Bilingualism: Language and Cognition 16 (2), 2013, 420–441 © Cambridge University Press 2012. The online version of this article is published within an Open Access environment subject to the conditions of the Creative Commons Attribution-NonCommercial-ShareAlike licence <http://creativecommons.org/licenses/by-nc-sa/2.5/>. The written permission of Cambridge University Press must be obtained for commercial re-use. doi:10.1017/S1366728912000405

The timing and magnitude of Stroop interference and facilitation in monolinguals and bilinguals*

EMILY L. CODERRE
School of Psychology, University of Nottingham
WALTER J. B. VAN HEUVEN
School of Psychology, University of Nottingham
KATHY CONKLIN
School of English, University of Nottingham

(Received: February 28, 2011; final revision received: June 15, 2012; accepted: June 26, 2012; first published online 20 November 2012)

Executive control abilities and lexical access speed in Stroop performance were investigated in English monolinguals and two groups of bilinguals (English–Chinese and Chinese–English) in their first (L1) and second (L2) languages. Predictions were based on a bilingual cognitive advantage hypothesis, implicating cognitive control ability as the critical factor determining Stroop interference; and two bilingual lexical disadvantage hypotheses, focusing on lexical access speed. Importantly, each hypothesis predicts different response patterns in a Stroop task manipulating stimulus onset asynchrony (SOA). There was evidence for a bilingual cognitive advantage, although this effect was sensitive to a number of variables including proficiency, language immersion, and script. In lexical access speed, no differences occurred between monolinguals and bilinguals in their native languages, but there was evidence for a delay in L2 processing speed relative to the L1. Overall, the data highlight the multitude of factors affecting executive control and lexical access speed in bilinguals.

Keywords: Stroop, bilingualism, lexical access, executive control, interference, facilitation

Introduction

Since its introduction in the mid-1930s, the Stroop task has become a canonical paradigm of executive control and cognitive conflict (see MacLeod, 1991, for a review). The traditional task consists of a color word printed in colored ink that either matches the color word (e.g. "red" printed in red ink: a congruent condition), or mismatches the color word (e.g. "red" printed in blue: an incongruent condition). When participants are asked to ignore the written word and name the ink color, reaction times (RTs) are reliably slower in incongruent conditions than in congruent or neutral conditions. Because reading is such a highly practiced process, an incongruent condition creates conflict between the word and the color, requiring increased cognitive control to resolve the conflict, and resulting in longer naming latencies for incongruent conditions.[1]

The Stroop task indexes cognitive control abilities and language skills, both of which are affected by bilingualism. We first provide an overview of three hypotheses regarding bilingual cognitive control and lexical access speed, and discuss how these factors might influence Stroop performance in bilinguals. We then introduce a stimulus onset asynchrony (SOA) manipulation as a method of separating these factors.

The bilingual cognitive advantage hypothesis

One phenomenon in the bilingualism literature is that bilinguals tend to outperform monolinguals on the Stroop task, showing smaller interference effects than their monolingual counterparts (e.g. Bialystok, 2009; Bialystok, Craik & Luk, 2008). This BILINGUAL COGNITIVE ADVANTAGE also extends more generally to non-linguistic executive control such as the Simon task (Bialystok et al., 2008; Bialystok, Craik, Klein & Viswanathan, 2004), anti-saccade tasks (e.g. Bialystok & Viswanathan, 2009), the dimensional card sort task (Bialystok & Martin, 2004), and the attentional network and flanker tasks (Costa, Hernández, Costa-Faidella & Sebastián-Gallés, 2009; Costa, Hernández & Sebastián-Gallés. 2008). The bilingual cognitive advantage hypothesis stems from the theory of non-selective access to an integrated bilingual lexicon. This proposes that both languages of bilinguals are activated in parallel, even in completely monolingual contexts, such that bilinguals cannot completely "turn off" one language (e.g., Martin, Dering, Thomas & Thierry,

* The authors would like to thank Andrew Kelly and Lifang Wang for their help with data collection, and the anonymous reviewers for their helpful comments on earlier versions of this manuscript.

[1] Stroop INTERFERENCE EFFECTS are referred to here as the difference between incongruent and control conditions, and Stroop FACILITATION EFFECTS as the difference between control and congruent conditions. STROOP EFFECTS (incongruent vs. congruent conditions) are a combination of interference and facilitation effects and therefore less appropriate for measuring Stroop interference.

Address for correspondence:
Emily L. Coderre, School of Psychology, The University of Nottingham, University Park, Nottingham NG7 2RD, UK
lpxec1@nottingham.ac.uk

2009; Midgley, Holcomb, van Heuven & Grainger, 2008; van Heuven, Dijkstra & Grainger, 1998; van Heuven, Schriefers, Dijkstra & Hagoort, 2008). Non-selective access necessitates a highly effective mechanism of control over the non-relevant language to avoid cross-language comprehension or production errors. This constant exertion of control over languages leads to an overall enhancement of executive control processes (Bialystok, 2001), resulting in better performance (i.e. less interference) for bilinguals than monolinguals on the Stroop and other conflict tasks. Importantly, as the bilingual advantage extends to non-language control tasks, the advantage is not purely language-based, but benefits executive processing more generally.

The bilingual L1 lexical disadvantage hypothesis

In contrast to the advantage exhibited on control tasks, a BILINGUAL DISADVANTAGE is reported in lexical access tasks such as picture naming (Gollan, Montoya, Fennema-Notestine & Morris, 2005; Kaushanskaya & Marian, 2007), lexical decision (Ransdell & Fischler, 1987) and lexical retrieval or verbal fluency (Bialystok et al., 2008), where bilinguals' performance is poorer than monolinguals', even when bilinguals perform in their native language. This disadvantage can also be attributed to non-selective access to an integrated bilingual lexicon (Bialystok, 2009; Bialystok et al. 2008): the presence of two (or more) lexicons activated in parallel results in competition and interactions between potential lexical candidates, creating delays in lexical access and a performance disadvantage compared to monolinguals (e.g. van Heuven et al., 1998).

One alternative theory of the bilingual disadvantage is the REDUCED FREQUENCY HYPOTHESIS (Pyers, Gollan & Emmorey, 2009), also referred to as the WEAKER LINKS HYPOTHESIS (Gollan et al., 2005), which attributes the disadvantage in language production to the frequency of language use. This theory suggests that compared to monolinguals, bilinguals use both languages less often, including their first language (L1). This reduced frequency of use leads to weaker ties between words and concepts (Gollan & Acenas, 2004; Gollan, Montoya, Cera & Sandoval, 2008; Gollan et al. 2005; see also the Revised Hierarchical Model (Kroll & Stewart, 1994) and models of concept mediation and word association (Potter, So, van Eckhardt & Feldman, 1984)). Weaker ties lead to slower lexical access and thus longer RTs in lexical access tasks. Importantly, as bilinguals also use their L1 less often than monolinguals, this hypothesis predicts slower lexical processing for bilinguals even in a native language. This theory, referred to here as the BILINGUAL L1 LEXICAL DISADVANTAGE HYPOTHESIS, thus explains the bilingual disadvantage in the L1 in terms of frequency of language use.

The bilingual L2 lexical disadvantage hypothesis

By the same reasoning, the reduced frequency hypothesis further predicts delayed lexical access in the less-dominant second language (L2) compared to the more-dominant L1 due to the reduced frequency of L2 use (Gollan et al., 2008). Similarly, the TEMPORAL DELAY ASSUMPTION of the BIA+ model of bilingual word recognition (Dijkstra & van Heuven, 2002; van Heuven & Dijkstra, 2010) states that in unbalanced bilinguals, the activation speed of orthographic, phonological and semantic codes is delayed in the L2 compared to the L1 since subjective word frequencies are lower in the L2. This L2 delay, referred to here as the BILINGUAL L2 LEXICAL DISADVANTAGE HYPOTHESIS, is supported by extensive empirical evidence, both in production and comprehension, and across a range of bilingual proficiencies and ages of acquisition (e.g. Ardal, Donald, Meuter, Muldrew & Luce, 1990; Coderre, Conklin & van Heuven, 2012; Moreno & Kutas, 2005; Newman, Tremblay, Nichols, Neville & Ullman, 2012; Proverbio, Adorni & Zani, 2009; see Runnqvist, Strijkers, Sadat & Costa, 2011, and van Heuven & Dijkstra, 2010, for reviews).

Stroop performance in monolinguals and bilinguals

When considering Stroop performance in monolinguals and bilinguals, an interesting question arises regarding the mechanisms governing interference effects. Do bilinguals show an advantage because they have superior executive control processes (as suggested by the bilingual cognitive advantage hypothesis), and are thus better able to resolve Stroop conflict? Or is the slower lexical access (as proposed by the bilingual lexical disadvantage hypotheses) generating less interference? Both possibilities would predict improved performance for bilinguals on a Stroop task, but previous research has not yet investigated the contributions of these two factors in determining Stroop performance.

The interaction of executive control and bilingual lexical access speed

Because each theory places the locus of Stroop interference on different factors, they are not mutually exclusive and may interact. For example, using a proactive interference (PI) task, a complex verbal task which requires executive control, Bialystok and Feng (2009), found no bilingual advantage unless controlling for vocabulary knowledge. Similarly, Carlson and Meltzoff (2008) reported a bilingual advantage on a variety of executive function tasks in young children, but only when verbal ability, age, and socio-economical status (SES) were controlled for. These studies demonstrate that both vocabulary knowledge and executive control abilities have implications for bilingual performance on linguistic cognitive control tasks. Since bilinguals are disadvantaged in lexical access speed but advantaged in cognitive control, a better understanding of how these two factors interact on a language-based executive function task like the Stroop task is needed to better interpret the bilingual advantage. The goal of the current experiments was to separate the influences of executive control abilities and lexical access speed in the Stroop task by manipulating the stimulus onset asynchrony of the word and color information.

SOA manipulation in the Stroop task

One of the many variations of the Stroop task is stimulus onset asynchrony (SOA) manipulation (e.g. Dyer, 1973; Glaser & Glaser, 1982; Roelofs, 2010), which presents the word and color stimuli at different times. The amount of interference at each SOA provides insight into the speed of processing of each dimension. Negative SOAs present the irrelevant stimulus (e.g. the word, in a color-naming task) before the relevant stimulus (the color), whereas positive SOAs present the irrelevant stimulus after the relevant stimulus. In a series of seminal experiments, Glaser and Glaser (1982) used nine SOAs (from –400 ms to +400 ms in 100 ms steps) to investigate the timing of interference between color and word stimuli. The most interference (incongruent RT minus control RT) occurred at the 0 ms SOA, but significant interference occurred at negative SOAs even out to –400 ms (see Figure 1 below for examples of the interference and facilitation effects from their original data). In positive SOAs, interference was diminished but still significant at +200 ms, but all effects were gone by +400 ms, as the word appeared too late to influence color-naming. SOA manipulation has been explored in monolingual Stroop tasks (Appelbaum, Boehler, Won, Davis & Woldorff, 2012; Appelbaum, Meyerhoff & Woldorff, 2009; Dyer, 1973, 1974; Glaser & Glaser, 1982, 1989; Roelofs, 2003) and applied to other tasks such as picture-word interference (e.g. Glaser & Düngelhoff, 1984; Schriefers, Meyer & Levelt, 1990) and word translation tasks (e.g. Miller & Kroll, 2002). SOA manipulation has not yet been used in bilinguals, yet it is a useful means of investigating Stroop effects in bilingualism.

Specifically, the Stroop SOA paradigm provides a unique method for investigating the impact of a cognitive advantage and/or lexical disadvantage on bilingual Stroop performance. The Stroop task indexes both lexical and executive control abilities within one task, enabling the investigation of these factors in bilingualism. Decades of research have established the Stroop task as a reliable measure of reading ability, as the strength of interference indicates how strongly reading interferes with color-naming. In bilinguals, the magnitude of interference in each language is therefore indicative of the strength of language connections. The use of high-frequency color words also allows for a measure of lexical access speed in bilinguals without a frequency confound. SOA manipulation also provides a wider temporal spectrum than typical Stroop tasks using a static 0 ms SOA. Importantly, the identification of interference effects at negative SOAs may be indicative of lexical processing delays (see predictions below). Crucially, each of the theories previously discussed predicts a different pattern of RTs in SOA manipulation.

The current studies

In three experiments, a Stroop SOA task using five SOAs (±400 ms, ±200 ms, and 0 ms) was administered to monolinguals and to two groups of bilinguals: English–Chinese (L1 English, L2 Mandarin), and Chinese–English (L1 Mandarin, L2 English) in their L1 and L2. Chinese and English were chosen to avoid the issue of cognate effects in color words translations. In most European languages the color words are highly similar, either in orthography, phonology, or both. This overlap could cause language-related facilitation and/or interference effects, which would be confounded with conflict-related effects. The factors of cognitive control and lexical access speed are examined in the context of the above-mentioned theories in order to more accurately explain the differences in Stroop performance: (i) between monolinguals and bilinguals in their native languages (monolinguals vs. bilinguals' L1); and (ii) between bilinguals' two languages (L1 vs. L2).

A manual Stroop task, in which participants indicate the color using a button-box, was employed in all groups to avoid the issue of overt word production processes. A vocal-response Stroop task, which requires participants to name the color aloud, involves not just word recognition but also verbal response processes, which are both influenced by L2 proficiency, age of L2 acquisition, word frequency, and semantic context (e.g. Gollan, Slattery, Goldenberg, van Assche, Duyck & Rayner, 2011; Thornburgh & Ryalls, 1998). To avoid these complexities of L2 production and eliminate the influence of overt word production processes, a manual task is used in the current studies, allowing for the investigation of how word recognition processes, as modulated by language proficiency, determine Stroop performance.

Below we outline the general predictions that the bilingual cognitive advantage hypothesis and the bilingual lexical disadvantage hypotheses would make for monolingual and bilingual performance on the Stroop SOA task. More specific predictions will be made after the presentation of Experiment 1, which establishes a "baseline" effects in monolinguals.

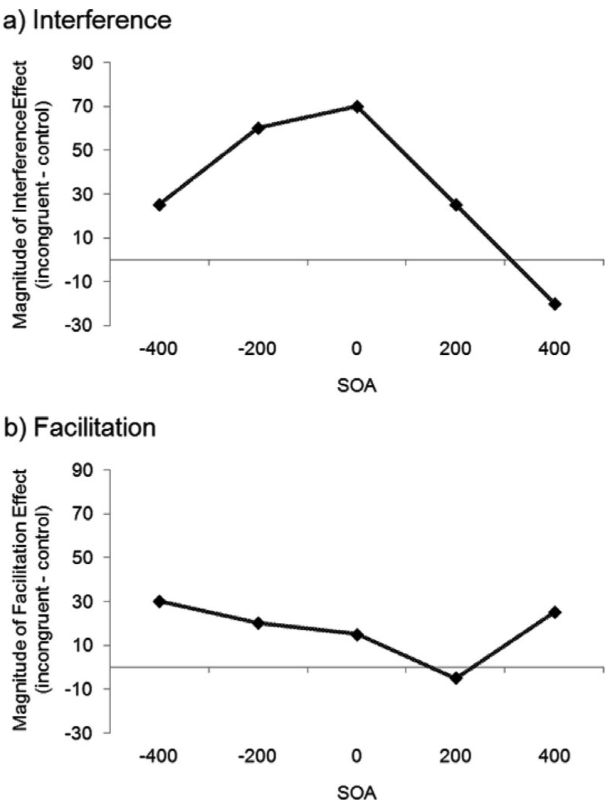

Figure 1. (a) Interference and (b) facilitation effects taken from Glaser and Glaser's (1982) original color-naming Stroop study (Experiment 1), for the five SOAs used in the current study.

Monolinguals versus bilinguals' L1

The bilingual L1 lexical disadvantage hypothesis predicts that relative to monolinguals, the peak interference should be shifted to negative SOAs for a bilingual L1. "Bilinguals are essentially less proficient or fluent than monolinguals" (Gollan et al., 2005, p. 1231) due to the reduced frequency of language use, which leads to delayed lexical access. If this is the case, a pre-exposure of the word in negative SOAs should allow this delayed lexical access a head-start. This head-start would create a peak interference effect at an earlier time window (e.g. –200 ms SOA vs. 0 ms SOA). Bilinguals' L1 should therefore show a negative shift in Stroop interference effects relative to monolinguals such that the peak interference effects occur at earlier negative SOAs (see Figure 1a for an example of monolingual interference patterns; a negative shift in interference effects would shift this line to the left). By the same rationale, a negative shift may also be apparent in the Stroop facilitation effects (Figure 1b).

According to the bilingual cognitive advantage hypothesis, bilinguals have more efficient cognitive control processes than monolinguals, and therefore should show smaller Stroop interference effects than monolinguals. It may also be the case that proficiency affects executive control abilities within bilingual groups, such that high-proficiency bilinguals show better cognitive control, and therefore less interference, than low-proficiency bilinguals. Importantly, enhanced executive control would predict no differences in the overall pattern of SOA interference effects. That is, bilinguals and monolinguals should show peak interference effects at the same SOAs, but bilinguals should show a smaller magnitude of interference (i.e. the pattern of Figure 1a would be shifted down).

The bilingual advantage has been shown to be sensitive to the degree of conflict present (Costa et al., 2009), so in some SOAs (particularly positive SOAs in which little interference is elicited in monolinguals) bilinguals will likely not show an advantage because of a floor effect in interference. Only in SOAs which require the most cognitive control, such as the 0 ms SOA, is the bilingual advantage expected to appear, as these high-conflict conditions should emphasize bilinguals' superior cognitive control abilities (e.g. Costa et al., 2009). The bilingual cognitive advantage hypothesis therefore predicts a decrease in interference for bilinguals as compared to monolinguals, especially – or perhaps only – at SOAs requiring the most cognitive control. Importantly, however, interference should not be modulated by SOA between the groups (i.e. all groups should show the peak interference effects at the same SOAs), as this theory does not take into account speed of lexical processing.

As well as smaller interference effects, bilinguals may also experience smaller facilitation effects. Previous work has found that populations with impaired cognitive control, such as adults with Alzheimer's disease (e.g. Weaver Cargin, Maruff, Collie, Shafiq-Antonacci & Masters, 2007) and children (e.g. Zelazo, Craik & Booth, 2004), show not only increased Stroop interference but also increased facilitation effects compared to normal adults (Spieler, Balota & Faust, 1996; Wright & Wanley, 2003). Ignoring the word in a color-naming Stroop task requires cognitive control, so individuals with decreased cognitive control abilities may have more "inadvertent word reading" errors (MacLeod & MacDonald, 2000; Wright & Wanley, 2003), leading to increased facilitation. Therefore if bilinguals have increased cognitive control abilities, they may show a decrease in facilitation compared to monolinguals (i.e. the pattern in Figure 1b would be shifted down), in the same way that they show decreased interference effects.

Bilinguals' L1 versus L2

Previous research using bilingual Stroop tasks has reported that the L1 tends to elicit more Stroop interference than the L2, although this depends on the proficiency of the participants and the similarity of the languages (Brauer, 1998; Chen & Ho, 1986; Fang, Tzeng & Alva, 1981; Tzelgov, Henik & Leiser, 1990). According

Table 1. *Subjective proficiency scores (scale: 1 = very poor, 10 = very fluent) and participant demographics of Experiments 1–3. Numbers in parentheses indicate standard deviation.*

						Subjective L2 proficiency scores				
Experiment	Participants	n	Age	L1	First contact L2	Speaking	Listening	Reading	Writing	Overall
1	English monolinguals	23	23.0 (4.1)	English	N/A	N/A	N/A	N/A	N/A	N/A
2	English–Chinese bilinguals	15	21.8 (2.4)	English	10.3 (7.3)	7.0 (1.3)	7.0 (1.6)	6.7 (1.6)	5.4 (2.5)	6.5 (1.3)
3	Chinese–English bilinguals	22	21.0 (1.6)	Mandarin Chinese	11.0 (2.7)	6.4 (1.3)	6.6 (1.4)	7.1 (1.4)	6.3 (1.5)	6.6 (1.2)

to the bilingual L2 lexical disadvantage hypothesis, word information in the L2 is less easily accessed (i.e. delayed) in unbalanced bilinguals. If the L2 experiences delayed lexical access, then a pre-exposure of the word in negative SOAs should create a negative shift in the peak interference effects for the L2 compared to the L1, by the same mechanism described earlier in the predictions for the L1 lexical disadvantage hypothesis. If decreasing proficiency results in an increased amount of time needed for lexical access, then this "negative shift" in interference effects should be even more pronounced: i.e. the negative shift should increase with decreasing proficiency. The bilingual L2 lexical disadvantage hypothesis thus predicts a language-dependent negative shift in interference effects, and possibly also facilitation effects, in the L2 compared to the L1 due to slower lexical access resulting from reduced frequency of language use.

The bilingual cognitive advantage hypothesis claims that bilinguals have better cognitive control than monolinguals, but generally does not mention differences within a bilingual's two languages. As it is based solely on the efficiency of executive control and not on speed of lexical processing, this hypothesis should predict no differences between the L1 and L2 in SOA modulation, because executive control processes likely occur at a fixed rate within individuals. It may be, however, that because the L1 is the stronger language it is more difficult to ignore and therefore requires more control, which would predict slightly larger Stroop effects in the L1 than the L2 within individuals. Importantly, though, these effects will not be modulated by SOA, because similar amounts of cognitive control will be required at each SOA.

As mentioned, these hypotheses point to different mechanisms being the critical factor in Stroop performance (cognitive control in the bilingual cognitive advantage hypothesis, and frequency of use in the bilingual lexical disadvantage hypotheses) and are therefore not mutually exclusive. Thus an alternative prediction is that that these factors will interact, such that bilinguals experience weaker language ties in their L1 as compared to monolinguals and even weaker ties in their L2, but still better cognitive control. If bilinguals have weaker language ties in their L1 because they use each language less often than monolinguals, this should generate the same type of negative shift in interference effects based on SOA manipulation in the L1 compared to monolinguals. The negative shift of the L2 will be even more pronounced due to weaker ties in the L1 and for monolinguals. In both L1 and L2, however, there should be a reduction in interference effects due to the enhanced cognitive control of bilinguals over monolinguals (i.e. the pattern in Figure 1a would be shifted to the left and down, and more so in the L2 than the L1).

To establish a "baseline" comparison for the bilingual participants, a manual Stroop SOA task was administered to English monolinguals in Experiment 1.[2] Based on the results of this experiment, more specific predictions are then made for Experiments 2 and 3, which administered a manual Stroop SOA task to English–Chinese bilinguals and Chinese–English, respectively, in their L1 and L2.

Experiment 1: English monolinguals

Methods

Participants

Twenty-four English monolinguals from the University of Nottingham, England, participated in Experiment 1 (see Table 1 above). One subject was removed from the analysis because of having learned another language besides English from birth. The remaining 23 participants were 14 females and 9 males, who all were right-handed,

[2] An additional study was run with five SOAs (±400 ms, ±200 ms and 0 ms) in a vocal modality with 12 monolingual English speakers as a pilot study. The results showed significant interference at the –400 ms, –200 ms and 0 ms SOAs. Interference was greatest at the 0 ms SOA, and decreased with increasing negative SOAs. Significant facilitation was found at –400 ms and –200 ms but not at 0 ms. The vocal task thus replicated the results of Glaser and Glaser (1982).

https://doi.org/10.1017/S1366728912000405 Published online by Cambridge University Press

reported no color-blindness, had normal or corrected-to-normal vision, and were not fluent in any other language besides English.

Materials and design

Word stimuli were the words *red*, *green*, and *blue* in lowercase font. Control word stimuli consisted of *xxxx*; this was included as a non-word, non-color control condition. As SOA manipulation in the Stroop task necessitates a spatial separation of the color and word stimuli, words were presented inside a colored rectangle, as in Glaser and Glaser (1982). Color stimuli consisted of red, green and blue filled rectangles of 284 × 142 pixels with a smaller black-filled rectangle inside. Word stimuli were presented in white ink inside the black rectangle. Congruent stimuli presented the same word and color (e.g. *red* surrounded by a red rectangle). Incongruent stimuli presented non-matching words and colors (e.g. *green* surrounded by a blue rectangle). Control stimuli presented *xxxx* surrounded by red, green or blue rectangles.

Procedure

Stimuli were presented using E-Prime software (Psychology Software Tools, Pittsburgh, PA). There were a total of 10 blocks in the experimental session, two for each SOA. The order of SOA block presentation was counterbalanced, and congruency was randomized within blocks. Each block consisted of 54 stimuli, with 18 each of congruent, incongruent and control trials, resulting in a total of 540 trials. In each trial, a fixation cross appeared for 500 ms, followed by a blank screen for 300 ms, and then the word and/or color stimuli appeared.

Five SOAs were included: –400 ms, –200 ms, 0 ms, +200 ms, and +400 ms. In negative SOAs, the word stimulus appeared on the screen alone first for either 400 ms or 200 ms, followed by the colored rectangle. In positive SOAs, the colored rectangle appeared on the screen alone for either 200 ms or 400 ms before the word stimulus appeared in the center of the rectangle. In the 0 ms SOA, the word stimulus and the colored rectangle appeared simultaneously. Both stimuli remained on the screen until participants made a response; if no response was made in 2000 ms, the stimuli disappeared and the next trial began. Participants were asked to respond manually to the color using an external button-box (right index finger for red, right middle finger for green, right ring finger for blue). To help with the initial finger-to-color mappings, the button-box was labeled with color patches. Once the experiment began, subjects were instructed to keep their gaze fixated in the center of the screen. Participants conducted a brief practice session to familiarize them with the mappings between buttons and colors. Following the completion of the experimental session, participants completed a language background questionnaire and received an inconvenience allowance of £4 for their participation.

Results and discussion

Data were first trimmed to remove incorrect responses (3.8%) and outliers (RTs of less than 250 ms or greater than 1500 ms: 0.4%). The mean number of errors per condition ranged between 0.2% and 0.4%. Because error rates were very low no error analyses were conducted.

A 3 (congruency: congruent, control, incongruent) × 5 (SOA) repeated-measures ANOVA showed a significant main effect of congruency ($F(2,44) = 71.31$, $p < .0001$) and of SOA ($F(4,88) = 5.49$, $p < .01$), and a significant interaction between SOA and congruency ($F(8,176) = 21.33$, $p < .0001$). Paired-sample *t*-tests were run between the incongruent vs. control and control vs. congruent conditions to compare interference and facilitation effects, respectively, at each SOA. The *p*-values of these paired-sample *t*-tests were Bonferroni-corrected, and only significant results ($p < .05$) after correction are reported. Bonferroni corrections were applied due to the large number of tests performed (10 per group). The *t*-tests revealed significant interference at –200 ms (74 ms; $t(22) = 10.78$, $p < .0001$), and 0 ms (57 ms; $t\ (22) = 5.68$, $p < .0001$). Significant facilitation was found at –400 ms (51 ms; $t(22) = 6.46$, $p < .0001$) and at –200 ms SOA (27 ms; $t(22) = 3.29$, $p < .05$). As can be seen in Figure 2a, the –200 ms SOA elicited the longest incongruent RT ($M = 639$ ms, $SE = 21$ ms), and the greatest amount of interference ($M = 74$ ms, $SE = 7$ ms).

The results of Experiment 1, which administered a Stroop task with five SOAs in a manual response modality, clearly showed a different pattern to those of Glaser and Glaser (1982) and Roelofs (2010). The greatest interference effects were not found at 0 ms SOA but at the –200 ms SOA, indicating a negative shift of interference effects due to the response modality. Previous Stroop literature has reported performance differences depending on response modality (Barch, Braver, Akbudak, Conturo, Ollinger & Snyder, 2001; MacLeod, 1991; Weekes & Zaidel, 1996), such that vocal responses generally elicit overall slower RTs than manual responses. Importantly, these results are in line with recent research that has also found a negatively-shifted peak in interference effects with a manual response (Appelbaum et al., 2009; Coderre & van Heuven, 2012).

Pre-exposure of the word at –200 ms SOA in a manual task may interfere with the faster motor response, while simultaneous presentation at 0 ms SOA interferes with the slower vocal response. For example, in both manual and vocal modalities, processing of the word stimulus presumably proceeds at the same rate. But in a vocal task, the color stimulus must be translated into a vocal response, requiring additional time than a manual response in which

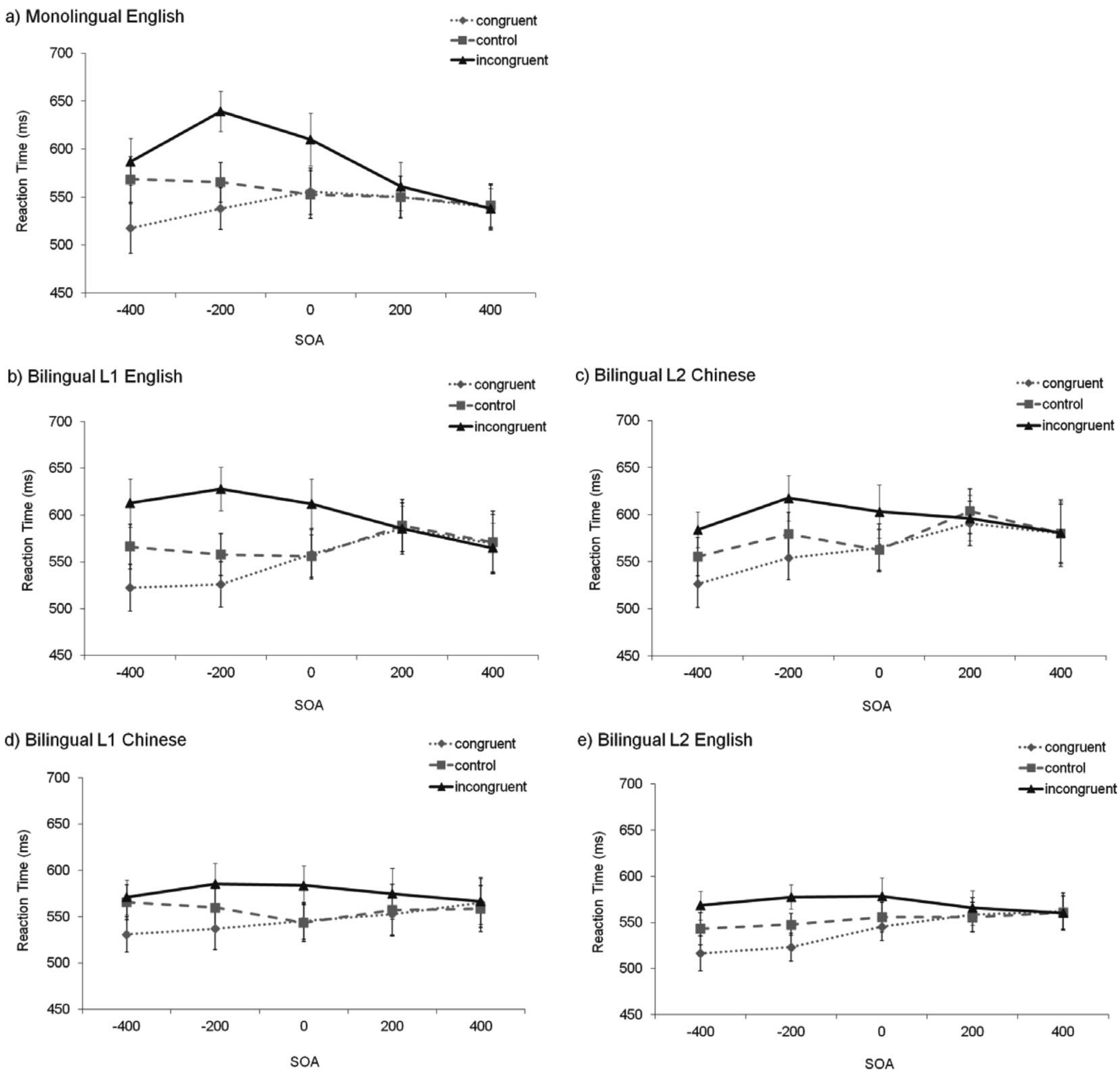

Figure 2. Mean RTs (ms) for the (a) monolinguals (Experiment 1); (b) bilingual L1 English (Experiment 2); (c) bilingual L2 Chinese (Experiment 2); (d) bilingual L1 Chinese (Experiment 3); and (e) bilingual L2 English (Experiment 3).

the color is mapped directly to a motor response. With simultaneous color and word presentation in a vocal response, word semantics and vocalization of the color word run at similar time courses and therefore cause stronger interference. However, pre-exposure of the word in the vocal modality gives semantic processing a head-start such that it is inhibited by the time the color word is vocalized, resulting in smaller interference. In contrast, when the color and word are presented simultaneously in a manual modality, the faster and more direct color mapping allows for a response to be selected before the word is fully active, creating smaller interference. With a pre-exposed word in a manual task, the time courses of word and color processing once again line up, creating maximal interference at a more negative SOA. The results of Experiment 1 thus demonstrate that response modality influences the RT patterns seen in a Stroop SOA task, and that the peak interference is negatively shifted in a manual task such that it occurs at the –200 ms SOA rather than the 0 ms SOA.

Experiment 2: English–Chinese bilinguals in L1 and L2

We next investigate the contributions of executive control abilities and speed of lexical access on Stroop

https://doi.org/10.1017/S1366728912000405 Published online by Cambridge University Press

performance in bilinguals. In light of the results of Experiment 1, specific predictions for each hypothesis can now be made about the patterns of interference elicited by a manual SOA Stroop task.

Predictions for bilinguals

The bilingual cognitive advantage hypothesis predicts smaller interference effects for bilinguals in both L1 and L2 compared to monolinguals due to enhanced executive control processes, but only at the SOA which experiences the greatest cognitive demand (see Figure 3a). As shown by Experiment 1, this is the –200 ms SOA. Facilitation effects may also be decreased in bilinguals, which is expected to occur where facilitation is generally greatest in monolinguals: at the –400 ms SOA (Figure 3b). Importantly, because the bilingual cognitive advantage hypothesis makes no claims about the speed of lexical access, interference and facilitation effects should not be shifted by SOA modulation.

The bilingual lexical disadvantage hypotheses predict a negative shift of peak interference due to the relatively decreased proficiency of both the L2 and L1 in bilinguals compared to monolinguals. Therefore, the bilingual L1 should experience the most interference at an earlier negative SOA than monolinguals, and the bilingual L2 at an earlier negative SOA than the L1 (see Figure 3c). Importantly, the magnitude of interference effects will be the same for each group, and only the latency of peak interference will be shifted according to the varying strengths of language ties in each group. Due to the latencies of the SOAs used in the current studies, it may be difficult to find the precise point where each group shows the peak interference; for example, lexical access in the bilingual L1 may only be delayed by 50 ms or 100 ms, making the peak interference effect fall between the –200 ms and –400 ms SOAs. In this case a relative plateau of interference is predicted between the SOAs, as in the bilingual L1 line in Figure 3c. Because peak facilitation occurs at the –400 ms SOA, negative shifts in facilitation effects may not be observed as this would be outside the range of SOAs (Figure 3d).

A combination of above hypotheses predicts that bilinguals will show less interference overall than monolinguals, but that interference and facilitation will also be shifted to negative SOAs in the L1 and L2 compared to monolinguals such that the peak interference and facilitation effects occur earlier for bilinguals (see Figure 3, panels (e) and (f)).

Experiment 2 tested a group of bilinguals with the same native language as the English monolinguals of Experiment 1, to compare the impact of bilingualism on executive control abilities and lexical access speed.

Methods

Participants

Participants were 15 English–Chinese bilinguals (English L1, Mandarin Chinese L2) from the University of Nottingham in Nottingham, England (10 female, 5 male). All were right-handed, reported no color-blindness, and had normal or corrected-to-normal vision. Participants were native English speakers who rated themselves dominant in English but also proficient in Mandarin. All participants completed a language background questionnaire after the testing session (see Table 1 above). Most spoke other languages (N = 11), including Spanish, French, German, and Malay, and some (N = 4) considered themselves more proficient in their other languages than in Chinese (overall self-rated proficiency in other languages: 6.3, on a 10-point scale). Their overall self-reported Chinese proficiency was 6.5 ($SD = 1.3$; averaged across reading, writing, speaking and listening self-reported proficiency scores), and their first contact with Chinese occurred at a mean age of 10.3 years ($SD =$ 7.3 years).

Materials and design

Word stimuli for the English (L1) version of the task were identical to Experiment 1, with the exception that the English control stimulus was *%%%%*.[3] Word stimuli for the Chinese (L2) version of the task consisted of the simplified Chinese characters 红, 绿, and 蓝 (“red”, “green” and “blue”, respectively). The Chinese control character was *%*, in order to match the approximate physical size of a character. Characters were printed in white font against a black background. Participants were instructed to ignore the English word or Chinese character and press a button corresponding to the color of the rectangle. Each participant performed the Chinese and English versions of the Stroop task on two consecutive days, and language order was counterbalanced between participants. Bilinguals therefore had more practice with the task than monolinguals; however, practice effects in the Stroop task are highly specific and transient (MacLeod, 1991), so this should not affect the results. As in Experiment 1, five SOAs were used; SOA was blocked, and the order of SOAs was counterbalanced between participants.

[3] Although the control stimuli used different symbol strings between Experiment 1 (*xxxx*) and Experiment 2 (*%%%%*), the magnitude of the interference should not be affected. To confirm, we compared the control conditions between the monolinguals and the bilingual L1 English. A five-way (SOA) ANOVA on the control RTs only with language as a between-subjects factor showed no main effects of SOA or language group (all Fs < 1), although there was a weak trend of an interaction ($F(4,143) = 2.02$, $p = .09$). The lack of a main effect of group indicates that the difference in control conditions between the experiments does not affect the results.

https://doi.org/10.1017/S1366728912000405 Published online by Cambridge University Press

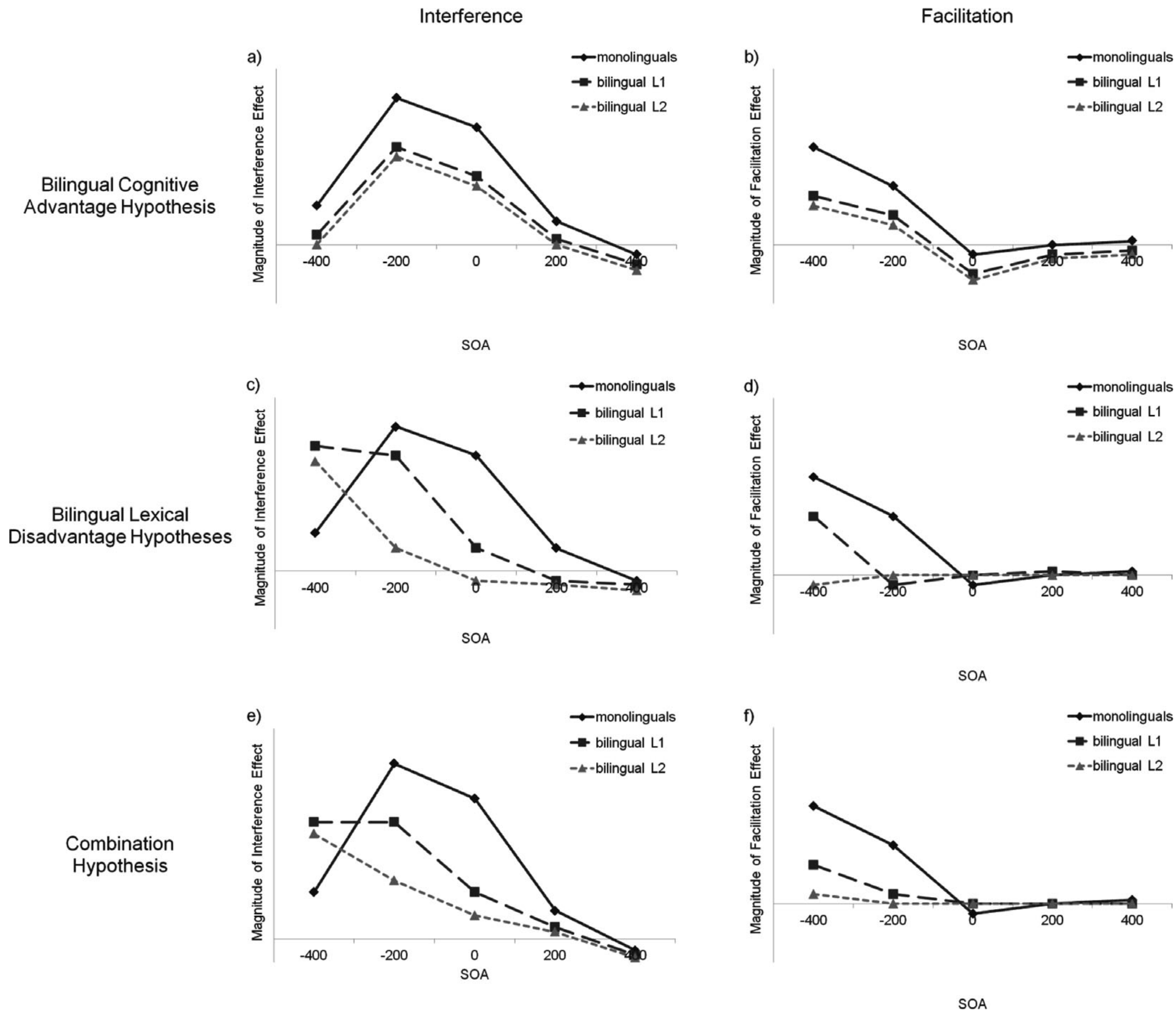

Figure 3. Panels (a) and (b): The bilingual cognitive advantage hypothesis predicts that bilinguals will have better cognitive control than monolinguals, but this control will be unaffected by SOA manipulation, resulting in an overall downward shift in (a) interference and (b) facilitation effects. Panels (c) and (d): The bilingual lexical disadvantage hypotheses predict that earlier negative SOAs will cause more (c) interference and (d) facilitation in the weaker language due to word pre-exposure, such that peak effects will be negatively shifted in the L1 vs. monolinguals and in L2 vs. L1. Panels (e) and (f): A combination of all three hypotheses would lead to a slightly negative shift in bilingual L1, even more of a negative shift in bilingual L2, but still overall reduced (e) interference and (f) facilitation as compared to monolinguals. Monolingual data are based on the interference and facilitation effects from Experiment 1.

Procedure

The procedure was similar to that of Experiment 1. Following the completion of the experiment, participants completed a language background questionnaire and received an inconvenience allowance of £8.

Results and discussion

English (L1) data

Overall, the mean RT pattern for the L1 English (Figure 2b above) appeared very similar to that of the monolinguals. A 3 (congruency) × 5 (SOA) repeated-measures ANOVA showed significant main effects of congruency ($F(2,26) = 60.23$, $p < .0001$) but not of SOA ($F(5,54) = 1.69$, $p = .17$, and a significant interaction of congruency by SOA ($F(8,109) = 14.31$, $p < .0001$). Paired-sample t-tests with Bonferroni corrections showed significant interference effects at –400 ms ($M = 47$ ms; $t(14) = 5.11$, $p < .001$), –200 ms ($M = 70$ ms; $t(14) = 7.33$, $p < .01$), and 0 ms ($M = 50$ ms; $t(14) = 6.23$, $p < .01$). Significant facilitation effects occurred at –400 ms ($M = 44$ ms; $t(14) = 4.10$, $p < .05$) and –200 ms

https://doi.org/10.1017/S1366728912000405 Published online by Cambridge University Press

(32 ms; $t(14) = 3.70$, $p < .05$) SOAs. As can be seen in Figure 2b, the slowest incongruent RT occurred at –200 ms SOA ($M = 618$ ms, $SE = 23$ ms), and the most interference occurred at –200 ms ($M = 70$ ms, $SE = 10$ ms).

Chinese (L2) data

The pattern of mean RTs for the L2 Chinese (Figure 2c) appeared different than those of the English conditions, especially with regards to the control condition. A 3 (congruency) × 5 (SOA) repeated-measures ANOVA showed a main effect of congruency ($F(2,24) = 17.93$, $p < .001$) and SOA ($F(4,53) = 3.45$, $p < .05$), and an interaction of congruency and SOA ($F(8,106) = 4.57$, $p < .001$). Paired-sample *t*-tests revealed significant interference effects at –200 ms (38 ms; $t(14) = 3.44$, $p < .05$) with a trend at 0 ms (41 ms; $t(14) = 3.03$, $p = .09$ corrected; $p = .009$ uncorrected). There were no significant facilitation effects after correction at any SOA (all *p*s > .17), although the largest effect occurred at –400 ms ($M = 29$ ms, $SE = 11$ ms). As can be seen in Figure 2c, the slowest incongruent RT occurred at –200 ms SOA ($M = 617$ ms, $SE = 24$ ms), and the most interference occurred at 0 ms SOA ($M = 41$ ms, $SE = 13$ ms).

The data therefore revealed similar patterns of RTs for the English Stroop task when performed in a native language. However, the pattern of RTs for L2 Chinese was somewhat different: most notably, slightly stronger interference effects were found at the 0 ms SOA than the –200 ms SOA. It is unclear whether this discrepancy arose from differences in native and non-native processing due to bilingualism, or from differences in linguistic processing between Chinese and English. To investigate these effects, Experiment 3 tested native Chinese speakers with English as their L2.

Experiment 3: Chinese–English bilinguals in L1 and L2

Methods

Participants

Twenty-four Chinese–English bilinguals (Mandarin Chinese L1, English L2) were tested in Experiment 2. Two participants were removed from the final analyses because they rated themselves as English-dominant. The remaining 22 included 19 females and 3 males, all of whom were right-handed, reported no color-blindness, and had normal or corrected-to-normal vision. Of the 22 participants, twelve were tested at the University of Nottingham Ningbo campus in Ningbo, China, and ten were tested at the University of Nottingham in England. The University of Nottingham Ningbo campus is an English-immersion environment, in which all classes are taught in English, so all participants tested in China were immersed in their non-native language, despite being in their native country. The subjects who were tested in England had all just arrived in the country from the Ningbo campus and had been living in England for no more than two months. Analyses of the data with location as a between-subjects factor showed no significant effects of testing environment on SOA or congruency effects (all *p*s > .25), so all Chinese–English bilinguals were considered together in subsequent analyses. All participants were native Mandarin Chinese speakers from mainland China who rated themselves dominant in Chinese but also proficient in English (see Table 1). Although some had learned other languages (N = 9), none rated themselves as very proficient in any foreign language (overall self-rated proficiency in other languages = 2.3). Most participants (N = 20) also spoke a Chinese dialect from their hometown, but as these dialects use simplified Chinese characters like Mandarin, this should not be a confounding factor. Their overall self-reported English proficiency was 6.6 ($SD = 1.2$), and they had their first contact with English at a mean age of 11 years ($SD = 2.7$ years). Importantly, they did not differ statistically from the English–Chinese bilinguals in either their overall self-rated proficiency scores ($p = .89$) or their age of L2 acquisition ($p = .76$).

Materials and design

The materials and design were identical to that of Experiment 2.[4]

Procedure

The procedure was identical to that of Experiment 2.

Results and discussion

Incorrect responses (2.6% for L1, 2.3% for L2) and outliers (RTs of less than 250 ms or greater than 1500 ms: 0.3% for L1, 0.2% for L2) were removed before statistical analysis. The mean number of errors per condition for L1 Chinese ranged between 0.1% and 0.2%. The mean number of errors per condition for L2 English ranged between 0.1% and 0.4%. Overall, error rates were again very low and therefore no error analyses are reported.

[4] To confirm again that the use of different control conditions did not affect the magnitude of interference between in Experiment 1 and Experiment 3, the control conditions were compared between the monolinguals and the bilingual L1 Chinese. A five-way (SOA) ANOVA on the control RTs with language as a between-subjects factor showed a main effect of SOA ($F(4,172) = 2.47$, $p < .05$), but no effect of language group ($F(1,43) < 1$) and no interaction ($F(4,172) = 1.06$, $p = .38$). The lack of a main effect of group indicates that the difference in control conditions between the experiments does not affect the results.

L1 Chinese data

The pattern of RTs in the L1 Chinese data (Figure 2d) exhibited a similar pattern to that of the L2 Chinese data from Experiment 2. A 3 (congruency) × 5 (SOA) repeated-measures ANOVA showed a significant main effect of congruency ($F(2,42) = 22.18$, $p < .0001$) but not SOA ($F(4,84) < 1$), and an interaction of SOA and congruency ($F(8,168) = 5.53$, $p < .0001$). Paired-sample *t*-tests with Bonferroni corrections indicated significant interference effects at –200 ms (26 ms; $t(21) = 3.17$, $p < .05$), and 0 ms (40 ms; $t(21) = 3.62$, $p < .05$). Significant facilitation effects occurred at –400 ms (35 ms; $t(21) = 4.37$, $p < .01$) and –200 ms (23 ms; $t(21) = 4.49$, $p < .01$). As can be seen in Figure 2d, the slowest incongruent RT occurred at –200 ms SOA ($M = 585$ ms, $SE = 22$ ms), and the most interference occurred at 0 ms SOA ($M = 40$ ms, $SE = 11$ ms), as in the L2 Chinese data.

L2 English data

The L2 English data (Figure 2e) showed a similar pattern as that of the monolingual and L1 English data in Experiments 1 and 2. A 3 (congruency) × 5 (SOA) repeated-measures ANOVA revealed a significant main effect of congruency ($F(2,42) = 19.72$, $p < .0001$) but not SOA ($F(4,84) = 1.44$, $p = .23$), and an interaction of SOA and congruency ($F(8,168) = 4.71$, $p < .0001$). Paired-sample *t*-tests showed significant interference at –200 ms (30 ms; $t(21) = 3.18$, $p < .05$). Significant facilitation occurred at –200 ms (25 ms; $t(21) = 3.89$, $p < .01$) only. As can be seen in Figure 2e, the longest incongruent RT occurred at 0 ms SOA ($M = 578$ ms, $SE = 20$ ms), and the most interference at –200 ms ($M = 30$ ms, $SE = 9$ ms).

Comparison of all three groups (Experiments 1–3)

To investigate the initial questions regarding the bilingual advantage and delay hypotheses, all three groups were directly compared: English monolinguals from Experiment 1; English–Chinese bilinguals from Experiment 2; and Chinese–English bilinguals from Experiment 3. An initial visual comparison of the raw RT patterns for each group (Figure 2) reveals an interesting effect of script, in that the English and Chinese languages elicit different RT patterns, independent of language status (native or non-native). This indicates that there may be an underlying effect of script driving the resultant RT patterns. Specifically, the control condition seems to behave differently in Chinese and English, especially at the 0 ms SOA. As the groups were compared on interference effects (incongruent minus control), this script difference could be driving some of the effects that are discussed in the next section. Comparisons of the Stroop effect (incongruent minus congruent) may seem more appropriate; however, as SOA manipulation also affects the amount of facilitation (e.g. increased facilitation in the –400 ms SOA), interpretations of the Stroop effects are just as difficult. These script differences, and the implications they hold for our interpretation of the data, will be examined in the "General discussion" section below.

Evaluation of the bilingual cognitive advantage hypothesis

To investigate the bilingual advantage on conflict processing, monolingual performance was compared to bilingual performance in both native languages: monolinguals vs. L1 English; and monolinguals vs. L1 Chinese. Monolingual English and L1 English interference effects (Figure 4 below, panels (a) and (b)) were compared at the –200 ms SOA using an independent-samples *t*-test (with Welch corrections to the degrees of freedom where the assumption of sphericity was violated), which showed no significant difference in the magnitude of interference ($p = .76$). Therefore when comparing the two groups of native English speakers, no bilingual advantage was observed. For English monolinguals vs. L1 Chinese (Figure 4, panels (a) and (d)), the largest interference effects (–200 ms SOA for monolinguals and 0 ms SOA in L1 Chinese bilinguals) were compared using an independent-samples *t*-test. This comparison showed a significant effect of bilingual status ($t(35.1) = 2.56$, $p < .05$), such that the L1 Chinese showed smaller interference effects ($M = 40$ ms, $SE = 5$ ms) than monolinguals ($M = 74$ ms, $SE = 3$ ms). Therefore a bilingual advantage in the L1 compared to monolinguals occurred for the Chinese–English bilinguals but not the English–Chinese bilinguals.

If bilingualism leads to generalized enhanced cognitive control, this should be evident in both the L1 and L2 in bilingual participants. To test whether the bilingual advantage also occurred in the second language, the magnitude of interference was compared between monolinguals and bilingual L2s. Comparing monolinguals versus L2 English (Figure 4, panels (a) and (e)), there was a significant reduction in interference effects for bilinguals' L2 English ($M = 30$ ms, $SE = 9$ ms) compared to monolinguals ($M = 74$ ms, $SE = 3$ ms; $t(38.8) = 3.79$, $p < .001$) at the –200 ms SOA (the SOA of peak interference for both groups). Thus, the Chinese–English bilinguals showed smaller interference effects in both their L1 and L2 compared to monolinguals, demonstrating a bilingual advantage across languages. To compare monolinguals versus L2 Chinese (Figure 4, panels (a) and (c)), once again the SOA which showed the largest interference effects in each group was selected (–200 ms SOA for monolinguals, 0 ms SOA for L2 Chinese). There was a significant difference in interference magnitude ($t(21.4) = 2.22$, $p < .05$) such that the interference effects were smaller for L2 Chinese ($M = 41$ ms, $SE = 13$ ms) than for monolinguals

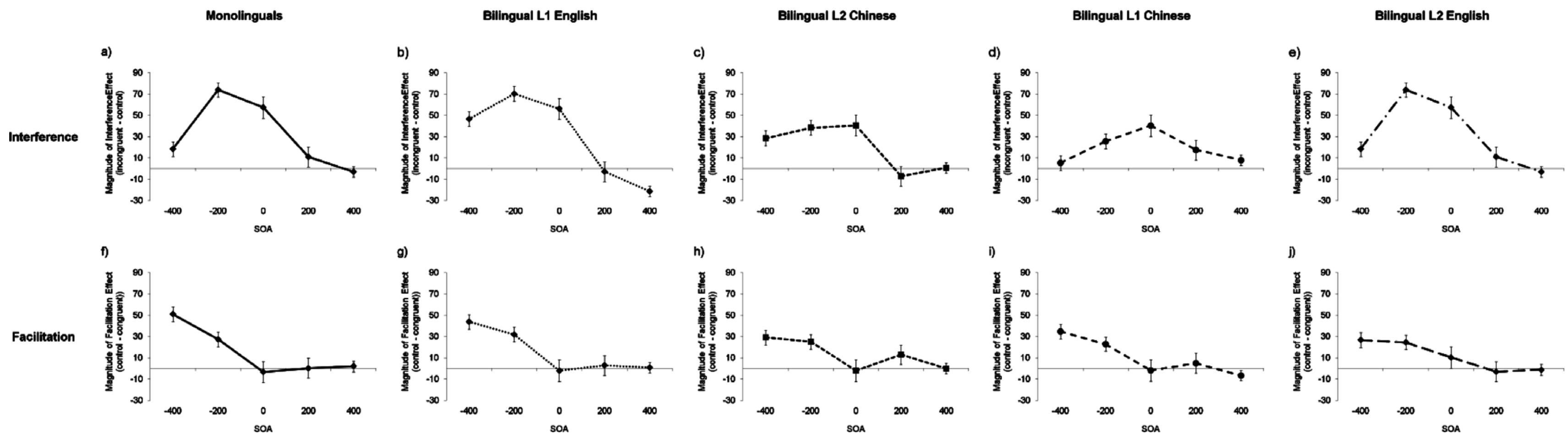

Figure 4. Comparison of the magnitude of interference (panels (a)–(e)) and facilitation (panels (f)–(j)) effects in monolinguals, L1 English, L2 Chinese, L1 Chinese, and L2 English.

($M = 74$ ms, $SE = 3$ ms). This might suggest enhanced cognitive control in the L2 for the English–Chinese bilinguals. However, as a similar advantage was not found in the L1 English, it is unclear whether the decreased interference effects in the L2 result from enhanced cognitive control, or from less interference from the second language due to the relatively reduced proficiency. In other words, because an advantage in interference effects did not occur in BOTH languages for the English–Chinese bilinguals, this cannot be interpreted as a general bilingual cognitive advantage.

Turning to the facilitation effects, Figure 4 (panels (f)–(j)) illustrates that at the –400 ms SOA monolinguals showed the greatest facilitation ($M = 51$ ms, $SE = 8$ ms; Figure 4f), followed by the bilingual L1 Chinese ($M = 37$ ms, $SE = 10$ ms; Figure 4i), L1 English ($M = 44$ ms, $SE = 11$ ms; Figure 4g), L2 Chinese ($M = 29$ ms, $SE = 11$ ms; Figure 4h) and L2 English ($M = 27$ ms, $SE = 12$ ms; Figure 4j). If, as mentioned previously, larger Stroop facilitation effects are indicative of poorer cognitive control, this pattern fits with the prediction of monolinguals showing poorer performance, and supports other findings of greater cognitive control abilities in the Chinese–English bilinguals. However, none of the groups differed statistically in the magnitude of the facilitation effects (though there was a trend between the monolinguals and the L2 English: $p = .09$ uncorrected). Therefore the facilitation effects offered no significant evidence for the bilingual cognitive advantage.

According to the bilingual cognitive advantage hypothesis, increased language proficiency should lead to enhanced cognitive control abilities (Bialystok, Craik & Ruocco, 2006; Linck, Hoshino & Kroll, 2008). To investigate whether the bilingual advantage is mediated by proficiency within bilinguals, the L2 performance in each bilingual group was compared using a median split of self-rated L2 proficiency (Figures 5 and 6) below.

English–Chinese bilingual performance split by L2 (Chinese) proficiency

In L1 English, there were no significant differences between the low- and high-proficiency groups in either interference (all *p*s > .42) or facilitation effects (all *p*s > .25) at any SOAs. Thus, Chinese proficiency did not influence the Stroop interference and facilitation patterns in L1 English.

In L2 Chinese, there was an interesting effect of Chinese proficiency on the overall mean RTs. Specifically, the low-proficiency bilinguals (Figure 5a) showed an RT pattern similar to English (Figure 2, panels (a) and (b)), while the high-proficiency bilinguals (Figure 5b) showed an RT pattern similar to native Chinese (Figure 2d). This suggests that with increasing proficiency comes more native-like processing of the foreign language. However, there were no significant differences in interference between the low- and high-proficiency groups at any SOA (all *p*s > .68; Figure 6a), indicating that the overall magnitude of interference in the L2 was not affected by proficiency. No differences in facilitation effects were found in the L2 Chinese between low-proficiency ($M = 27$ ms, $SE = 20$ ms) and high-proficiency ($M = 31$ ms, $SE = 9$ ms) bilinguals ($p = .87$).

Chinese–English bilingual performance split by L2 (English) proficiency

In L1 Chinese, there were no significant differences between the low- and high-proficiency groups at any SOAs either in interference (all *p*s > .32) or facilitation effects (all *p*s > .21). Thus, English proficiency did not influence the Stroop interference and facilitation patterns in L1 Chinese.

For L2 English, the RT patterns for the low-proficiency bilinguals (Figure 5c) appeared similar to those of English (Figure 2, panels (a) and (b)), but the high-proficiency participants (Figure 5d) showed a different pattern, unlike English or Chinese. There was a difference in interference magnitude at the –200 ms SOA between proficiency groups ($t(15.9) = 2.38$, $p < .05$; Figure 6b) such that low-proficiency participants showed greater interference effects ($M = 50$ ms, $SE = 15$ ms) than high-proficiency bilinguals ($M = 10$ ms, $SE = 8$ ms). There was also a trend towards a significant effect at the 0 ms SOA ($t(15.2) = 1.97$, $p = .07$). Reduced facilitation effects also occurred in the high-proficiency L2 English bilinguals ($M = 9$ ms, $SE = 12$ ms) compared to the low-proficiency group ($M = 44$ ms, $SE = 19$ ms), though this was only a statistical trend ($t(106.4) = 1.72$, $p = .08$). These data provide support for the bilingual advantage hypothesis but not for the lexical disadvantage hypotheses. According to the bilingual L2 lexical disadvantage hypothesis, the L2 experiences delayed lexical access relative to the L1 due to the relatively reduced proficiency; this predicts that in the L2, low-proficiency bilinguals would show smaller interference effects than high-proficiency bilinguals because the word generates less interference. However, in the present data, the low-proficiency bilinguals actually showed stronger interference effects. This indicates that Stroop performance is largely driven by executive control abilities, which are related to proficiency in a second language, with increasing proficiency leading to stronger L2 activation and increased language conflict, and therefore a greater need for cognitive control.

To summarize the analyses of the bilingual cognitive advantage hypothesis, the present data revealed a bilingual advantage for the Chinese–English participants, (i) in L1 Chinese compared to English monolinguals; and (ii) in L2 English between high and low-proficiency participants. There was a reduction in interference effects in the English–Chinese L2 compared to monolinguals; however, a similar reduction did not occur in the L1 for this group.

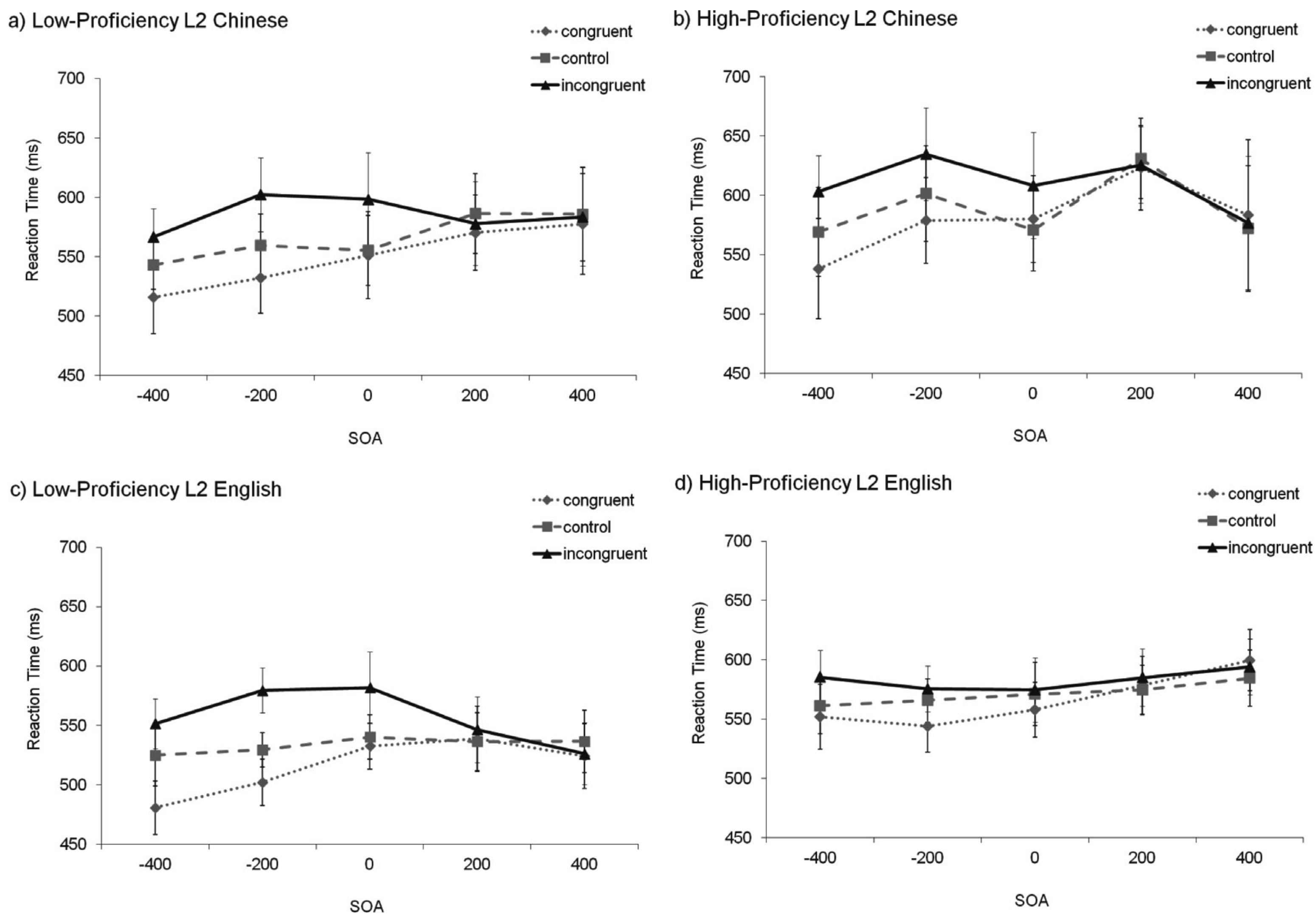

Figure 5. Mean RTs (ms) after the L2-proficiency split in each bilingual group. (a) low-proficiency English–Chinese bilinguals in L2 Chinese; (b) high-proficiency English–Chinese bilinguals in L2 Chinese; (c) low-proficiency Chinese–English bilinguals in L2 English; (d) high-proficiency Chinese–English bilinguals in L2 English.

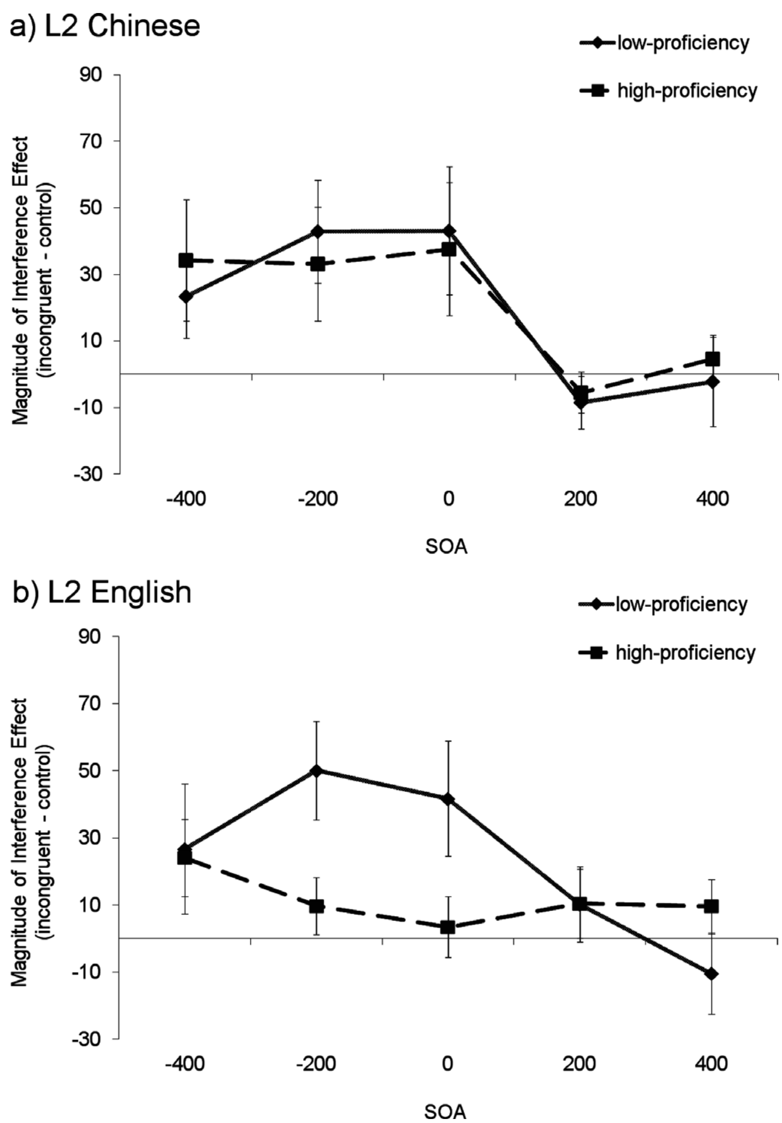

Figure 6. Mean L2 interference effects for each bilingual group split by proficiency. (a) low- vs. high-proficiency English–Chinese bilinguals on the L2 Chinese Stroop task; (b) low- vs. high-proficiency Chinese–English bilinguals on the L2 English Stroop task.

Thus, there was no bilingual advantage in the English–Chinese bilinguals, as enhanced cognitive control abilities did not occur for both languages of these bilinguals.

Evaluation of the bilingual lexical disadvantage hypotheses

To investigate the potential impact of lexical access speed, a series of comparisons were conducted using peak interference and facilitation effects (Figure 4 above) as indicators of the window of lexical access. Latency analyses were conducted by identifying the SOA that exhibited the maximum interference and facilitation for each group and comparing groups using *t*-tests. First, between-subjects analyses were conducted for each language to investigate the effects of bilingualism and proficiency on the latency of interference within the same language. Next, the monolinguals were compared to the bilingual L1 in each group to investigate the bilingual L1 lexical disadvantage hypothesis; and finally, the L1 and L2 were contrasted for the two groups of bilinguals to investigate the bilingual L2 lexical disadvantage hypothesis.

In all groups performing in English (i.e. monolinguals, L1 English, L2 English), the peak interference occurred at –200 ms SOA (Figure 4, panels (a), (b), and (e)). Comparisons revealed no significant differences between the peak interference of English monolinguals and L1 English bilinguals ($t(27.2) = 1.01$, $p = .32$) or between English monolinguals and L2 English bilinguals ($t(36.6) = 1.49$, $p = .14$). There was also no difference

https://doi.org/10.1017/S1366728912000405 Published online by Cambridge University Press

across bilinguals between L1 English and L2 English ($p = .64$). This indicates that RTs were affected similarly in the English script, regardless of bilingual status.

The two Chinese conditions were also compared (L1 Chinese vs. L2 Chinese; Figure 4, panels (c) and (d)) to evaluate the effects of native-speaker status on the latency of peak interference effects in the logographic script. There was a strong trend towards a significant difference in the peak interference latency between L1 Chinese and L2 Chinese bilinguals ($t(30.5) = 1.98$, $p = .06$) such that the majority of L1 Chinese bilinguals showed peak interference effects at the 0 ms SOA, whereas the peak interference effects of the L2 Chinese bilinguals were more evenly spread over the –200 ms and 0 ms SOAs. The issue of how script affects the latency of interference effects will be addressed further in the "General discussion" section below.

In addressing the bilingual L1 lexical disadvantage hypothesis, as reported above there were no differences between the peak interference of English monolinguals and L1 English bilinguals ($t(27.2) = 1.01$, $p = .32$). There was a significant difference between the peak interference effects for monolinguals relative to L1 Chinese bilinguals, ($t(37.3) = 2.27$, $p < .05$) such that monolinguals experienced peak interference at the –200 ms SOA, whereas the L1 Chinese elicited maximal interference at the 0 ms SOA (Figure 4, panels (a) and (d)). Contrary to the negative shift predicted by the bilingual lexical disadvantage hypotheses, this is a positive shift in bilinguals relative to monolinguals. However, as already mentioned, there were differences across the groups in the patterns of interference for English and Chinese scripts: it may be that the difference in peak interference latency between monolinguals and L1 Chinese is due to these script differences rather than proficiency-driven differences in lexical access speed (see "General discussion").

To investigate the bilingual L2 disadvantage hypothesis within bilingual groups, the SOA generating the maximum interference effect was identified for each subject in both L1 and L2, and compared using paired-sample *t*-tests. Comparing L1 and L2 within the Chinese–English subjects, a negative shift in interference effects occurred in L2 English compared to L1 Chinese ($t(21) = 4.16$, $p < .001$) such that the L1 elicited the peak interference at the 0 ms SOA while the L2 showed maximal interference at the –200 ms SOA (Figure 4, panels (d) and (e)). In contrast, peak interference latencies in the English–Chinese bilinguals did not differ between L1 and L2 ($t(14) = 0.54$, $p = .60$; Figure 4, panels (b) and (c)). Therefore the Chinese–English bilinguals, but not the English–Chinese bilinguals, demonstrated a significant negative shift in peak interference effects as predicted by the bilingual L2 lexical disadvantage hypothesis.

An L2 proficiency split was again performed to investigate whether proficiency modulates the latency of peak interference effects within bilinguals' L2. Comparing the high- and low-proficiency participants in each bilingual group, there were no differences in peak interference effects either in English–Chinese ($t(12.8) = 1.42$, $p = .18$) or Chinese–English ($t(16.7) = 0.36$, $p = .72$) bilinguals. For the facilitation effects, all groups showed maximum effects at the –400 ms SOA and there were no significant differences between any of the groups (all $ps > .16$).

To summarize the analyses of the bilingual lexical disadvantage hypotheses, the data do not provide any evidence for a delay in lexical access between bilinguals' L1 and monolinguals, and is instead indicative of notable differences between the English and Chinese scripts. However, when comparing similar languages, a difference in peak interference occurred between L2 Chinese relative to L1 Chinese. There was also a significant shift between the L1 and L2 within Chinese–English bilinguals such that L2 English experienced peak interference effects at –200 ms, while the L1 Chinese peaked at 0 ms.

General discussion

The current experiments administered a Stroop SOA task to monolinguals and bilinguals to investigate how lexical access speed and executive control abilities modulate Stroop performance. Experiment 1, with monolinguals, demonstrated that an SOA Stroop task using a manual response modality shows peak interference effects at the –200 ms SOA, instead of the 0 ms as in a vocal task. This indicates that interference effects occur earlier in a manual Stroop task due to the nature of the faster motor response. In the English–Chinese bilingual data of Experiment 2, both the L1 English and L2 Chinese showed a roughly similar RT pattern as the monolingual data, in that negative SOAs experienced significant interference effects but no facilitation or interference effects remained at positive SOAs. However, the overall RT pattern was noticeably different between English and Chinese, a finding that was replicated in Experiment 3 with Chinese–English bilinguals. We first review the evidence found for the bilingual cognitive advantage hypothesis, then focus on the bilingual lexical disadvantage hypotheses, before discussing the effect of script and its potential influence on the results seen here.

The bilingual cognitive advantage hypothesis

The current data revealed some evidence for the bilingual cognitive advantage hypothesis, although the benefit in executive control was sensitive to proficiency and language immersion. No evidence was found for a bilingual advantage in the English–Chinese bilinguals: there were smaller interference effects in the L2 compared to the English monolinguals, but not in the

https://doi.org/10.1017/S1366728912000405 Published online by Cambridge University Press

L1, rendering it unclear whether the effect in the L2 was due to enhanced cognitive control abilities or less interference from the weaker language. In contrast, the Chinese–English bilinguals showed significantly smaller interference effects in both the L1 and L2 compared to monolinguals, which could be indicative of enhanced cognitive control abilities in this group. Alternatively, the difference in script could contribute to this effect in the L1 Chinese. Some previous studies have found that Chinese elicits smaller Stroop interference effects than alphabetic languages (e.g. van Heuven, Conklin, Coderre, Guo & Dijkstra, 2011), in which case it may not be enhanced cognitive control leading to the reduction in L1 Chinese interference effects, but rather the effect of processing the Chinese language. However, other studies directly comparing monolingual speakers of Chinese and alphabetic languages have shown that Chinese elicits larger interference effects (Biederman & Tsao, 1979; Saalbach & Stern, 2004; Tsao, Wu & Feustel, 1981). Thus though the present results indicate a bilingual advantage in the L1 Chinese compared to English monolinguals, the possibility remains that script differences may influence the magnitude of the interference effect (see section on the Effects of Script below).

In the Chinese–English bilinguals, the L2 proficiency split revealed smaller interference effects in L2 English for high-proficiency bilinguals than low-proficiency bilinguals. According to the bilingual L2 lexical disadvantage hypothesis, lower proficiency in the L2 predicts less interference from the distracting word due to weaker language ties; however, these data show LARGER interference effects for the low-proficiency group in L2 English performance. Thus this effect is not due to weaker L2 language ties in the low-proficiency group, but to increased executive control abilities (and therefore smaller interference) in the high-proficiency group. Although this is not a bilingual advantage in the traditional sense, as it is within-group rather than compared to monolinguals, it indicates that high-proficiency bilinguals may be more adept at controlling interference from their L2, demonstrating superior cognitive control abilities with increasing language proficiency.

Therefore a bilingual advantage was observed for the L1 Chinese bilinguals but not for the L1 English bilinguals, despite similar levels of subjective proficiency (see Methods, Experiment 3). One potential explanation for this disparity is that the Chinese–English bilinguals were immersed in the foreign language, while the English–Chinese bilinguals were not. Living in the foreign country and/or hearing the non-native language every day may have created more long-term and sustained language conflict, consequently boosting the bilingual advantage in the Chinese–English bilinguals. There is a shortage of research on the role of immersion on the bilingual advantage, but one study by Linck, Kroll & Sunderman (2009) demonstrated that L2 speakers immersed in the foreign-language environment had reduced access to their L1, suggesting enhanced L1 inhibition (see Green, 1998). In other words, bilinguals immersed in their weaker language (usually L2) engage cognitive control to a greater extent on a daily basis due to the need to avoid interference from the dominant language, which would predict larger cognitive advantages for immersed bilinguals. In order to fully explore how immersion experience affects the bilingual advantage, future research should seek to balance the immersion background of participants, by including for example groups of L2-immersed and non-immersed bilinguals.

Overall, the above results highlight the sensitivity of bilingual executive control abilities to factors such as proficiency and immersion in the non-native language. This is in line with the current literature on the bilingual cognitive advantage. One recent review (Hilchey & Klein, 2011) has revealed that the bilingual advantage in interference effects (i.e. differences between incongruent and congruent conditions) is actually not often found, and is very sensitive to the task used and the type of executive control being tested. The more robust finding is a "global RT advantage", such that bilinguals perform faster on all trial types, incongruent AND congruent, than monolinguals (e.g. Kousaie & Phillips, 2012; see review in Tao, Marzecová, Taft, Asanowicz & Wodniecka, 2011). This global RT advantage is hypothesized to reflect enhanced monitoring and maintenance abilities in bilinguals, rather than enhanced cognitive control. To investigate the bilingual global RT advantage in the current data, a *post-hoc* comparison of global RTs for each group was performed. The results revealed that the English–Chinese bilinguals were slowest overall, in Chinese first ($M = 578$ ms, $SE = 25$ ms) and then in English ($M = 573$ ms, $SE = 26$ ms); the monolinguals ($M = 560$ ms, $SE = 24$ ms) and L1 Chinese ($M = 560$ ms, $SE = 22$ ms) were the same; and L2 English was the fastest ($M = 555$ ms, $SE = 17$ ms). As the monolinguals fell in the middle of this distribution, this does not support the bilingual global RT advantage hypothesis.

To summarize, the data of the current experiments provided some evidence for the bilingual cognitive advantage hypothesis, primarily in the reduced magnitude of L1 Chinese interference compared to monolinguals and in the L2 English proficiency split. Above all, the results highlight that the bilingual advantage is elusive, and sensitive to a number of variables including proficiency, language immersion, and script.

The bilingual lexical disadvantage hypotheses

The BILINGUAL L1 LEXICAL DISADVANTAGE hypothesis (i.e. reduced frequency hypothesis, relating to processing differences between monolinguals and the bilingual L1)

https://doi.org/10.1017/S1366728912000405 Published online by Cambridge University Press

predicted a negative shift in peak interference and facilitation effects for bilinguals in their L1 compared to monolinguals due to reduced frequency of use. The current data did not show a negative shift; in fact, L1 Chinese showed a positive shift (peak interference at 0 ms rather than –200 ms) compared to monolinguals. As mentioned previously, the English and Chinese scripts elicited different peak interference latencies, with maximal interference occurring at 0 ms in Chinese and at –200 ms in English. The difference in peak interference between monolinguals and the L1 Chinese is therefore more likely due to script differences than to differences in lexical access speed. Importantly, in comparing within the same language (English monolinguals vs. L1 English bilinguals), there were no differences in the latency of peak interference, indicating no difference in lexical access speed in a native language for bilinguals compared to monolinguals, as has been reported previously (Coderre et al., 2012).

The BILINGUAL L2 LEXICAL DISADVANTAGE HYPOTHESIS (i.e. temporal delay assumption, relating to the L2 and L1 in bilinguals) predicts a negative shift in interference effects for the L2 compared to the L1 due to reduced proficiency. In the within-groups comparison of L1 Chinese vs. L2 English, the peak interference effect in the L1 Chinese occurred at 0 ms, while the L2 English peaked at negative SOAs. As discussed, this could be due to differences in script. However, in the between-groups comparison of the L1 Chinese vs. the L2 Chinese, the L2 Chinese showed overall peak interference at more negative SOAs than the L1 Chinese. This supports the temporal delay assumption of the BIA+ model (Dijkstra & van Heuven, 2002) and is in line with previous literature supporting a temporal delay in activation of L2 representations in unbalanced bilinguals (e.g., Coderre et al., 2012; Newman et al., 2012; see for a review Dijkstra & van Heuven, 2002; van Heuven & Dijkstra, 2010). Therefore these data provide support for delays in lexical access due to reduced proficiency in the L2.

The effects of script

The present experiments were conducted with Chinese and English bilinguals to minimize the effects of phonological and orthographic overlap between languages. However, as is apparent from Figure 2, English and Chinese elicited different RT patterns in SOA manipulation. In particular, English experienced the peak interference effects at the –200 ms SOA, whereas Chinese peak interference occurred at the 0 ms SOA. This pattern was similar across English–Chinese and Chinese–English bilinguals, suggesting this is not an effect of proficiency but rather an underlying difference in the mechanisms of language processing in these two scripts.

The question of how linguistic processing differs between alphabetic and logographic languages is a highly contentious issue in the literature. In visual word reading in particular, a central question has been whether lexical access involves phonology. In alphabetic scripts like English, in which letters map directly onto sounds, phonology is activated as a part of word recognition, and plays a critical role in this process (e.g. Frost, 1998; see review in Perfetti, Liu & Tan, 2005). In Chinese, however, the role of phonology is more debatable. Being a logographic system, Chinese does not have letters that map onto sounds; rather, each character has a specific pronunciation. The same pronunciation is shared by many other characters, creating a large number of homophones. In other words, phonology is much less reliable in Chinese. For this reason it has been hypothesized that Chinese may use a "direct-access" route in word recognition, proceeding directly from orthography to semantics and bypassing phonology altogether (Saalbach & Stern, 2004; Taft & van Graan, 1998).

Although this direct-access hypothesis seems logical, a large body of accumulating evidence has demonstrated that this is not the case: phonology is in fact activated – obligatorily and even in the absence of lexical activation – in Chinese word recognition (Chua, 1999; Guo, Peng & Liu, 2005; Liu, Perfetti & Hart, 2003; Perfetti et al., 2005; Perfetti & Zhang, 1995; Saalbach & Stern, 2004; Spinks, Liu, Perfetti & Tan, 2000; Tan, Laird, Li & Fox, 2005; Xu, Pollatsek & Potter, 1999). The lexical constituency model of word recognition (Perfetti et al., 2005) argues that phonological access is a key constituent of word recognition in all languages, regardless of whether its activation is helpful or not. The degree to which it is useful and contributes to lexical access, however, can be mediated by script. For example, Tan and Perfetti (1997), using a phonologically-mediated priming paradigm in Chinese, demonstrated that the mediation effect was determined by homophone density: the more homophones a Chinese character had, the smaller the mediation priming. They proposed that phonology is accessed in the presence of a large number of homophones, but the phonological activation does not aid in semantic access due to a very distributed spread of activation. In the presence of fewer homophones, the activation of phonology can aid in semantic access, making phonology a more central part of visual word recognition. Extending this proposal to Chinese and English more generally, in the presence of a large number of homophones in Chinese, phonology is activated but is not helpful, so lexical access is effectively a direct link between orthography and semantics (supporting the direct access hypothesis). In English, consisting of fewer homophones, the role of phonology is more pronounced and lexical access is phonologically mediated.

In the current data, Chinese generated peak interference effects at the 0 ms SOA, while English experienced peak interference at the –200 ms SOA. If

https://doi.org/10.1017/S1366728912000405 Published online by Cambridge University Press

the role of phonology is the major difference between English and Chinese word recognition, this suggests that phonological access and/or mediation occurs at different speeds in each language. Specifically, as the peak interference occurred at a more negative SOA in English, English phonological mediation may be slower than Chinese (in the same way that delayed lexical access leads to negatively-shifted interference effects). This supports a previous finding by Saalbach and Stern (2004) reporting faster activation of phonology in Chinese than in English. The stronger influence of phonology in English could add an extra step in processing, making lexical access slower, whereas in Chinese the more direct pathway from orthography to semantics speeds up word recognition, creating different timings of lexical interference. More research that directly compares how phonology is mediated in alphabetic and logographic languages is needed, with particular regards to the timecourse of phonological activation in each writing system.

Another factor to consider is the use of a manual task in the current study rather than a vocal task, which may have downplayed the influence of phonology. Naming aloud necessarily requires access to phonology, but in a manual task this activation could theoretically be bypassed altogether. If so, the use of a manual task may have diminished the role of phonology in English, making it more like Chinese. Manual tasks in Chinese have still reported phonological access (e.g. Liu et al., 2003; Xu et al., 1999), indicating that the use of a manual modality does not eliminate phonological processing; however, this is a potentially important point to consider in future research.

As seen in Figure 2 (panels (c) and (d)), the larger interference effects in the 0 ms SOA in Chinese were driven by the control condition, which experienced a reduction in RT at the 0 ms SOA compared to the –200 ms or +200 ms SOAs. It is unclear why this only occurred in one SOA. One possibility is that because Chinese characters are more spatial in nature, participants exerted more attentional resources in the control condition in order to determine whether or not it was a Chinese character. The use of blocked SOAs in the current task design may also have affected the strategies that participants adopted: for example, in the 0 ms SOA, participants may not have invested as much effort into decoding the character because it occurred simultaneously with the target color, whereas in negative SOAs, the pre-exposure of the word allowed time to decipher the character. This explanation is not entirely satisfactory, as if participants were not attending to the character in the 0 ms SOA the incongruent condition would not have elicited a longer RT. The cross-linguistic similarity of bilinguals' languages may also have an impact on processing speed (e.g. Bates, D'Amico, Jacobsen, Székely, Andonova, Devescovi, Herron, Lu, Pechmann, Pléh, Wicha, Federmeier, Gerdjikova, Gutierrez, Hung, Hsu, Iyer, Kohnert, Mehotcheva, Orozco-Figueroa, Tzeng & Tzeng, 2003; Liu, Hao, Li & Shu, 2011), so it remains to be seen whether this pattern can be observed with other logographic languages, or whether other scripts are also able to modulate the RT pattern of SOA manipulation.

A brief discussion of the differences between production and comprehension modalities in relation to the bilingual lexical disadvantage is warranted here. Previous studies investigating bilingual lexical processing delays have mainly focused on production delays in the L1 (e.g. Gollan et al., 2005; Pyers et al., 2009) and comprehension delays in the L2 (e.g. Ardal et al., 1990; Moreno & Kutas, 2005; Newman et al., 2012). Recent evidence suggests that language processing delays associated with bilingualism are more robust in production than comprehension (e.g. Gollan et al., 2008, 2011; Runnqvist et al., 2011). For example, Gollan et al. (2011) speculated that bilingual processing delays have a larger impact on production than comprehension because production requires more exposures, or more lifetime practice, in order to reach high levels of proficiency. Therefore although the current data did not show evidence of an L1 delay in comprehension, an L1 delay may be elicited with a vocal color-naming task or an alternative language production paradigm. However, the source of this delay might be different from the delay found in comprehension, so more research is needed regarding how bilingualism differentially affects lexical delays in production and comprehension.

A limitation of the current study is the use of only five SOAs at 200 ms intervals. Only five SOAs were included due to length considerations for the experiment, but this may not have provided enough fine-grained measurement to identify differences in lexical access speeds between L1 and L2. Recent work by our group using a Stroop SOA task with electroencephalography (EEG) in Chinese–English bilinguals (Coderre et al., 2012) revealed that lexical processing is delayed by 100 ms in bilinguals' L2. Other studies using EEG have also provided compelling evidence for a temporal delay of L2 relative to L1 (Newman et al., 2012; see van Heuven & Dijkstra, 2010, for an overview). Therefore future research using more fine-grained SOA intervals is encouraged.

Conclusions

To summarize, the current study used an SOA manipulation of the Stroop task to investigate differences in Stroop performance between monolinguals and bilinguals in relation to factors of enhanced cognitive control (as proposed by the bilingual cognitive advantage hypothesis) and delayed lexical access (as proposed by the bilingual lexical disadvantage hypotheses). The bilingual cognitive advantage hypothesis predicted smaller Stroop interference for bilinguals in both their L1 and L2 compared to monolinguals. The Chinese–English bilinguals demonstrated reduced interference effects for both languages compared to monolinguals; however, no bilingual advantage was observed in the English–Chinese bilinguals. L2 proficiency also affected the magnitude of interference in the Chinese–English bilinguals. These results support the bilingual cognitive advantage hypothesis, highlighting its sensitivity to language proficiency and effects of language immersion.

The bilingual L1 lexical disadvantage hypothesis predicted delayed lexical access in the L1 compared to monolinguals, reflected as a negative shift in the peak interference effects. This hypothesis was not supported in either bilingual group. The bilingual L2 lexical disadvantage hypothesis further predicted delayed lexical access between L1 and L2, reflected as a negative shift in interference effects between L2 and L1. This pattern was found in the L1 Chinese vs. L2 English and the L2 Chinese vs. L1 Chinese. Thus the data support a delay in bilinguals' L2 relative to L1, but not in bilinguals' L1 relative to monolinguals.

In conclusion, the data support the L2 lexical disadvantage hypothesis and the bilingual cognitive advantage hypothesis, suggesting that both executive control abilities and lexical access speed contribute to bilingual Stroop performance. Overall, the data promotes future consideration of the multitude of factors, including script, choice of SOA, and language experience, that interact in determining bilingual performance on linguistic executive control tasks.

References

Appelbaum, L. G., Boehler, C., Won, R., Davis, L., & Woldorff, M. G. (2012). Strategic orientation of attention reduces temporally predictable stimulus conflict. *Journal of Cognitive Neuroscience, 24 (9),* 1834–1848.

Appelbaum, L. G., Meyerhoff, K. L., & Woldorff, M. G. (2009). Priming and backwards influences in the human brain: Processing interactions during the Stroop interference effect. *Cerebral Cortex, 19 (11),* 2508–2521.

Ardal, S., Donald, M. W., Meuter, R., Muldrew, S., & Luce, M. (1990). Brain responses to semantic incongruity in bilinguals. *Brain and Language, 39,* 187–205.

Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: Effects of response modality and processing domain. *Cerebral Cortex, 11,* 837–848.

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., Herron, D., Lu, C. C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., & Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review, 10 (2),* 344–380.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy and cognition*. New York: Cambridge University Press.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12 (1),* 3–11.

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19 (2),* 290–303.

Bialystok, E., Craik, F. I. M., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34 (4),* 859–873.

Bialystok, E., Craik, F. I. M., & Ruocco, A. C. (2006). Dual-modality monitoring in a classification task: The effects of bilingualism and ageing. *Quarterly Journal of Experimental Psychology, 59,* 1968–1983.

Bialystok, E., & Feng, X. (2009). Language proficiency and executive control in proactive interference: Evidence from monolingual and bilingual children and adults. *Brain and Language, 109 (2–3),* 93–100.

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science, 7 (3),* 325–339.

Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition, 112,* 494–500.

Biederman, I., & Tsao, Y. C. (1979). On processing Chinese ideographs and English words: Some implications from Stroop-test results. *Cognitive Psychology, 11,* 125–132.

Brauer, M. (1998). Stroop interference in bilinguals: The role of similarity between the two languages. In A. F. Healy & L. E. Bourne Jr. (eds.), *Foreign language learning: Psycholinguistic studies on training and retention*, pp. 317–337. Mahwah, NJ: Lawrence Erlbaum.

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science, 11 (2),* 282–298.

Chen, H. C., & Ho, C. (1986). Development of Stroop interference in Chinese–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12 (3),* 397–401.

Chua, F. K. (1999). Phonological recoding in Chinese logograph recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25 (4),* 876–891.

Coderre, E. L., Conklin, K., & van Heuven, W. J. B. (2012). The effects of bilingualism on the timecourse of early word recognition. Ms., University of Nottingham.

Coderre, E. L., & van Heuven, W. J. B. (2012). Modulations of the executive control network by stimulus onset asynchrony in a Stroop task. Ms., University of Nottingham.

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition, 113,* 135–149.

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

https://doi.org/10.1017/S1366728912000405 Published online by Cambridge University Press

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5 (3),* 175–197.

Dyer, F. N. (1973). Interference and facilitation for color naming with separate bilateral presentations of the word and color. *Journal of Experimental Psychology, 99,* 314–317.

Dyer, F. N. (1974). Stroop interference with long preexposures of the word: Comparison of pure and mixed preexposure sequences. *Bulletin of the Psychonomic Society, 3,* 8–10.

Fang, S. P., Tzeng, O. J. L., & Alva, L. (1981). Intralanguage vs. interlanguage Stroop effects in two types of writing systems. *Memory & Cognition, 9 (6),* 609–617.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123 (1),* 71–99.

Glaser, W. R., & Düngelhoff, F. J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 640–654.

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance, 8 (6),* 875–894.

Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General, 118,* 13–42.

Gollan, T. H., & Acenas, L.-A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish–English and Tagalog–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30 (1),* 246–269.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58 (3),* 787–814.

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition, 33 (7),* 1220–1234.

Gollan, T. H., Slattery, T. J., Goldenberg, D., van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General, 140 (2),* 186–209.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1,* 67–81.

Guo, T., Peng, D., & Liu, Y. (2005). The role of phonological activation in the visual semantic retrieval of Chinese characters. *Cognition, 98,* B21–B34.

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review, 18 (4),* 625–658.

Kaushanskaya, M., & Marian, V. (2007). Bilingual language processing and interference in bilinguals: Evidence from eye tracking and picture naming. *Language Learning, 57 (1),* 119–163.

Kousaie, S., & Phillips, N. A. (2012). Ageing and bilingualism: Absence of a "bilingual advantage" in Stroop interference in a nonimmigrant sample. *The Quarterly Journal of Experimental Psychology, 65 (2),* 356–369.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149–174.

Linck, J. A., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *Mental Lexicon, 3 (3),* 349–374.

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immerse in a second language: Evidence for the role of inhibition in second-language learning. *Psychological Science, 20,* 1507–1515.

Liu, Y., Hao, M., Li, P., & Shu, H. (2011). Timed picture naming norms for Mandarin Chinese. *PLoS One, 6 (1),* 1–7.

Liu, Y., Perfetti, C. A., & Hart, L. (2003). ERP evidence for the time course of graphic, phonoligical, and semantic information in Chinese meaning and pronunciation decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29 (6),* 1231–1247.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109 (2),* 163–203.

MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences, 4 (10),* 383–391.

Martin, C. D., Dering, B., Thomas, E. M., & Thierry, G. (2009). Brain potentials reveal semantic priming in both the "active" and the "non-attended" language of early bilinguals. *NeuroImage, 47,* 326–333.

Midgley, K., Holcomb, P. J., van Heuven, W. J. B., & Grainger, J. (2008). An electrophysiologival investigation of cross-language effects of orthographic neighborhood. *Brain Research, 1246,* 123–135.

Miller, N., & Kroll, J. F. (2002). Stroop effects in bilingual translation. *Memory & Cognition, 30 (4),* 614–628.

Moreno, E. M., & Kutas, M. (2005). Processing semantic anomalies in two languages: An electrophysiological exploration in both languages of Spanish–English bilinguals. *Cognitive Brain Research, 22,* 205–220.

Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. J., & Ullman, M. T. (2012). The influence of language proficiency on lexical semantic processing in native and late learners of English. *Journal of Cognitive Neuroscience, 24,* 1205–1223.

Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The Lexical Constituency Model: Some implications of research on Chinese for general theories of reading. *Psychological Review, 112 (1),* 43–59.

Perfetti, C. A., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 24–33.

Potter, M. C., So, K. F., von Eckhardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and more proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23,* 23–38.

Proverbio, A. M., Adorni, R., & Zani, A. (2009). Inferring native language from early bio-electrical activity. *Biological Psychology, 80,* 52–63.

Pyers, J. E., Gollan, T. H., & Emmorey, K. (2009). Bimodal bilinguals reveal the source of tip-of-the-tongue states. *Cognition, 113,* 323–329.

Ransdell, S. E., & Fischler, I. (1987). Memory in a monolingual mode: When are bilinguals at a disadvantage? *Journal of Memory and Language, 26,* 392–405.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review, 110 (1),* 88–125.

Roelofs, A. (2010). Attention, temporal predictability, and the time course of context effects in naming performance. *Acta Psychologia, 133,* 146–153.

Runnqvist, E., Strijkers, K., Sadat, J., & Costa, A. (2011). On the temporal and functional origin of L2 disadvantages in speech production: A critical review. *Frontiers in Psychology, 2,* 1–8.

Saalbach, H., & Stern, E. (2004). Differences between Chinese morphosyllabic and German alphabetic readers in the Stroop interference effect. *Psychonomic Bulletin & Review, 11 (4),* 709–715.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture–word interference studies. *Journal of Memory and Language, 29,* 86–102.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22 (2),* 461–479.

Spinks, J. A., Liu, Y., Perfetti, C. A., & Tan, L. H. (2000). Reading Chinese characters for meaning: The role of phonological information. *Cognition, 76,* B1–B11.

Taft, M., & van Graan, F. (1998). Lack of phonological mediation in a semantic categorization task. *Journal of Memory and Language, 38,* 203–224.

Tan, L. H., Laird, A. R., Li, K., & Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. *Human Brain Mapping, 25,* 83–91.

Tan, L. H., & Perfetti, C. A. (1997). Visual chinese character recognition: Does phonological information mediate access to meaning? *Journal of Memory and Language, 37,* 41–57.

Tao, L., Marzecová, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: The role of age of acquisition. *Frontiers in Psychology, 2,* 1–19.

Thornburgh, D. F., & Ryalls, J. H. (1998). Voice onset time in Spanish–English bilniguals: Early versus late learners of English. *Journal of Communication Disorders, 31,* 215–229.

Tsao, Y. C., Wu, M. F., & Feustel, T. (1981). Stroop interference: Hemispheric difference in Chinese speakers. *Brain and Language, 13,* 372–378.

Tzelgov, J., Henik, A., & Leiser, D. (1990). Controlling Stroop interference: Evidence from a bilingual task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16 (5),* 760–771.

van Heuven, W. J. B., Conklin, K., Coderre, E. L., Guo, T., & Dijkstra, T. (2011). The influence of cross-language similarity on within- and between-language Stroop effects in trilinguals. *Frontiers in Psychology, 2,* 1–16.

van Heuven, W. J. B., & Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Reseach Reviews, 64,* 104–122.

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39,* 458–483.

van Heuven, W. J. B., Schriefers, H., Dijkstra, T., & Hagoort, P. (2008). Language conflict in the bilingual brain. *Cerebral Cortex, 18,* 2706–2716.

Weaver Cargin, J., Maruff, P., Collie, A., Shafiq-Antonacci, R., & Masters, C. (2007). Decline in verbal memory in non-demented older adults. *Journal of Clinical and Experimental Neuropsychology, 29 (7),* 706–718.

Weekes, N. Y., & Zaidel, E. (1996). The effects of procedural variations on lateralized Stroop effects. *Brain and Cognition, 31,* 308–330.

Wright, B. C., & Wanley, A. (2003). Adults' versus children's performance on the Stroop task: Interference and facilitation. *British Journal of Psychology, 94,* 475–485.

Xu, Y., Pollatsek, A., & Potter, M. C. (1999). The activation of phonology during silent Chinese word reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25 (4),* 838–857.

Zelazo, P. D., Craik, F. I. M., & Booth, L. (2004). Executive function across the life span. *Acta Psychologica, 115,* 167–183.