unrelated distractors will reduce RTs on unrelated trials and consequently will reduce the difference in RT between the unrelated and translation trials (i.e., the translation effect), reduce the difference in RT between the unrelated and control trials (i.e., the lexicality effect), but *increase* the difference in RT between the unrelated and semantically related trials (i.e., the semantic effect). If the unrelated distractor words are more strongly inhibited by the high- than the low-proficiency participants, the difference in RT between unrelated and translation trials (i.e., the translation effect) will decrease more with increasing RT for the high- than the low-proficiency participants, as empirically observed. Moreover, the difference in RT between unrelated and control trials (i.e., the lexicality effect) will decrease more with increasing RT for the high- than the low-proficiency participants, also as observed. However, if the unrelated distractor words are more strongly inhibited by the high- than the low-proficiency participants, the difference in RT between unrelated and semantically related trials (i.e., the semantic effect) will *increase* more with increasing RT for the high- than the low-proficiency participants, as empirically observed.

The observation that the delta plots level off differentially depending on the type of effect (i.e., semantic vs. translation and lexicality) suggests that individuals engage inhibition differently depending on the prevailing circumstances. This implies that participants monitored progress on task performance on each trial and allocated attentional inhibition depending on their assessment of the distractor type (cf. Kahneman, 1973; Roelofs, 2007, 2008).

Costa and Santesteban (2004) argued that the use of inhibition depends on language proficiency. In their view, while non-balanced bilinguals use inhibition to achieve language selectivity, balanced bilinguals do not use inhibition to accomplish this. As indicated earlier, this view of Costa and Santesteban (2004) has been challenged by Verhoef et al. (2009), who observed that whether inhibition is engaged may depend on the amount of preparation time in language switching in a single population of non-balanced Dutch–English bilinguals. The present results suggest that the use of inhibition by non-balanced bilinguals may not only depend on the preparation time, but also on the type of distractor. This corroborates the view that inhibition is an optional mechanism that is engaged depending on the prevailing circumstances.

CONCLUSION

The present study provides evidence that inhibition is a mechanism of attentional control in bilingual language performance. In a bilingual picture–word interference experiment, the magnitude of semantic, translation, and lexicality effects decreased with increasing RT, more so when the mean distractor effect is smaller (presumed to reflect strong inhibition) than when it is larger (presumed to reflect weaker inhibition). This suggests that the influence of inhibition is largest toward the RT distribution tail, corresponding to what is observed in non-linguistic domains. Moreover, the delta plots suggested that more inhibition was applied by high- than low-proficiency individuals on unrelated trials than on the other trial types. These results support the view that inhibition is a domain-general mechanism that may be optionally engaged depending on the prevailing circumstances.

REFERENCES

Abutalebi, J., and Green, D. (2007). Bilingual language production: the neurocognition of language representation and control. *J. Neurolinguistics* 20, 242–275.

Aron, A. R. (2007). The neural basis of inhibition in cognitive control. *Neuroscientist* 13, 214–228.

Aron, A. R., Robbins, T. W., and Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci. (Regul. Ed.)* 8, 170–177.

Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303.

Bub, D. N., Masson, M. E. J., and Lalonde, C. E. (2006). Cognitive control in children: Stroop interference and suppression of word reading. *Psychol. Sci.* 17, 351–357.

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D.,

Nieuwenhuis, S. T., La Heij, W., and Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 302–312.

Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86.

Costa, A., Miozzo, M., and Caramazza, A. (1999). Lexical selection in bilinguals: do words in the bilingual's two lexicons compete for selection? *J. Mem. Lang.* 41, 365–397.

Costa, A., and Santesteban, M. (2004). Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. *J. Mem. Lang.* 50, 491–511.

De Jong, R., Berendsen, E., and Cools, R. (1999). Goal neglect and inhibitory limitations: dissociable causes

of interference effects in conflict situations. *Acta Psychol. (Amst)* 101, 379–394.

De Jong, R., Liang, C.-C., and Lauber, E. (1994). Conditional and unconditional automaticity: a dual-process model of effects of spatial stimulus-response correspondence. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 731–750.

Dell'Acqua, R., Job, R., Peressotti, F., and Pascali, A. (2007). The picture-word interference effect is not a Stroop effect. *Psychon. Bull. Rev.* 14, 717–722.

Duncan, J. (2010). *How Intelligence Happens.* Yale: Yale University Press.

Finkbeiner, M., Almeida, J., Janssen, N., and Caramazza, A. (2006). Lexical selection in bilingual speech production does not involve language suppression. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1075–1089.

Forstmann, B. U., Jahfari, S., Scholte, H. S., Wolfensteller, U., van den Wildenberg, W. P. M., and Ridderinkhof, K. R. (2008a). Function and structure of

the right inferior frontal cortex predict individual differences in response inhibition: a model-based approach. *J. Neurosci.* 28, 9790–9796.

Forstmann, B. U., van den Wildenberg, W. P. M., and Ridderinkhof, K. R. (2008b). Neural mechanisms, temporal dynamics, and individual differences in interference control. *J. Cogn. Neurosci.* 20, 1854–1865.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81.

Hermans, D., Bongaerts, T., De Bot, K., and Schreuder, R. (1998). Producing words in a foreign language: can speakers prevent interference from their first language? *Biling. Lang. Cogn.* 1, 213–229.

Indefrey, P., and Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs

(transformation or not) and towards logit mixed models. *J. Mem. Lang.* 59, 434–446.

Kahneman, D. (1973). *Attention and Effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kroll, J. F., Bobb, S. C., Misra, M., and Guo, T. (2008). Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychol. (Amst.)* 128, 416–430.

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization.* New York: Oxford University Press.

Proctor, R. W., Miles, J. D., and Baroni, G. (2011). Reaction time distribution analysis of spatial correspondence effects. *Psychon. Bull. Rev.* 18, 242–266.

Ridderinkhof, K. R. (2002). "Activation and suppression in conflict tasks: empirical clarification through distributional analyses," in *Attention and Performance XIX: Common Mechanisms in Perception and Action,* eds W. Prinz and B. Hommel (Oxford: Oxford University Press), 494–519.

Ridderinkhof, K. R., Scheres, A., Oosterlaan, J., and Sergeant, J. (2005). Delta plots in the study of individual differences: new tools reveal response inhibition deficits in AD/HD that are eliminated by methylphenidate treatment. *J. Abnorm. Psychol.* 114, 197–215.

Ridderinkhof, R. K., Van den Wildenberg, W. P. M., Wijnen, J., and Burle, B. (2004). "Response inhibition in conflict tasks is revealed in delta plots," in *Cognitive Neuroscience of Attention,* ed. M. I. Posner (New York: The Guilford Press), 369–377.

Roelofs, A. (2003). Goal-referenced selection of verbal action: modeling attentional control in the Stroop task. *Psych. Rev.* 110, 88–125.

Roelofs, A. (2007). Attention and gaze control in picture naming, word reading, and word categorizing. *J. Mem. Lang.* 57, 232–251.

Roelofs, A. (2008). Dynamics of the attentional control of word retrieval: analyses of response time distributions. *J. Exp. Psychol. Gen.* 137, 303–323.

Roelofs, A., and Verhoef, K. (2006). Modeling the control of phonological encoding in bilingual speakers. *Biling. Lang. Cogn.* 9, 167–176.

Sharma, D., Booth, B., Brown, R., and Huguet, P. (2010). Exploring the temporal dynamics of social facilitation in the Stroop task. *Psychon. Bull. Rev.* 17, 52–58.

Van den Wildenberg, W. P. M., Wylie, S. A., Forstmann, B. U., Burle, B., Hasbroucq, T., and Ridderinkhof, R. K. (2011). To head or to heed? Beyond the surface of selective action inhibition: a review. *Front. Hum. Neurosci.* 4:222. doi: 10.3389/fnhum.2010.00222

Van Maanen, L., van Rijn, H., and Borst, J. P. (2009). Stroop and picture-word interference are two sides of the same coin. *Psychon. Bull. Rev.* 16, 987–999.

Verhoef, K., Roelofs, A., and Chwilla, D. (2009). Role of inhibition in language switching: evidence from event-related brain potentials in overt picture naming. *Cognition* 110, 84–99.

Wundt, W. (1904). *Principles of Physiological Psychology.* London: Swan Sonnenschein.